

**Smith Washington Arauco
Canchumuni**

**Método Híbrido Baseado em Filtro
de Kalman e Modelos Generativos
de Aprendizagem Profunda no
Ajuste de Histórico sob Incertezas
para Modelos de Fácies Geológicas**

TESE DE DOUTORADO

**DEPARTAMENTO DE ENGENHARIA ELÉTRICA
Programa de Pós-graduação em Engenharia
Elétrica**

Rio de Janeiro
Setembro de 2017



Smith Washington Arauco Canchumuni

**Método Híbrido Baseado em Filtro de Kalman
e Modelos Generativos de Aprendizagem
Profunda no Ajuste de Histórico sob Incertezas
para Modelos de Fácies Geológicas**

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio.

Orientador : Prof. Marco Aurélio Cavalcanti Pacheco
Co-orientador: Prof. Alexandre Anozé Emerick

Rio de Janeiro
Setembro de 2017



Smith Washington Arauco Canchumuni

**Método Híbrido Baseado em Filtro de Kalman
e Modelos Generativos de Aprendizagem
Profunda no Ajuste de Histórico sob Incertezas
para Modelos de Fácies Geológicas**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Aurélio Cavalcanti Pacheco

Orientador

Departamento de Engenharia Elétrica– PUC-Rio

Prof. Alexandre Anozé Emerick

Co-orientador

Petróleo Brasileiro - Rio de Janeiro – Matriz

Prof. Marco Antonio Guimarães Dias

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Ruy Luiz Milidui

Departamento de Informática – PUC-Rio

Prof.^a Luciana Faletti Almeida

CEFET – RJ

Prof. Eugênio da Silva

UEZO

Prof. Leonardo Alfredo Forero Mendoza

UERJ

Farid Salomao Shecaira

Petrobras

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro Técnico Científico PUC-Rio

Rio de Janeiro, 6 de Setembro de 2017

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Smith Washington Arauco Canchumuni

Graduou-se em Engenharia Mecatrônica pela Universidade Nacional de Engenharia (UNI, Peru). Fez mestrado no Departamento de Engenharia Mecânica da PUC-Rio, especializando-se na área de mecânica aplicada, sistemas dinâmicos, robótica, métodos de otimização, métodos numéricos, desenvolvimento de sistemas de software utilizando diversos linguagens com experiência na área de Engenharia Mecânica e Inteligência Computacional.

Ficha Catalográfica

Arauco Canchumuni, Smith Washington

Método Híbrido Baseado em Filtro de Kalman e Modelos Generativos de Aprendizagem Profunda no Ajuste de Histórico sob Incertezas para Modelos de Fácies Geológicas / Smith Washington Arauco Canchumuni; orientador: Marco Aurélio Cavalcanti Pacheco; co-orientador: Alexandre Anozé Emerick. – Rio de Janeiro: PUC-Rio, Departamento de Engenharia Elétrica, 2017.

v., 169 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Ajuste de histórico. 3. Quantificação da incertezas. 4. Ensemble Kalman. 5. Aprendizado Profundo. 6. Atualização de fácies.
I. Cavalcanti Pacheco, Marco Aurélio. II. Anozé Emerick, Alexandre. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Agradecimentos

Em primeiro lugar, gostaria de agradecer especialmente ao meu orientador Prof. Marco Aurélio C. Pacheco e ao co-orientador Dr. Alexandre Emerick. Agradecimentos especiais ao meu co-orientador Alexandre, pela aceitação, auxílio, dedicação, paciência, revisões, sugestões, contribuições substanciais ao trabalho e por ter me guiado durante esta caminhada com seu extenso conhecimento.

Eu tive a sorte e a oportunidade de contar com meus competentes colegas do laboratório ICA que, sem sombra de dúvida, contribuíram significativamente para esta tese. Eles também foram os responsáveis por criar um ambiente agradável e estimulante para trabalhar durante esses meus quatro anos de pesquisa.

Também gostaria de agradecer aos meus pais, pelo amor, pela confiança, pelo apoio incondicional, pela compreensão nos momentos de ausência e pelo apoio à minha formação acadêmica, mesmo a muitos quilômetros de distância, que tem sido essencial para encontrar forças e continuar lutando pelos meus objetivos. Palavras não podem expressar a imensidão da gratidão que tenho por eles. À minha família, em geral, pela força e confiança.

Por último, mas não menos importante, gostaria de agradecer à CAPES pelo apoio financeiro e à PUC-Rio pela bolsa de isenção de mensalidades do doutorado.

Resumo

Arauco Canchumuni, Smith Washington; Cavalcanti Pacheco, Marco Aurélio; Anozé Emerick, Alexandre. **Método Híbrido Baseado em Filtro de Kalman e Modelos Generativos de Aprendizagem Profunda no Ajuste de Histórico sob Incertezas para Modelos de Fácies Geológicas**. Rio de Janeiro, 2017. 169p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Os métodos baseados no filtro de Kalman têm tido sucesso notável na indústria do petróleo nos últimos anos, especialmente, para resolver problemas reais de ajuste de histórico. No entanto, como a formulação desses métodos é baseada em hipóteses de gaussianidade e linearidade, seu desempenho é severamente degradado quando a geologia *a priori* é descrita em termos de distribuições complexas (e.g. modelos de fácies). A tendência atual em soluções para o problema de ajuste de histórico é levar em consideração modelos de reservatórios mais realistas com geologia complexa. Assim, a modelagem de fácies geológicas desempenha um papel importante na caracterização de reservatórios, como forma de reproduzir padrões importantes de heterogeneidade e facilitar a modelagem das propriedades petrofísicas das rochas do reservatório. Esta tese introduz uma nova metodologia para realizar o ajuste de histórico de modelos geológicos complexos. A metodologia consiste na integração de métodos baseados no filtro de Kalman em particular o método conhecido na literatura como *Ensemble Smoother with Multiple Data Assimilation* (ES-MDA), com uma parametrização das fácies geológicas por meio de técnicas baseadas em aprendizado profundo (*Deep Learning*) em arquiteturas do tipo autoencoder. Um autoencoder sempre consiste em duas partes, o codificador (modelo de reconhecimento) e o decodificador (modelo gerador). O procedimento começa com o treinamento de um conjunto de realizações de fácies por meio de algoritmos de aprendizado profundo, através do qual são identificadas as principais características das imagens de fácies geológicas, permitindo criar novas realizações com as mesmas características da base de treinamento com uma reduzida parametrização dos modelos de fácies na saída do codificador. Essa parametrização é regularizada no codificador para fornecer uma distribuição gaussiana na saída, a qual é utilizada para atualizar os modelos de fácies de acordo com os dados observados do reservatório, através do método ES-MDA. Ao final, os modelos atualizados são reconstruídos através do aprendizado profundo (decodificador), com o objetivo de obter modelos finais que apresentem características similares às da base de treinamento.

Os resultados, em três casos de estudo com 2 e 3 fácies, mostram que a parametrização de modelos de fácies baseada no aprendizado profundo consegue reconstruir os modelos de fácies com um erro inferior a 0,3%. A metodologia proposta gera modelos geológicos ajustados que conservam a descrição geológica *a priori* do reservatório (fácies com canais curvilíneos), além de ser consistente com o ajuste dos dados observados do reservatório.

Palavras-chave

Ajuste de histórico; Quantificação da incertezas; Ensemble Kalman;
Aprendizado Profundo; Atualização de fácies.

Abstract

Arauco Canchumuni, Smith Washington; Cavalcanti Pacheco, Marco Aurélio (Advisor); Anozé Emerick, Alexandre (Co-Advisor). **Hybrid Method Based into Kalman Filter and Deep Generative Model to History Matching and Uncertainty Quantification of Facies Geological Models**. Rio de Janeiro, 2017. 169p. Tese de doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Kalman filter-based methods have had remarkable success in the oil industry in recent years, especially to solve several real-life history matching problems. However, as the formulation of these methods is based on the assumptions of gaussianity and linearity, their performance is severely degraded when *a priori* geology is described in terms of complex distributions (e.g., facies models). The current trend in solutions for the history matching problem is to take into account more realistic reservoir models, with complex geology. Thus the geological facies modeling plays an important role in the characterization of reservoirs as a way of reproducing important patterns of heterogeneity and to facilitate the modeling of the reservoir rocks' petrophysical properties. This thesis introduces a new methodology to perform the history matching of complex geological models. This methodology consists of the integration of Kalman filter-based methods, particularly the method known in the literature as Ensemble Smoother with Multiple Data Assimilation (ES-MDA), with a parameterization of the geological facies through techniques based on deep learning in autoencoder type architectures. An autoencoder always consists of two parts, the encoder (recognition model) and the decoder (generator model). The procedure begins with the training of a set of facies realizations via deep generative models, through which the main characteristics of geological facies images are identified, allowing for the creation of new realizations with the same characteristics of the training base, with a low dimension parametrization of the facies models at the output of the encoder. This parameterization is regularized at the encoder to provide Gaussian distribution models in the output, which is then used to update the models according to the observed data of the reservoir through the ES-MDA method. In the end, the updated models are reconstructed through deep learning (decoder), with the objective of obtaining final models that present characteristics similar to those of the training base.

The results, in three case studies with 2 and 3 facies, show that the parameterization of facies models based on deep learning can reconstruct facies models with an error lower than 0.3%. The proposed methodology generates final geological models that preserve *the a priori* geological description of the reservoir (facies with curvilinear channels), besides being consistent with the adjustment of the observed data of the reservoir.

Keywords

History matching; Uncertainty quantification; Ensemble Kalman;
Deep learning; Facies update.

Sumário

1	Introdução	25
1.1	Objetivos	30
1.2	Objetivos Específicos	30
1.3	Resumo dos Capítulos	31
2	Revisão Bibliográfica e Fundamentos Teóricos	33
2.1	Ajuste de Histórico	33
2.2	Métodos Baseados no Filtro de Kalman	36
2.2.1	Filtro de Kalman	36
2.2.2	Extended Kalman Filter	38
2.2.3	Ensemble Kalman Filter	38
2.2.4	Ensemble Smoother	40
2.2.5	ES-MDA	41
2.2.6	IAGS	43
2.2.7	GMMDA	44
2.2.8	Localização do Ganho de Kalman	46
2.2.9	Aplicações dos Métodos Baseados no Filtro de Kalman no Ajuste de Histórico	47
2.2.10	Atualização de Modelos de Fácies	49
2.3	Análise de Componentes Principais	49
2.3.1	Representação de Modelos Geológicos Mediante a Análise de Componentes Principais	50
2.3.2	Parametrização OPCA	52
2.4	Geoestatística	53
2.4.1	Simulações Geoestatísticas	54
2.4.2	Simulação Gaussiana Truncada	54

2.4.3	Simulação Plurigaussiana Truncada	55
2.4.4	Geoestatística Multiponto	56
3	Estudo Comparativo de Métodos Baseados no Filtro de Kalman	58
3.1	Caso de Estudo 1D	59
3.1.1	Resultados	60
3.2	Caso de Estudo PUNQ-S3	66
3.2.1	Resultados	67
3.3	Caso de Estudo BRUGGE	70
3.3.1	Resultados	70
4	Metodologia de Atualização de Fácies Baseado em Análise de Componentes Principais	77
4.1	Procedimento	77
4.2	Aplicação no Modelo PUNQ-S3	81
4.2.1	ES-MDA	84
4.2.2	ES-MDA-PCA	85
5	Deep Learning	90
5.1	Modelos Generativos Profundos	91
5.1.1	Autoencoder	92
5.1.2	Variações do Autoencoder	94
5.1.2.1	Denoising Autoencoder	95
5.1.2.2	Sparse Autoencoder	96
5.1.2.3	Contractive Autoencoder	97
5.1.3	Restricted Boltzmann Machine	97
5.1.3.1	Estrutura de RBM	97
5.1.3.2	Divergência Contrastiva	100
5.1.3.3	Generalizações do RBM	101
5.1.4	Deep Belief Networks	103
5.1.4.1	Treinamento de DBN	103

5.1.4.2	DBN como Autoencoder	105
5.2	Treinamento de Modelos de Fácies	106
5.2.1	Arquitetura das Redes Utilizadas	107
5.2.2	Resultados	108
5.2.2.1	PUNQ-S3	109
5.2.2.2	MOD-3F	113
5.2.2.3	MOD45	114
5.2.2.4	MOD100	115
5.2.2.5	BRUGGE	116
5.2.2.6	Erro de Reconstrução	117
6	Metodologia de Atualização de Fácies Baseado em Deep Learning	120
6.1	Procedimento	120
6.1.1	Etapa de Aprendizado	121
6.1.2	Etapa de Representação	122
6.1.3	Etapa de Atualização	123
6.2	Casos de Estudo	124
6.2.1	PUNQ-S3	125
6.2.1.1	ES-MDA-DL	126
6.2.2	MPS45	130
6.2.2.1	ES-MDA	131
6.2.2.2	ES-MDA-DL	132
6.2.2.3	ES-MDA-OPCA	134
6.2.3	MPS100	136
6.2.3.1	ES-MDA	138
6.2.3.2	ES-MDA-DL	140
6.2.3.3	ES-MDA-OPCA	141
6.2.4	BRUGGE	144
6.2.4.1	ES-MDA	145
6.2.4.2	ES-MDA-DL	146

6.2.4.3 ES-MDA-OPCA	148
7 Conclusões e Recomendações	156
7.1 Conclusões	156
7.2 Recomendações para Trabalhos Futuros	158
Referências bibliográficas	159

Lista de figuras

Figura 1 - Fluxo de trabalho do ajuste de histórico considerando a geologia do reservatório.	26
Figura 2 - Modelo de um reservatório com quatro fácies, cada fácies é representada por uma cor diferente, na qual apresenta propriedades petrofísicas semelhantes.	28
Figura 3 - Ajuste de Histórico e previsão da Produção.	33
Figura 4 - Procedimento de simulação gaussiana truncada para geração de fácies: na parte acima, à esquerda, o campo gaussiano; na direita, a correspondente representação da fácies para o campo gaussiano; na parte inferior, o histograma com os respectivos cortes para gerar as fácies.	55
Figura 5 - Procedimento de simulação gaussiana para geração de três fácies: na parte de cima os 2 campos gaussianos, na parte inferior esquerda a correspondente representação de fácies para dos campos gaussianos e na parte de baixo, à direita, o histograma com os respectivos cortes para gerar as fácies com a respectiva RTR.	56
Figura 6 - Modolo do caso “1D” [53].	59
Figura 7 - Distribuição <i>a posteriori</i> do logaritmo natural da permeabilidade do caso de estudo “1D” obtido por MCMC [53].	60
Figura 8 - Distribuição <i>a posteriori</i> do logaritmo natural da permeabilidade após o procedimento de ajuste de histórico com 10 iterações, para o caso de estudo “1D”.	61
Figura 9 - Distribuição <i>a posteriori</i> da vazão de produção de água (Q_w) ao final do ajuste de histórico com 10 iterações, para o caso de estudo “1D”.	62

- Figura 10 - Distribuição *a posteriori* da produção de água acumulada (W_P) ao final do ajuste de histórico com 10 iterações, para o caso de estudo “1D”. 62
- Figura 11 - Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de ES-MDA, para o caso de estudo “1D”. 64
- Figura 12 - Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de GMMDA, para o caso de estudo “1D”. 64
- Figura 13 - Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de IAGS, para o caso de estudo “1D”. 65
- Figura 14 - Diagrama de caixas da função objetivo das experiências realizadas no caso de estudo “1D”. 65
- Figura 15 - Diferença absoluta do valor do logaritmo natural da permeabilidade com o valor real e com o valor de referência (MCMC), para todas as experiências realizadas no caso de estudo “1D”. 66
- Figura 16 - Modelos do logaritmo natural da permeabilidade horizontal da primeira realização do primeiro conjunto, para o caso de estudo PUNQ-S3. 68
- Figura 17 - Modelos do logaritmo natural da permeabilidade horizontal da realização 33 do primeiro conjunto, para o caso de estudo PUNQ-S3. 69
- Figura 18 - Diagrama de caixas da função objetivo do caso de estudo PUNQ-S3. 70
- Figura 19 - Curvas de produção dos modelos iniciais (cinza) e finais (azul). A coluna da esquerda apresenta a pressão no poço produtor P1. A coluna da direita apresenta a razão gás-óleo (RGO) do poço produtor P5, para o caso de estudo PUNQ-S3. 71

Figura 20 - Diagrama de caixas da função objetivo para o caso de estudo BRUGGE mediante o método GMDA variando o parâmetro de controle h .	72
Figura 21 - Modelo de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE, correspondente ao membro número 1 do conjunto.	72
Figura 22 - Modelo de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE, correspondente ao membro número 33 do conjunto.	73
Figura 23 - Modelos de desvio padrão de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE.	73
Figura 24 - Diagrama de caixas da função objetivo dos três métodos estudados, para o caso de estudo BRUGGE.	73
Figura 25 - Curvas de pressão dos modelos iniciais (cinza) e finais (azul) para os poços P10 e P15, para o caso de estudo BRUGGE.	74
Figura 26 - Curvas de produção dos modelos iniciais (cinza) e finais (azul). A coluna da esquerda apresenta a vazão de água (Q_w) do poço P5. A coluna da direita apresenta a vazão de óleo (Q_o) do poço P14, para o caso de estudo BRUGGE.	75
Figura 27 - Procedimento do método ES-MDA-PCA.	78
Figura 28 - Partição de um modelo com três fácies.	78
Figura 29 - Simulação sequencial gaussiana condicionada à fácies binárias.	79
Figura 30 - Modelo de fácies de referência na linha superior, a linha central o logaritmo da permeabilidade e a linha inferior os histogramas correspondentes.	82
Figura 31 - Curvas de produção do produtor 11 (curvas azuis). Histórico de produção (asteriscos vermelhos).	83

- Figura 32 - Primeiras três realizações de modelos de fácies, logaritmo da permeabilidade horizontal e seus histogramas do conjunto *a priori*, para o caso de estudo PUNQ-S3 modificado. 85
- Figura 33 - Primeiras três realizações do logaritmo da permeabilidade horizontal do conjunto final (*a posteriori*) utilizando ES-MDA, para o caso de estudo PUNQ-S3 modificado. 86
- Figura 34 - Curvas de produção resultantes do ES-MDA aplicado ao caso de estudo PUNQ-S3 modificado. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11. 86
- Figura 35 - Primeiras três realizações de modelos de fácies, logaritmo da permeabilidade horizontal e seus histogramas do conjunto *a posteriori* utilizando ES-MDA-PCA, para o caso de estudo PUNQ-S3 modificado. 87
- Figura 36 - Curvas de produção resultantes do ES-MDA-PCA aplicado ao caso de estudo PUNQ-S3 modificado. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11. 88
- Figura 37 - Estrutura geral de uma rede AE, mapeando uma entrada \mathbf{x} a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} . A rede AE apresenta duas componentes: o *encoder* $f(\mathbf{x})$ e o *decoder* $g(\mathbf{h})$. 91
- Figura 38 - Funções de ativação comumente utilizadas, $\phi(x)$. 93
- Figura 39 - Estrutura de uma arquitetura AE com 3 camadas ocultas, mapeando uma entrada \mathbf{x} a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} (representação latente). 94
- Figura 40 - Estrutura geral de uma rede DAE, onde é acrescentado um ruído na entrada \mathbf{x} e mapeando a entrada ruidosa $\tilde{\mathbf{x}}$ a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} . 95
- Figura 41 - Representação gráfica de uma RBM com três unidades visíveis e duas ocultas. 98

Figura 42 - Esquerda: Uma rede RBM. Direita: Uma rede DBN com 2 camadas ocultas, com pesos atados $\mathbf{W}^2 = \mathbf{W}^{1T}$.	105
Figura 43 - Representação gráfica de uma rede DBN (direita) de 4 camadas escondidas, acoplando 4 redes RBMs (esquerda).	105
Figura 44 - Representação gráfica de uma RBM com três unidades visíveis e duas ocultas.	106
Figura 45 - Arquitetura da rede AE para o treinamento do caso de estudo PUNQ-S3.	109
Figura 46 - Arquitetura da Rede RBM para o treinamento do caso de estudo PUNQ-S3.	109
Figura 47 - Arquitetura da rede DBN como AE para o treinamento do caso de estudo PUNQ-S3.	110
Figura 48 - Imagens de fâcies de treinamento do caso de estudo PUNQ-S3.	110
Figura 49 - Imagens de fâcies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede AE.	111
Figura 50 - Imagens de fâcies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede RBM.	111
Figura 51 - Imagens de fâcies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede DBN como AE.	112
Figura 52 - Imagens de fâcies da base de treinamento do caso de estudo MOD-3F.	113
Figura 53 - Imagens fâcies da base de teste do caso de estudo MOD-3F, utilizando uma rede DBN como AE.	114
Figura 54 - Imagens de fâcies da base de teste do caso de estudo MOD45, utilizando uma rede DBN como AE.	115
Figura 55 - Imagens de fâcies da base de teste do caso de estudo MOD100, utilizando uma rede DBN como AE.	116
Figura 56 - Reconstrução de duas imagens de fâcies do caso de estudo MPS100, usando uma arquitetura AE.	116

Figura 57 - Imagens de fácies de teste do caso de estudo BRUGGE, utilizando uma rede DBN como AE.	117
Figura 58 - Procedimento do método ES-MDA-DL.	121
Figura 59 - Arquitetura geral do aprendizado profundo de modelos de fácies geológicas.	122
Figura 60 - Acoplamento de permeabilidade em cada fácies a um modelo de permeabilidade geral de acordo com a imagem de fácies (esquerda).	124
Figura 61 - Primeiras três realizações do conjunto inicial (<i>a priori</i>) do modelo PUNQ-S3.	126
Figura 62 - Primeiras três realizações do conjunto final (<i>a posteriori</i>) do modelo PUNQ-S3, utilizando ES-MDA-DL.	127
Figura 63 - Curvas de produção resultantes do método ES-MDA-DL aplicado ao reservatório PUNQ-S3. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11.	128
Figura 64 - Função objetivo dos métodos estudados, para o caso de estudo PUNQ-S3.	129
Figura 65 - Modelo de permeabilidade de referência (parte de cima) e as primeiras cinco realizações do conjunto inicial (<i>a priori</i>) para o caso de estudo MPS45.	130
Figura 66 - Primeiras cinco realizações de mapas de permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo MPS45 aplicando o método ES-MDA.	131
Figura 67 - Curvas de produção resultantes do método ES-MDA aplicado ao caso de estudo MPS45. Vazão de produção de óleo e água, para os poços produtores 1, 3 e 4.	132
Figura 68 - Primeiras cinco realizações de mapas de permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo MPS45 aplicando o método ES-MDA-DL.	133

Figura 69 - Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo MPS45. Taxa de produção de óleo e água para os poços produtores 1, 3 e 4.	133
Figura 70 - Média e as primeiras cinco realizações de mapas de permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo MPS45 aplicando o método ES-MDA-OPCA.	134
Figura 71 - Curvas de produção resultantes do método ES-MDA-OPCA aplicado ao caso de estudo MPS45. Taxa de produção de óleo e água para os poços produtores 1, 3 e 4.	135
Figura 72 - Função objetivo dos métodos estudados, para o caso de estudo MPS45.	135
Figura 73 - Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo MPS45.	136
Figura 74 - Mapa de logaritmo natural da permeabilidade de referência para o caso de estudo MPS100.	137
Figura 75 - Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto inicial (<i>a priori</i>) para o caso de estudo MPS100.	138
Figura 76 - Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final do caso de estudo MPS1000 aplicando o método ES-MDA.	138
Figura 77 - Curvas de produção do método ES-MDA aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.	139
Figura 78 - Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo MPS1000 aplicando o método ES-MDA-DL.	140
Figura 79 - Curvas de produção do método ES-MDA-DL aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.	141

Figura 80 - Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo MPS1000 aplicando o método ES-MDA-OPCA.	142
Figura 81 - Função objetivo dos métodos estudados, para o caso de estudo MPS100.	142
Figura 82 - Curvas de produção do método ES-MDA-OPCA aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.	143
Figura 83 - Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo MPS100.	144
Figura 84 - Mapa de logaritmo natural da permeabilidade de referência para o caso de estudo BRUGGE.	144
Figura 85 - Primeiras três realizações do logaritmo natural da permeabilidade do conjunto inicial (<i>a priori</i>) para o caso de estudo BRUGGE.	145
Figura 86 - Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) para caso de estudo BRUGGE aplicando o método ES-MDA.	146
Figura 87 - Curvas de produção resultantes do método ES-MDA aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.	147
Figura 88 - Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) para caso de estudo BRUGGE aplicando o método ES-MDA-DL.	148
Figura 89 - Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.	149
Figura 90 - Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) para caso de estudo BRUGGE aplicando o método ES-MDA-OPCA.	150

Figura 91 - Curvas de produção resultantes do método ES-MDA-OPCA aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.	151
Figura 92 - Função objetivo dos métodos estudados, para o caso de estudo BRUGGE.	151
Figura 93 - Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo BRUGGE.	152
Figura 94 - Influência do número de assimilações para o caso de estudo BRUGGE com o método ES-MDA-DL.	152
Figura 95 - Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (<i>a posteriori</i>) do caso de estudo BRUGGE aplicando o método ES-MDA-DL, com um número de 20 assimilações.	153
Figura 96 - Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo BRUGGE, com um número de 20 assimilações. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.	154

Lista de tabelas

Tabela 1 - Fluxo de Trabalho do ES-MDA	42
Tabela 2 - Fluxo de Trabalho do IAGS	45
Tabela 3 - Fluxo de Trabalho do GMMDA	46
Tabela 4 - Casos de Estudo para os Métodos Baseados no ES	59
Tabela 5 - Variação de h em GMMDA para o Caso de Estudo “1D”	63
Tabela 6 - Variação de h em GMMDA para o Caso de Estudo PUNQ-S3	67
Tabela 7 - PUNQ-S3: A Média e o Desvio Padrão das Propriedades Petrofísicas por Tipo de Fácies	82
Tabela 8 - Ruído para as Observações	84
Tabela 9 - Resumo das Bases de Treinamento	107
Tabela 10 - Configuração das Redes AEs para Modelos de Fácies	108
Tabela 11 - Tempo de Treinamento do Caso de Estudo PUNQ-S3 em CPU vs GPU	112
Tabela 12 - Erro de Reconstrução das Redes AEs para Modelos de Fácies	118
Tabela 13 - Casos de Estudo para os testes do Método ES-MDA-DL	125
Tabela 14 - Parâmetros Gerais para os Testes do Método ES-MDA-DL	125

Lista de Abreviaturas

Abreviaturas em Inglês

EKF	– <i>Extended Kalman Filter</i>
EnKF	– <i>Ensemble Kalman Filter</i>
ES	– <i>Ensemble Smoother</i>
ES-MDA	– <i>Ensemble Smoother with Multiple Data Assimilation</i>
ES-MDA-DL	– <i>Ensemble Smoother with Multiple Data Assimilation and Deep Learning</i>
ES-MDA-PCA	– <i>Ensemble Smoother with Multiple Data Assimilation and Principal Component Analysis</i>
ES-MDA-OPCA	– <i>Ensemble Smoother with Multiple Data Assimilation and Optimization-Based Principal Component Analysis</i>
RML	– <i>Randomized Maximum Likelihood</i>
IS	– <i>Importance Sampling</i>
IAGS	– <i>Iterative Adaptive Gaussian Mixture</i>
GMMDA	– <i>Gaussian Mixture with Multiple Data Assimilation</i>
POD	– <i>Proper Orthogonal Decomposition</i>
PCA	– <i>Principal Component Analysis</i>
KPCA	– <i>Kernel-Principal Component Analysis</i>
OPCA	– <i>Optimization-Based Principal Component Analysis</i>
CDF-PCA	– <i>Cumulative Distribution Function Principal Component Analysis</i>
Pluri-PCA	– <i>Plurigaussian-Principal Component Analysis</i>
RTR	– <i>Rock Type Rule</i>
SGeMS	– <i>Stanford Geostatistical Modeling Software</i>
CMG	– <i>Computer Modelling Group</i>
HMMs	– <i>Hidden Markov Models</i>

DL	– <i>Deep Learning</i>
ML	– <i>Machine Learning</i>
RBM	– <i>Restricted Boltzmann Machines</i>
GRBM	– <i>Restricted Boltzmann Machines Gaussian</i>
SGD	– <i>Stochastic Gradient Descent</i>
DBN	– <i>Deep Belief Networks</i>
AE	– <i>Autoencoder</i>
DAE	– <i>Denoising Autoencoder</i>
SAE	– <i>Sparse Autoencoder</i>
CAE	– <i>Contractive Autoencoder</i>
MAP	– <i>Maximum A Posteriori</i> (Estimativa máximo <i>a posteriori</i>)
SVD	– <i>Singular Value Decomposition</i> (Decomposição em Valores Singulares)
MPS	– <i>Multiple Point Statistics</i> (Estatística de Multiponto)
MCMC	– <i>Markov Chain Monte Carlo</i> (Monte Carlo via cadeias de Markov)

Abreviaturas em Português

E&P	– Exploração e Produção
AHA	– Ajuste de Histórico Assistido
FO	– Função Objetivo
CPs	– Componentes Principais
IA	– Inteligência Artificial
DC	– Divergência Contrastiva
FDP	– Função Densidade de Probabilidade
RGO	– Razão Gás-Óleo

1

Introdução

Na área de Exploração e Produção (E&P) da indústria de petróleo, um grande desafio para os engenheiros é prever o comportamento dos reservatórios durante a produção. Normalmente essa previsão é feita através de modelos numéricos conhecidos como “simuladores de reservatório”. A partir da simulação, os engenheiros podem prever o desempenho do reservatório sob diversas condições operacionais, permitindo assim a otimização do desenvolvimento e gerenciamento do campo. Portanto, é imprescindível a construção de modelos confiáveis. A construção desses modelos envolve o trabalho de diversos geocientistas e engenheiros que buscam integrar todas as informações disponíveis do campo para caracterizar as propriedades da rocha e fluido do reservatório. Contudo, mesmo com a disponibilidade de técnicas modernas de caracterização, a quantidade de informações disponíveis sobre o campo é insuficiente. Além disso, os dados costumam conter ruído e podem, às vezes, até ser inconsistentes. Como resultado, existe uma incerteza significativa nas propriedades da rocha e fluidos usadas na construção dos modelos. Entretanto, com o início da produção do campo novas informações sobre o reservatório vão sendo reveladas, através da medição das vazões de produção de fluidos, medições de pressão e até novas aquisições sísmicas. O processo de calibração dos modelos de simulação para honrar dados oriundos da produção é conhecido na indústria como “ajuste de histórico”. A Figura 1 ilustra o processo de ajuste de histórico.

Matematicamente, o ajuste de histórico pode ser classificado como um problema inverso cujo objetivo é estimar parâmetros incertos do modelo, a partir de medições ruidosas das suas variáveis de saída. Tipicamente o número de dados disponíveis é muito menor que o número de parâmetros do modelo, o que torna o ajuste de histórico um problema mal posto onde diversos modelos distintos são capazes de reproduzir o histórico dentro de uma mesma tolerância. Tradicionalmente o ajuste de histórico é feito por engenheiros através de um processo lento e tedioso de “tentativa-e-erro”. Contudo, está se popularizando o uso de programas computacionais que propõem automatizar esse processo. Essas ferramentas são normalmente conhecidas como “ajuste de histórico assistido” (AHA).

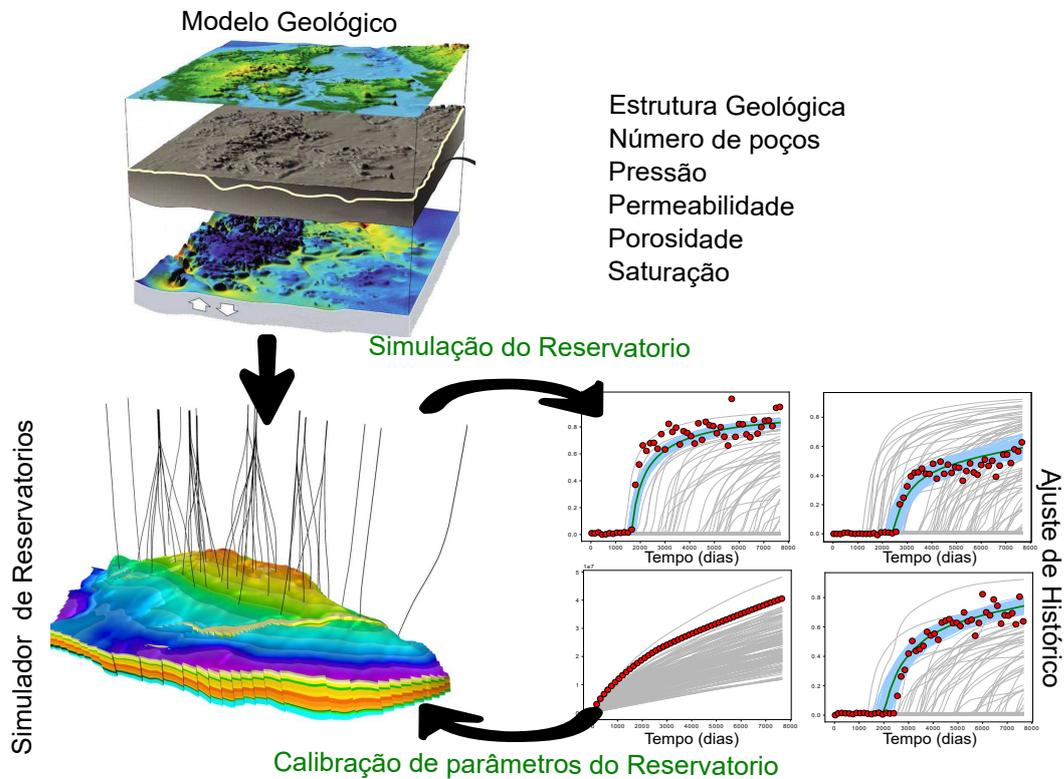


Figura 1: Fluxo de trabalho do ajuste de histórico considerando a geologia do reservatório.

A literatura técnicas de AHA é vasta, com os primeiros trabalhos na década de 60 [1, 2]. A maioria dessas técnicas formulam o ajuste de histórico como um problema de otimização, o qual envolve a minimização de uma função objetivo que quantifica a distância entre os dados observados no campo e os dados simulados pelo modelo. Entretanto, as técnicas mais modernas de AHA se baseiam em uma formulação Bayesiana. Nessa formulação, o ajuste de histórico é visto como um processo de redução de incertezas através de incorporação de dados. A incerteza é modelada através de distribuições de probabilidade. O processo se inicia com uma distribuição *a priori*, que é atualizada a partir da combinação com uma função de verossimilhança para dar origem a uma distribuição dita *a posteriori*. Normalmente nesses casos o problema inverso é resolvido com o uso de técnicas do tipo Monte Carlo.

Um ajuste de histórico bem sucedido deve, além de gerar modelos consistentes com os dados observados, permitir quantificar as incertezas nas previsões. Um conjunto de técnicas Bayesianas que têm sido aplicadas com

relativo sucesso em problemas de ajuste de histórico são os métodos baseados no filtro de Kalman [3]. Dentre esses métodos, o *ensemble Kalman filter*¹ (EnKF) é o mais conhecido. Contudo existe uma quantidade enorme de variações desse método; ver, por exemplo, [4, 5, 6] e referências citadas nestes.

O EnKF é um método de assimilação sequencial de dados proposto por Evensen [7] como uma versão estocástica (Monte Carlo) do filtro de Kalman. O EnKF e suas variantes têm sido objeto de intensa investigação e aplicações em diversas áreas como, por exemplo, previsão do tempo [8], hidrologia [9], oceanografia [10] e engenharia de petróleo [5, 6]. Entre as vantagens normalmente atribuídas a esses métodos cabe destaque a simplicidade de implementação computacional e a capacidade de tratar problemas com uma quantidade muito grande de parâmetros a um custo computacional relativamente baixo. Por exemplo, na literatura são reportadas aplicações diversas variando na ordem de 10^0 a 10^7 parâmetros incertos.

Contudo, esses métodos possuem deficiências bem documentadas na literatura que podem limitar a sua aplicação em casos mais complexos. As deficiências desses métodos estão em geral associadas a dois aspectos: (i) hipóteses de gaussianidade e linearidade usadas na construção dos métodos; (ii) uso de conjuntos relativamente pequenos de modelos para fazer a assimilação. O primeiro problema decorre do fato de que apesar do EnKF ter sido proposto como um método de assimilação de dados em sistemas não lineares, as hipóteses de linearidade e gaussianidade estão implícitas na formulação desse método. Na prática observa-se que o desempenho do método vai sendo deteriorado à medida que essas hipóteses vão sendo violadas. O problema de não linearidade é normalmente contornado com o uso de formulações iterativas. O uso de iterações é particularmente importante no caso da formulação com assimilação simultânea de dados. Essas formulações são normalmente conhecidas na literatura como *Ensemble Smoother* (ES) [11, 12].

Nos últimos anos foram desenvolvidas diversas versões iterativas do ES. A razão para esse crescente interesse no ES está relacionado ao fato de que esse método aparenta ser mais robusto e eficiente que o EnKF para o ajuste de histórico. De fato, o esquema de assimilação sequencial do EnKF tende a ser problemático para ajuste de histórico pois exige que as variáveis primárias do simulador, tipicamente pressão e saturação de fluidos, sejam atualizadas em conjunto com os parâmetros estáticos. Contudo, devido ao fato da atualização

¹Ainda não existe uma tradução bem estabelecida em língua portuguesa para o termo *ensemble Kalman filter*. Uma tradução literal seria algo como “filtro de Kalman por conjuntos”. Por essa razão, e também pelo fato de simplificar as referências a outros trabalhos, decidiu-se adotar no presente texto o nome original em língua inglesa dos métodos sempre que não existir uma tradução bem estabelecida em português.

ser linear, podem surgir inconsistências entre parâmetros e variáveis primárias que vão se propagando ao longo do processo de ajuste, levando a uma perda de qualidade nos resultados. Esse problema é conhecido na literatura como inconsistência parâmetro-estado [13]. Além disso, frequentemente a atualização das variáveis primárias gera problemas de convergência nas simulações, o que evidentemente onera o processo. Como o ES não utiliza assimilação sequencial, esse problema é completamente eliminado. Contudo, para que o ES apresente bons resultados é indispensável o uso de um esquema iterativo. Dentre os ES iterativos propostos na literatura cabe citar os trabalhos de [14, 15, 16, 17, 18]. Dentre esses métodos, um que tem recebido bastante atenção é o ES-MDA [16]. Esse método aparece inclusive em uma ferramenta comercial de ajuste de histórico [19]. Recentemente, Stordal e Elsheikh [17] combinaram o método ES-MDA com *Importance Sampling* (IS) dando origem ao método *Gaussian Mixture with Multiple Data Assimilation* (GMMDA). Em teoria esse método relaxa um pouco a exigência de gaussianidade implícita na construção do ES-MDA, o que poderia levar a um desempenho melhor em problemas não-gaussianos. Isso foi ilustrado pelos autores em um exemplo com dimensões bastante reduzidas [17, 18]. Entretanto, não está claro se esse método é de fato superior ao ES-MDA quando aplicado a problemas realistas de ajuste de histórico.

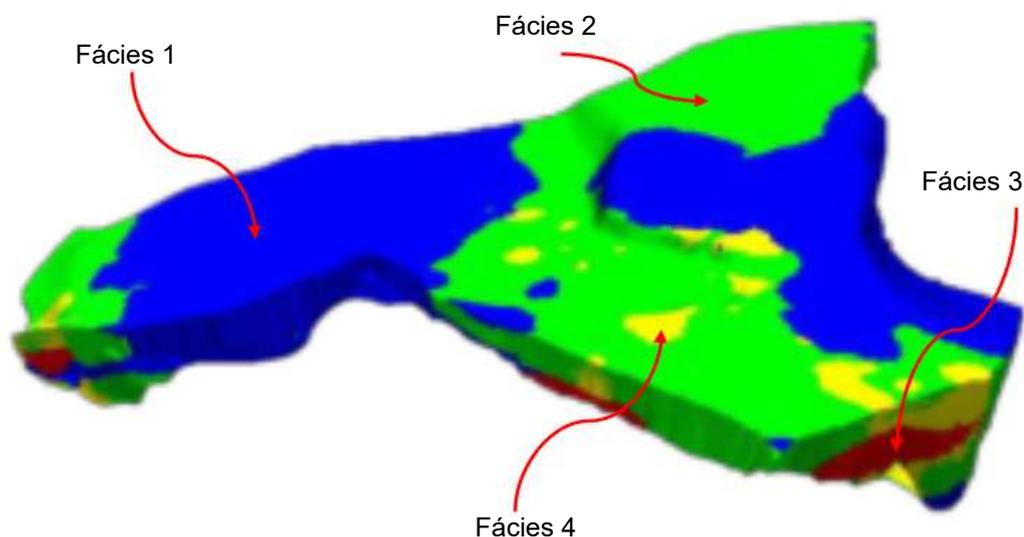


Figura 2: Modelo de um reservatório com quatro fácies, cada fácies é representada por uma cor diferente, na qual apresenta propriedades petrofísicas semelhantes.

Problemas associados com a não gaussianidade são particularmente evidentes quando esses métodos são aplicados a modelos geológicos mais complexos com geometrias deposicionais curvilíneas, como, por exemplo, canais

fluviais. Isso também é evidente quando o modelo geológico é composto por fácies. As fácies (Figura 2) são regiões onde o geólogo assume que as propriedades da rocha possuem características semelhantes. O desafio na atualização de modelos de fácies decorre do fato que essas são tipicamente variáveis categóricas, ou seja, variáveis que assumem categorias como valor, por exemplo, “areia”, “folhelho”, “interlaminado”, etc. Dessa forma é evidente que a distribuição *a priori* de fácies no modelo se afasta bastante de uma distribuição gaussiana. É importante observar que o ajuste de histórico de modelos de fácies é um desafio tecnológico para a maioria dos métodos de AHA e não apenas para os métodos baseados no filtro de Kalman. De fato, isto leva a utilizar abordagens que possam complementar os métodos baseados no filtro de Kalman para lidar com suas deficiências. Uma das abordagens muito utilizadas na literatura é a parametrização, não somente para reduzir a dimensão dos parâmetros do modelo, como também para conseguir uma transformação de variáveis categóricas em um novo espaço contínuo, preferencialmente um espaço onde a variável apresente uma distribuição aproximadamente gaussiana.

Parametrizações baseadas em análise de componentes principais (*Principal Component Analysis*, PCA [20]) têm tido aplicações com relativo sucesso no problema de ajuste de histórico. Embora ainda não exista metodologia que consiga parametrizar a grande complexidade dos modelos de fácies geológicas, já existem avanços nessa área. Por exemplo, Emerick [21] apresentou um estudo comparativo de diferentes métodos de parametrização baseada em PCA em conjunto com ES-MDA aplicados ao problema de ajuste de histórico de fácies geológicas.

Por outro lado, *Deep Learning* (DL)² é uma nova linha de pesquisa dentro da área de *Machine Learning* (ML), que foi introduzida com o objetivo de mover a aprendizagem de máquinas para mais perto de modelar problemas originais [22]. DL busca imitar a atividade das diferentes camadas de neurônios no córtex cerebral. Apesar dos algoritmos que podem simular a grande variedade de neurônios no neocórtex em uma rede neural artificial terem décadas de idade, devido a melhorias nas fórmulas matemáticas e o uso de computadores cada vez mais poderosos, cientistas agora podem modelar mais camadas de neurônios virtuais do que antes. A arquitetura das técnicas de DL permite criar modelos computacionais compostos de múltiplas camadas de processamento, para aprender automaticamente representações de dados com múltiplos níveis de abstração. Assim, o sistema consegue aprender funções complexas, mape-

²O termo *deep learning* pode ser traduzido para o português como “aprendizado profundo”. Contudo, o nome original em inglês é bastante difundido na literatura técnica. Por essa razão, decidiu-se usar o termo *deep learning* (DL) nessa tese. O mesmo vale para outros termos como, por exemplo, *machine learning* e *principal component analysis*.

ando a entrada para a saída diretamente a partir dos dados de treinamento, mediante um aprendizado não supervisionado. Isto é especialmente importante para aprender abstrações de nível superior [23].

Diversas arquiteturas de DL têm sido aplicadas com sucesso em diversos campos de pesquisa, como, por exemplo, reconhecimento de voz [24], a percepção de objetos [25], tradução automática [26], visão computacional [27], processamento de linguagem natural [28], dentre outros (todas áreas especialmente complexas para pesquisadores em inteligência artificial (IA) [29]). De fato, DL é um tema que está se tornando cada vez mais importante no campo da IA e rapidamente se tornou um dos campos mais solicitados para pesquisa [30].

Embora não existam aplicações diretas de DL no problema de ajuste de histórico, cabe tentar construir uma parametrização puramente a partir de dados que possam ser mais aptos a lidar com grandes não-linearidades e fortes restrições. Assim, a parametrização baseada em técnicas de DL pode ser útil para transformar imagens de modelos de fácies em um novo sistema, no qual esses novos parâmetros apresentam uma distribuição aproximadamente gaussiana.

1.1 Objetivos

O principal objetivo desse trabalho é o desenvolvimento de uma metodologia robusta para o ajuste de histórico de modelos de reservatório com descrição geológica através de modelos de fácies), que seja capaz de honrar os dados de produção e ao mesmo tempo conservar o realismo geológico.

1.2 Objetivos Específicos

- Investigação e estudo comparativo de diversos métodos iterativos baseados no ES aplicados a problemas com alta não linearidade e de grande dimensão.
- Desenvolvimento, implementação e avaliação de um método de parametrização baseado em PCA, para modelar adequadamente realizações de fácies geológicas, que sejam capazes de ser integradas com métodos baseado no filtro de Kalman.
- Estudo de técnicas baseadas no DL, para extração de características e geração de novos modelos de fácies geológicas.
- Implementação e avaliação de um método de parametrização usando técnicas de DL, para modelar adequadamente realizações de fácies geológi-

cas complexas, que sejam capazes de ser integradas com métodos baseados no filtro de Kalman.

1.3

Resumo dos Capítulos

O restante desta tese está organizada da seguinte forma:

- Capítulo 2 apresenta uma breve revisão bibliográfica e os fundamentos teóricos sobre ajuste de histórico, filtro de Kalman e suas variantes, métodos para atualização de modelos de fácies geológicas, métodos de parametrização baseados em PCA (aplicado no problema de ajuste de histórico) e finalmente conceitos básicos de geoestatística.
- Capítulo 3 apresenta uma investigação de métodos baseados no ES, apresentando uma comparação de três métodos (ES-MDA, IAGS e GMMDA). Ao final mostra-se os resultados comparando as três abordagens em três casos de estudo. O primeiro caso de estudo é um dedu altamente não linear mas de pequena dimensão. O segundo caso de estudo é um dedu muito conhecido na indústria do petróleo (PUNQ-S3) e o último caso de estudo é um dedu maior (BRUGGE), com aproximadamente 222750 parâmetros a serem atualizados.
- Capítulo 4 introduz a metodologia para a atualização de modelos de fácies geológicas baseado em PCA e é integrada com o método ES-MDA para realizar o ajuste de histórico de um dedu sintético com fácies geológicas, criado a partir do dedu PUNQ-S3. Ao final do capítulo são apresentados os resultados do ajuste de histórico, utilizando a metodologia de parametrização e sem parametrização.
- Capítulo 5 apresenta um estudo de técnicas baseadas no DL, especialmente de abordagens de aprendizado de extração de características de alto nível de abstração (*deep generative models*). O objetivo deste capítulo é mostrar diferentes técnicas capazes de realizar parametrizações ótimas de modelos de fácies geológicas, e métodos geradores de novas realizações de fácies. Mostra-se também os resultados obtidos, em imagens de fácies geológicas de três diferentes modelos de reservatórios sintéticos.
- Capítulo 6 apresenta o modelo híbrido que consiste na integração do método ES-MDA com parametrização baseada em DL (ES-MDA-DL). A metodologia proposta para realizar ajuste de histórico incluindo fácies geológicas é avaliada em quatro modelos de reservatórios com fácies geológicas complexas.

- Capítulo 7 apresenta um resumo das contribuições, conclusões e comentários finais da tese e serão discutidas possíveis direções de pesquisa futuras.

2.1

Ajuste de Histórico

Tradicionalmente, o ajuste de histórico é visto como um processo cujo objetivo é a determinação dos parâmetros do reservatório e aquíferos a partir de dados de produção medidos no campo [31]. Em estudos de avaliação de incerteza e risco de campos maduros, o processo de ajuste é fundamental para a compreensão da atuação dos mecanismos, seleção dos atributos, montagem e validação dos modelos a serem utilizados [32]. Em termos gerais, o ajuste de histórico de produção busca minimizar a incerteza nas curvas de produção (água, óleo e/ou gás) e pressão do reservatório, minimizando a diferença entre os dados simulados e observados ao longo da vida do campo. A complexidade geológica que ocorre na maioria dos reservatórios de petróleo é uma das principais responsáveis pela dificuldade do processo de ajuste [33].

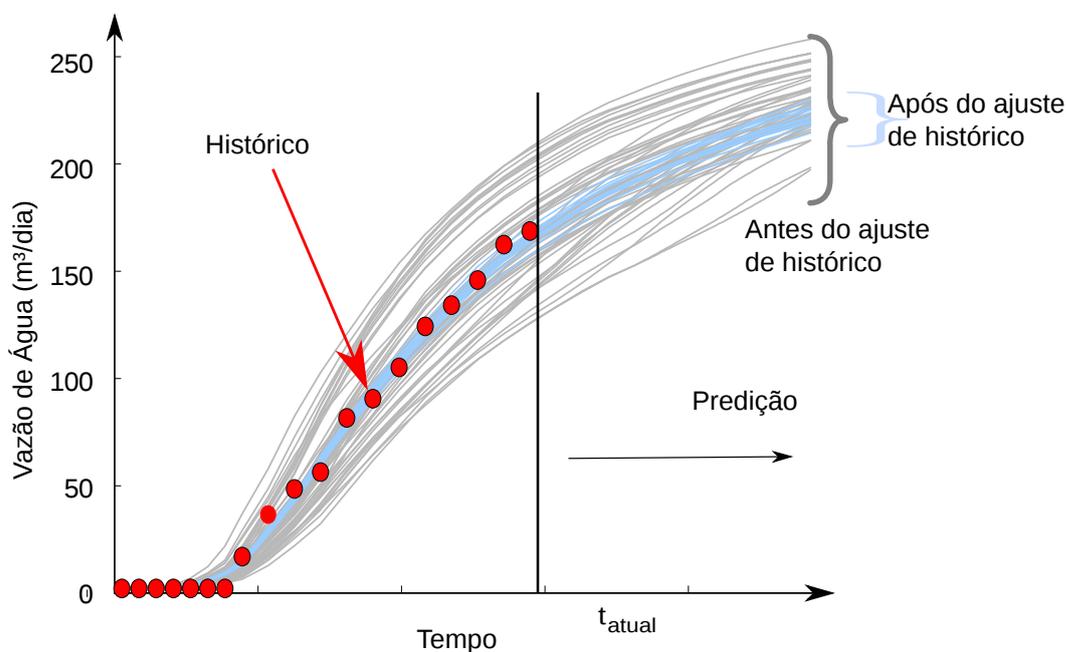


Figura 3: Ajuste de Histórico e previsão da Produção.

Dessa forma, o ajuste de histórico é comumente definido como um processo de otimização que consiste em encontrar um conjunto de parâmetros

do modelo do reservatório que minimize a diferença entre valores calculados e observados, como pressão e taxas de produção de fluidos. Oliver e Chen [6] apresentaram uma revisão geral do processo de ajuste de histórico, mostrando uma comparação entre diferentes métodos para resolver o ajuste de histórico. Os primeiros trabalhos que buscaram a automação do ajuste de histórico utilizaram métodos baseados em gradiente, que gastam muito tempo computacional encontrando a melhor combinação de parâmetros que minimizam a função objetivo. A definição dessa função objetivo é feita em termos de uma “distância” entre os dados observados e simulados. Entre as diferentes formulações, pode-se encontrar [34]:

- Formulação de mínimos quadrados

$$FO = (\mathbf{d}_{\text{obs}} - \mathbf{d})^T (\mathbf{d}_{\text{obs}} - \mathbf{d}). \quad (2-1)$$

- Formulação de mínimos quadrados ponderada

$$FO = (\mathbf{d}_{\text{obs}} - \mathbf{d})^T \mathbf{W} (\mathbf{d}_{\text{obs}} - \mathbf{d}). \quad (2-2)$$

- Formulação de mínimos quadrados ponderada e com termo de regularização

$$FO = (\mathbf{d}_{\text{obs}} - \mathbf{d})^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{d}) + (\mathbf{m} - \mathbf{m}_{\text{pr}})^T \mathbf{C}_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_{\text{pr}}), \quad (2-3)$$

onde \mathbf{d}_{obs} e \mathbf{d} representam vetores de dados observados e simulados, respectivamente, \mathbf{m} representa o vetor de parâmetros do reservatório (por exemplo porosidade, permeabilidade), \mathbf{m}_{pr} representa o vetor de parâmetros *a priori*, \mathbf{W} é uma matriz diagonal que atribui pesos a cada medição e $\mathbf{C}_{\mathbf{d}}$ e $\mathbf{C}_{\mathbf{m}}$ são as matrizes de covariâncias dos dados e dos parâmetros do modelo. Note que as matrizes $\mathbf{C}_{\mathbf{m}}$ e $\mathbf{C}_{\mathbf{d}}$ precisam ser positivas definidas para que as inversas existam.

Geralmente, reconhece-se que ajuste de histórico é um problema mal-posto e que várias combinações de parâmetros de entrada podem resultar em modelos igualmente bem ajustados. Além disso, os engenheiros de reservatório com experiência sabem que o comportamento de previsão de um único modelo de reservatório com o ajuste de histórico é problemático. Uma previsão precisa é muito desafiadora, já que as decisões que custam a companhia de petróleo milhões de dólares estão sendo feitas com base na previsão de um único modelo. O problema é, essencialmente, que não se pode reduzir a incerteza de um único modelo. Por exemplo na Figura 3, com um único modelo ajustado, não se pode saber se a propagação na resposta prevista do reservatório é ampla

ou estreita. Quando as decisões são feitas com base em um único modelo, a suposição implícita é que há zero (ou insignificante) incerteza, que pode não estar correta. Em vez de assumir uma incerteza particular espalhada em torno de um único modelo, pode ser melhor, em muitas circunstâncias, dedicar mais esforço em vários modelos ajustados usando diversos métodos que fornecem um conjunto de modelos ajustados.

Apesar de os primeiros trabalhos sobre ajuste automático de histórico serem da década de 60, ainda hoje não existem métodos que permitem um ajuste puramente automático, considerando todas as restrições impostas pela modelagem geológica. Na realidade, a ideia de ajuste automático foi substituída pelo conceito de AHA, em que o computador é usado para automatizar etapas do processo que é supervisionado pelos engenheiros do campo. A literatura sobre métodos para a solução do AHA é bastante vasta.

Oliver e Chen [6] apresentaram uma revisão recente sobre os principais métodos disponíveis na literatura. Devido à grande variedade, é difícil fazer uma classificação única desses métodos. Entretanto, pode-se dizer que existem pelo menos quatro classes principais de métodos:

1. **Métodos baseados em otimização com informação de derivadas:** Foram os primeiros métodos estudados para ajuste de histórico. Costumam ser bastante eficientes, embora requeiram a disponibilidade de informações das derivadas da função objetivo em relação aos parâmetros de ajuste, o que pode limitar a aplicabilidade desses métodos. Dentre as diversas publicações utilizando métodos de otimização com derivadas pode-se citar [35, 36, 37, 38].
2. **Métodos de otimização estocástica:** São uma grande classe de algoritmos de exploração que têm sido utilizados na indústria de petróleo e gás (e.g., *Evolutionary algorithms* [39], *Particle Swarm algorithms* [40], *Simulated annealing* [41]), notavelmente. Muitas vezes estes métodos de otimização podem ser considerados como uma “caixa preta”, onde o *loop* de otimização não requer conhecimento íntimo de como a resposta do modelo do reservatório é obtida. Uma grande desvantagem é que eles podem exigir centenas ou milhares de avaliações da resposta do modelo do reservatório, que não é viável em casos reais.
3. **Métodos baseados em planejamento de experimentos e modelos substitutos:** Nesses métodos, o simulador de reservatório é substituído por um modelo simplificado, normalmente uma função polinomial ou até mesmo uma rede neural. A maioria das ferramentas comerciais de ajuste de histórico utilizam essa estratégia. Entretanto, esses métodos são

limitados a problemas com um número bastante reduzido de parâmetros de ajuste do modelo. Há também muitos trabalhos aplicados ao problema de ajuste de histórico, como Kabir e Young [42] no campo “Meren”, Alessio et al [43] em um campo de carbonato de “Luconia”, Peake et al [44] no reservatório de “Minolish Oolite”, e King et al [45] no campo de “Nemba”.

4. **Métodos baseados no filtro de Kalman:** Métodos mais recentes na literatura que utilizam uma formulação tipo Monte Carlo do filtro de Kalman. Permitem resolver problemas com um número elevado de parâmetros de ajuste a um custo computacional razoável. Nesta tese será utilizado o método ES-MDA [16].

A despeito da grande variedade de métodos descritos na literatura, os métodos mais modernos buscam formular o problema de ajuste de histórico como um problema de inferência Bayesiana, onde o objetivo é estimar a distribuição de parâmetros incertos do modelo condicionado a dados ruidosos medidos em campo. Os métodos baseados no filtro de Kalman foram desenvolvidos dentro dessa filosofia Bayesiana.

2.2

Métodos Baseados no Filtro de Kalman

2.2.1

Filtro de Kalman

O filtro de Kalman [3] constitui um processo recursivo e eficiente de estimação de estados de um sistema dinâmico linear, dada uma série de medições ruidosas. Através da observação da variável denominada “variável de observação”, outra variável não observável denominada “variável de estado” pode ser estimada eficientemente.

O filtro de Kalman é baseado em um modelo dinâmico-linear representado por um vetor de estados (Equação 2-4), que evolui à medida que o filtro é processado. Nele, o estado atual do sistema é associado com uma incerteza, a qual é expressa mediante uma matriz de covariância (\mathbf{C}_m). Além do modelo dinâmico do sistema, o filtro tem como base a observação (medida) do sistema, que relaciona as medições mediante uma combinação linear dos estados

(Equação 2-6)

$$\mathbf{m}_t = \mathbf{A}\mathbf{m}_{t-1} + \mathbf{w}_{t-1}, \quad (2-4)$$

$$\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_m), \quad (2-5)$$

$$\mathbf{d}_t = \mathbf{H}\mathbf{m}_t + \mathbf{e}_t, \quad (2-6)$$

$$\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_d), \quad (2-7)$$

onde \mathbf{m}_t representa o vetor de estados no passo de tempo t , \mathbf{A} é uma matriz de transição de estados que mostra como o processo transita do estado $t - 1$ para o estado t , \mathbf{w}_{t-1} representa um ruído gaussiano do sistema de média igual a zero e matriz de covariância \mathbf{C}_m (Equação 2-5), \mathbf{d}_t representa o vetor de medição, \mathbf{H} é a função de medição que relaciona o vetor de estados com as observações e \mathbf{e}_t é o ruído gaussiano da medição de média igual a zero e matriz de covariância \mathbf{C}_d (Equação 2-7).

Através da Equação 2-4 percebe-se que o estado atual \mathbf{m}_t depende exclusivamente de condições anteriores. Em outras palavras, o estado presente de \mathbf{m}_t está diretamente relacionado com o seu estado anterior, \mathbf{m}_{t-1} , uma vez que este é multiplicado pela matriz \mathbf{A} e então somado a um ruído \mathbf{w}_{t-1} . O procedimento básico do filtro de Kalman consiste em duas etapas:

1. **Previsão:** Estimar as variáveis de estado utilizando o modelo dinâmico linear, ou seja

$$\mathbf{m}_t^f = \mathbf{A}\mathbf{m}_{t-1}^a, \quad (2-8)$$

$$\mathbf{C}_m^f = \mathbf{A}\mathbf{C}_m^a\mathbf{A}^T + \mathbf{C}_m. \quad (2-9)$$

2. **Correção:** Corrigir os estados através da incorporação das observações atuais, o que é feito através das equações

$$\mathbf{m}_t^a = \mathbf{m}_t^f + \mathbf{K}_t(\mathbf{d}_{\text{obst}} - \mathbf{H}\mathbf{m}_t^f), \quad (2-10)$$

$$\mathbf{K}_t = \mathbf{C}_m^f\mathbf{H}^T(\mathbf{H}\mathbf{C}_m^f\mathbf{H}^T + \mathbf{C}_d)^{-1}, \quad (2-11)$$

$$\mathbf{C}_m^a = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{C}_m^f. \quad (2-12)$$

No sistema de equações acima (Equações 2-8 até 2-12), os índices a e f indicam que a variável respectiva é atualizada e estimada, respectivamente¹. \mathbf{d}_{obst} é a medição observada e a matriz \mathbf{K} é conhecida como o ganho de Kalman.

No entanto, a aplicação desse tipo de filtro é limitada aos modelos lineares com ruídos aditivos e gaussianos. Dessa fora, sua aplicação na prática pode ser

¹Foi mantida a nomenclatura usual na literatura onde a significa *analysis* e f significa *forecast*.

difícil, devido ao fato da maioria dos problemas reais são não-lineares. No passado, foram desenvolvidas extensões do filtro de Kalman para casos menos restritos, usando-se técnicas de linearização.

2.2.2

Extended Kalman Filter

O *extended Kalman filter* (EKF) [46] foi proposto com o objetivo de superar essa limitação utilizando um esquema de linearização. O EKF tem sido o filtro não linear mais popular das últimas quatro décadas [47], com parte dessa popularidade certamente devida à simplicidade e ao baixo custo computacional. É sabido, porém, que o fato de aproximar as não linearidades do método de sistema dinâmico por meio de expansões em série de Taylor de ordem baixa pode dificultar sua aplicabilidade em situações caracterizadas por não linearidades significativas. Mais especificamente, a qualidade das estimativas do vetor de estado produzidas pelo EKF pode ser muito baixa, caso o sistema dinâmico apresente não linearidades pronunciadas. O EKF também requer o uso de matrizes jacobianas, e o cálculo destas matrizes não é trivial, por exemplo, na simulação de reservatórios cujo modelo é altamente não linear e de grande dimensão, ou seja, há muitos parâmetros a serem estimados. Uma explicação detalhada das limitações do EKF foi apresentada por Julier e Uhlmann [48] e Arulampalam e Ristic [49].

2.2.3

Ensemble Kalman Filter

Para lidar com as limitações do filtro de Kalman, EKF e suas variantes, existem outras variantes do filtro de Kalman para sistemas não lineares. Evensen [7] introduziu o EnKF como uma versão estocástica (Monte Carlo) do filtro de Kalman para problemas de grande dimensão (essencialmente, a matriz de covariância \mathbf{C}_m é substituída por uma aproximação utilizando um conjunto de amostras). Aqui, a matriz de covariância \mathbf{C}_m é só parcialmente propagada [5] (apenas as amostras são utilizadas e não a matriz \mathbf{C}_m). O EnKF ganhou popularidade por causa de sua simples formulação conceitual e relativa facilidade de execução, sendo implementado para resolver os dois principais problemas relacionados ao uso do EKF em modelos não lineares com grande dimensionalidade da variável de estado. Assim, não há necessidade do cálculo explícito de derivadas para construir o jacobiano usado na linearização das equações dinâmicas ou a relação entre as variáveis de estado e os dados. Além disso, não é preciso calcular e atualizar a estimativa da covariância, de modo que o método pode ser prático para grandes modelos [50]. Uma

descrição matemática detalhada da implementação do método em EnKF pode ser encontrado no trabalho apresentado por Evensen [51].

No EnKF, a média de um conjunto de realizações fornece a melhor estimativa, e o conjunto fornece uma estimativa empírica da densidade de probabilidade. A implementação do EnKF começa com a geração de um conjunto com N_e , tipicamente 50 – 300 para manter o custo computacional do processo gerenciável. Os modelos iniciais devem ser consistentes com o conhecimento prévio do estado inicial e a sua distribuição de probabilidade. Além disso, é usual aumentar o vetor de estado com um vetor de dados previstos (simulados), e esses dados podem ser funções não lineares das variáveis de estado. No tempo t , o vetor de estado aumentado (\mathbf{y}_t) para o modelo é expressado como

$$\mathbf{y}_t = [(\mathbf{m}_t)^T, \mathbf{g}(\mathbf{m}_t)^T]^T, \quad (2-13)$$

onde, \mathbf{m}_t consiste de variáveis do sistema a serem estimadas e $\mathbf{g}(\mathbf{m}_t)$ é o vetor de dados simulados para o vetor de estado \mathbf{m}_t .

Analogamente ao método do filtro de Kalman, o EnKF utiliza duas etapas sequenciais (previsão e correção) durante o processo de assimilação de dados. Na etapa de correção, a atualização de cada membro do conjunto é feita utilizando a fórmula do ganho de Kalman, ou seja,

$$\mathbf{y}_{t,j}^a = \mathbf{y}_{t,j}^f + \widehat{\mathbf{K}}(\mathbf{d}_j - \mathbf{H}\mathbf{y}_{t,j}^f), \quad (2-14)$$

$$\widehat{\mathbf{K}} = \widehat{\mathbf{C}}_{\mathbf{y}_t} \mathbf{H}^T (\mathbf{H} \widehat{\mathbf{C}}_{\mathbf{y}_t} \mathbf{H}^T + \mathbf{C}_d)^{-1}, \quad (2-15)$$

$$\mathbf{d}_{\text{obs},j} = \mathbf{d}_{\text{obs}} + \mathbf{e}_{t,j}, \quad (2-16)$$

para $j = 1, 2, \dots, N_e$, onde $\widehat{\mathbf{K}}$ é o ganho de Kalman estimada a partir do conjunto, $\mathbf{d}_{\text{obs},j}$ é o vetor de observações dos dados perturbado com a mesma função de distribuição do erro de medição e \mathbf{H} é o operador de medição, que extrai a dada simulada do vetor de estado \mathbf{y}_t

$$\mathbf{H} = [\mathbf{0} \quad \mathbf{I}], \quad (2-17)$$

onde \mathbf{I} apresenta a matriz identidade de dimensão igual ao número de observações. A matriz de covariância do vetor de estado $\widehat{\mathbf{C}}_{\mathbf{y}_t}$ ao tempo t é aproximada a partir dos membros do conjunto, de acordo com a expressão

$$\widehat{\mathbf{C}}_{\mathbf{y}_t} = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} (\mathbf{y}_{t,j}^f - \bar{\mathbf{y}}_t^f)(\mathbf{y}_{t,j}^f - \bar{\mathbf{y}}_t^f)^T, \quad (2-18)$$

onde $\bar{\mathbf{y}}_t^f$ é a média do conjunto.

O esquema de atualização sequencial do EnKF pode ser problemático em casos práticos de ajuste de histórico. O que se observa é que devido ao fato das equações de atualização serem lineares as correções feitas nos estados (variáveis primárias do simulador) se tornam inconsistentes com os parâmetros atualizados [52]. Em casos extremos, pode-se obter valores não físicos, por exemplo, pressões negativas ou saturações maiores que um. Uma solução simples para esse problema é iniciar as simulações da condição inicial (tempo zero) em cada passo de assimilação. Entretanto, essa solução onera excessivamente o processo de assimilação de dados. Alternativamente, a gente pode simplesmente abrir mão do processo de assimilação sequencial e aplicar as equações de atualização considerando todos os dados observados simultaneamente. Esse procedimento é conhecido como *ensemble smoother* (ES) [11, 12].

2.2.4

Ensemble Smoother

O ES [11, 12] foi proposto como uma metodologia de assimilação de dados para resolver problemas de estimação de parâmetros. O ES difere do EnKF ao computar uma atualização global no domínio do espaço-tempo em vez de usar atualizações recursivas no tempo. Enquanto que o EnKF assimila os dados observados sequencialmente no tempo, o ES assimila todas as observações uma só vez, utilizando as mesmas equações de atualização do EnKF, ou seja

$$\mathbf{y}_j^a = \mathbf{y}_j^f + \widehat{\mathbf{K}}(\mathbf{d}_{\text{obs},j} - \mathbf{H}\mathbf{y}_j^f) , \quad (2-19)$$

$$\widehat{\mathbf{K}} = \widehat{\mathbf{C}}_y \mathbf{H}^T (\mathbf{H} \widehat{\mathbf{C}}_y \mathbf{H}^T + \mathbf{C}_d)^{-1} , \quad (2-20)$$

$$\mathbf{d}_{\text{obs},j} = \mathbf{d}_{\text{obs}} + \mathbf{e}_j , \quad (2-21)$$

para $j = 1, 2, \dots, N_e$. A notação é a mesma à apresentada para o EnKF, sem fazer a assimilação de dados observados em cada passo de tempo. Por isso o índice t foi removido nas Equações 2-19 a 2-21.

A partir desse sistema de equações, pode-se reescrever de forma bastante compacta a equação para atualizar somente o vetor de parâmetros, \mathbf{m} , como

$$\mathbf{m}_j^a = \mathbf{m}_j^f + \widehat{\mathbf{C}}_{\text{md}}^f (\widehat{\mathbf{C}}_{\text{dd}}^f + \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^f) + \mathbf{e}_j) , \quad (2-22)$$

para $j = 1, 2, \dots, N_e$, onde $\widehat{\mathbf{C}}_{\text{md}}^f$ é a matriz de covariância cruzada entre o vetor *a priori* dos parâmetros do modelo, \mathbf{m}^f , e o vetor dos dados previstos, $\mathbf{g}(\mathbf{m}^f)$; $\widehat{\mathbf{C}}_{\text{dd}}^f$ é a matriz de covariância dos dados previstos de dimensão $N_d \times N_d$; \mathbf{d}_{obs} o vetor de dados observados; \mathbf{e}_j é o vetor de perturbação acrescentado aos dados observados e \mathbf{C}_d é a matriz de covariância do erro dos dados observados de

dimensão $N_d \times N_d$. As matrizes $\hat{\mathbf{C}}_{\mathbf{d}\mathbf{d}}^f$ e $\hat{\mathbf{C}}_{\mathbf{m}\mathbf{d}}^f$ são calculadas usando

$$\hat{\mathbf{C}}_{\mathbf{d}\mathbf{d}}^f = \frac{1}{N_e - 1} (\Delta \mathbf{D}^f) (\Delta \mathbf{D}^f)^T \quad (2-23)$$

e

$$\hat{\mathbf{C}}_{\mathbf{m}\mathbf{d}}^f = \frac{1}{N_e - 1} (\Delta \mathbf{M}^f) (\Delta \mathbf{D}^f)^T, \quad (2-24)$$

onde $\Delta \mathbf{D}^f$ é matriz de $N_d \times N_e$ e $\Delta \mathbf{M}^f$ é uma matriz de $N_m \times N_e$, na qual cada coluna corresponde a um elemento do conjunto subtraído de sua respectiva média, sendo N_d , N_e e N_m o número de dados observados, o número de elementos do conjunto e o número de parâmetros do modelo, respectivamente. As matrizes \mathbf{D} e \mathbf{M} têm a forma

$$\mathbf{D} = [\mathbf{g}(\mathbf{m}_1^f), \mathbf{g}(\mathbf{m}_2^f), \dots, \mathbf{g}(\mathbf{m}_{N_e}^f)], \quad (2-25)$$

$$\mathbf{M} = [\mathbf{m}_1^f, \mathbf{m}_2^f, \dots, \mathbf{m}_{N_e}^f]. \quad (2-26)$$

Portanto,

$$\Delta \mathbf{D}^f = \mathbf{D} - \bar{\mathbf{D}}^f, \quad (2-27)$$

$$\Delta \mathbf{M}^f = \mathbf{M} - \bar{\mathbf{M}}^f, \quad (2-28)$$

onde $\bar{\mathbf{D}}^f$ e $\bar{\mathbf{M}}^f$ representam a média do dado simulado e a média do vetor de parâmetros a estimar, respectivamente.

O uso do esquema de atualização do ES elimina o problema de inconsistência parâmetro-estado pelo fato de que não utiliza assimilação sequencial, além de apresentar custo computacional menor em comparação com o EnKF [12].

Entretanto, diversos resultados reportados na literatura [14, 15, 16, 53] mostram que a aplicação da Equação 2-22 uma única vez é insuficiente para gerar um conjunto de modelos adequadamente condicionados ao histórico de produção em problemas realistas de ajuste de histórico. Esse fato motivou o desenvolvimento de ES iterativos [14, 15, 16, 17, 18].

2.2.5 ES-MDA

Inspirado no conceito de assimilar todos os dados apenas uma vez, nasce o método ES-MDA proposto por Emerick e Reynolds [16]. Neste método, todos os dados são assimilados múltiplas vezes utilizando uma matriz de covariância

inflada, mediante um fator de inflação, sendo esse método uma abordagem iterativa do ES, o qual permite obter um bom ajuste de histórico.

Baseado na formulação do ES (Equação 2-22) e incorporando o fator de inflação α_k , obtém-se o modelo de atualização dos parâmetros a estimar mediante ES-MDA. Assim, o esquema de atualização, modificando a Equação 2-22 para a formulação do ES-MDA, é representado por

$$\mathbf{m}_j^{k+1} = \mathbf{m}_j^k + \hat{\mathbf{C}}_{\text{md}}^k (\hat{\mathbf{C}}_{\text{dd}}^k + \alpha_k \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^k) + \mathbf{e}_j). \quad (2-29)$$

Nota-se que os índices a e f da Equação 2-22 de atualização de ES são substituídos pelos índices $k + 1$ e k respectivamente, que correspondem aos índices do processo iterativo, onde $k = 1, 2, \dots, N_a$, sendo N_a o número de assimilações do processo iterativo. Para a implementação do ES-MDA é preciso escolher os valores do fator de inflação α_k em cada iteração. A única condição para a escolha do fator de inflação é a seguinte

$$\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1. \quad (2-30)$$

Tabela 1: Fluxo de Trabalho do ES-MDA

<i>Ensemble Smoother with Multiple Data Assimilation</i>	
Definir $\mathbf{m}_j^0 = \mathbf{m}_{\text{pr},j}$ para $j = 1, \dots, N_e$	
Para $k = 1$ a N_a	
$\mathbf{g}(\mathbf{m}_j^{k-1}) \leftarrow$ Executa \mathbf{m}_j^{k-1} desde o tempo zero.	
Calcular o vetor de perturbações para cada membro do conjunto: $\mathbf{e}_j = \sqrt{\alpha_k} \mathbf{C}_d^{1/2} \mathbf{z}_d$ onde $\mathbf{z}_d \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{N_d})$	
Calcular $\hat{\mathbf{C}}_{\text{dd}}^{k-1}$ e $\hat{\mathbf{C}}_{\text{md}}^{k-1}$, utilizando as Equações 2-23 e 2-24	
Atualizar cada membro do conjunto mediante: $\mathbf{m}_j^k = \mathbf{m}_j^{k-1} + \hat{\mathbf{C}}_{\text{md}}^{k-1} (\hat{\mathbf{C}}_{\text{dd}}^{k-1} + \alpha_k \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^{k-1}) + \mathbf{e}_j)$	
Fim	

A condição da Equação 2-30 foi proposta por Emerick e Reynolds [16] utilizando álgebra linear para garantir que o ES-MDA converge para o mesmo resultado do ES no caso de um problema linear-gaussiano. Entretanto, essa mesma condição pode ser derivada direta do teorema de Bayes [54]. Não existe um consenso na literatura de como escolher os valores de α_k . Na prática tem sido utilizado $\alpha_k = N_a$ [55, 54]. Um problema do método é que o número de iterações, N_a , deve ser escolhido *a priori*. Na prática, tem-se observado que N_a variando entre 4 e 10 é suficiente [54]. Existem versões mais recentes do

método onde α_k e N_a são escolhidos de forma automática, veja, por exemplo, [54, 56].

O fluxo de trabalho completo do algoritmo ES-MDA é apresentado na Tabela 1, onde \mathbf{m}_{pr} representa o vetor de parâmetros do modelo *a priori*.

2.2.6 IAGS

IAGS é um método iterativo que busca combinar o método ES com conceitos do filtro de partículas [57], acrescentando o uso de pesos para cada membro do conjunto e lidando, assim, com a hipótese de gaussianidade que utilizam os métodos baseados no filtro de Kalman. O método IAGS foi desenvolvido por Stordal [18] e utiliza equações similares ao ES-MDA, mas com pesos para cada membro do conjunto. Para implementar o método, define-se uma função de distribuição empírica

$$\widehat{\psi}^{N_e} * \phi_h(\mathbf{m}) = \sum_{j=1}^{N_e} W_j \mathcal{N}(\bar{\mathbf{m}}_j, h^2 \widehat{\mathbf{C}}_{\mathbf{m}}), \quad (2-31)$$

onde h é o parâmetro de largura de banda do Kernel gaussiano e $\widehat{\mathbf{C}}_{\mathbf{m}}$ é a matriz de covariância de dimensão $N_m \times N_m$. Os pesos, W , são correções calculadas mediante a

$$W_j = \frac{\mu(\mathbf{m}_j)}{\widehat{\psi}^{N_e} * \phi_h(\mathbf{m}_j)}, \quad (2-32)$$

onde $\mu(\mathbf{m})$ é a função densidade de probabilidade (FDP) do modelo. Assim, utilizando-se a distribuição empírica, Equação 2-31, dentro do teorema de Bayes, obtêm-se

$$p(\mathbf{m}|\mathbf{d}_{obs}) \propto \sum_{j=1}^{N_e} W_j^a \mathcal{N}(\mathbf{m}_j^a, \widehat{\mathbf{C}}_{\mathbf{m}}^a). \quad (2-33)$$

Assim, o esquema de atualização do vetor de parâmetros do modelo, \mathbf{m}^a , para o algoritmo de IAGS é modelado como

$$\mathbf{m}_j^a = \mathbf{m}_j^f + \widehat{\mathbf{C}}_{\mathbf{md}}^f (\widehat{\mathbf{C}}_{\mathbf{dd}}^f + h^{-2} \mathbf{C}_{\mathbf{d}})^{-1} (\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_j^f) + \mathbf{e}_j), \quad (2-34)$$

$$\widehat{\mathbf{C}}_{\mathbf{m}}^a = \widehat{\mathbf{C}}_{\mathbf{m}}^f - \widehat{\mathbf{C}}_{\mathbf{md}}^f (\widehat{\mathbf{C}}_{\mathbf{dd}}^f + h^{-2} \mathbf{C}_{\mathbf{d}}) \widehat{\mathbf{C}}_{\mathbf{md}}^{fT}, \quad (2-35)$$

$$W_j^a \propto W_j^f \mathcal{N}(\mathbf{g}(\mathbf{m}_j^f), h^2 \widehat{\mathbf{C}}_{\mathbf{dd}}^f + \mathbf{C}_{\mathbf{d}}). \quad (2-36)$$

Como cada membro do conjunto tem um peso associado as matrizes $\widehat{\mathbf{C}}_{\mathbf{md}}^f$

e $\hat{\mathbf{C}}_{\text{dd}}^f$ são calculadas usando

$$\Delta \mathbf{M}^f = [(\mathbf{m}_1^f - \bar{\mathbf{m}}^f)\sqrt{W_1}, \dots, (\mathbf{m}_{N_e}^f - \bar{\mathbf{m}}^f)\sqrt{W_{N_e}}], \quad (2-37)$$

$$\Delta \mathbf{D}^f = [(\mathbf{g}(\mathbf{m}_1^f) - \bar{\mathbf{d}})\sqrt{W_1}, \dots, (\mathbf{g}(\mathbf{m}_{N_e}^f) - \bar{\mathbf{d}})\sqrt{W_{N_e}}], \quad (2-38)$$

$$\hat{\mathbf{C}}_{\text{md}}^f = \frac{\Delta \mathbf{M}^f (\Delta \mathbf{D}^f)^T}{1 - \sum_{j=1}^{N_e} W_j^2}, \quad (2-39)$$

$$\hat{\mathbf{C}}_{\text{dd}}^f = \frac{\Delta \mathbf{D}^f (\Delta \mathbf{D}^f)^T}{1 - \sum_{j=1}^{N_e} W_j^2}. \quad (2-40)$$

Nas Equações 2-39 e 2-40, W_j é o peso associado a cada membro do conjunto e $\bar{\mathbf{m}}^f$ é a média ponderada do conjunto. Para lidar com o colapso dos pesos durante o processo iterativo, eles são reduzidos para pesos uniformes, utilizando a seguinte equação

$$\widehat{W}_j^\delta = \delta \widehat{W}_j + (1 - \delta) \frac{1}{N_e}. \quad (2-41)$$

O termo, δ é um fator de encolhimento dos pesos, $0 < \delta \leq 1$. Em teoria, ele pode ser escolhido aleatoriamente, mas usualmente é escolhido de forma adaptativa mediante a

$$\delta = \frac{1}{N_e \sum_{j=1}^{N_e} \widehat{W}_j^2}. \quad (2-42)$$

O fluxo de trabalho completo do algoritmo IAGS é apresentado na Tabela 2. O número de iterações, K , e o valor do parâmetro h não estão claramente definidos, pois eles dependem muito do problema ao qual foi aplicado.

2.2.7 GMMDA

O GMMDA [17] é uma formulação iterativa do ES que aproveita o conceito dos pesos introduzidos pela metodologia de IAGS, além da assimilação dos dados com um fator de inflação na matriz de covariância do erro das observações, similar à metodologia de ES-MDA.

A extensão do filtro de misturas gaussianas é direta utilizando o teorema de Bayes, mas com verossimilhança modificada, $l(\mathbf{m})^{(1/\alpha_k)}$, similar à Equação 2-33. A formulação do GMMDA para a atualização do vetor de parâmetros a

Tabela 2: Fluxo de Trabalho do IAGS

<i>Iterative Adaptive Gaussian Mixture Smoother</i>	
Definir $\mathbf{m}_j^0 = \mathbf{m}_{\text{pr},j}$ para $j = 1, \dots, N_e$	
Calcular o vetor de perturbações para cada membro do conjunto: $\mathbf{e}_j = \mathbf{C}_d^{1/2} \mathbf{z}_d$ Onde $\mathbf{z}_d \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{N_d})$	
Definir $\widehat{\psi}^{N_e} * \phi_h(\mathbf{m}) = \mu(\mathbf{m})$	
Definir $\widehat{W}_j^0 = \frac{1}{N_e}$	
Para $k = 1$ a K	
Se $k > 1$ então	
Set $\widehat{\psi}^{N_e} * \phi_h(\mathbf{m}) = \sum_{j=1}^{N_e} \widehat{W}_j^{k-1} \mathcal{N}(\bar{\mathbf{m}}_j^{k-1}, h^2 \widehat{\mathbf{C}}_{\text{mm}}^{k-1})$	
Fim	
$\mathbf{g}(\mathbf{m}_j^{k-1}) \leftarrow$ Executa \mathbf{m}_j^{k-1} desde o tempo zero.	
Calcular $\widehat{\mathbf{C}}_{\text{dd}}^{i-1}$ e $\widehat{\mathbf{C}}_{\text{md}}^{i-1}$, utilizando a Equação 2-39	
Atualizar cada membro do conjunto mediante:	
$\mathbf{m}_j^k = \mathbf{m}_j^{k-1} + \widehat{\mathbf{C}}_{\text{md}}^{k-1} (\widehat{\mathbf{C}}_{\text{dd}}^{k-1} + h_k^{-2} \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^{k-1}) + \mathbf{e}_j)$	
Calcular os pesos:	
$W_j^k = \frac{\mu(\mathbf{m}_j^{k-1})}{\widehat{\psi}^{N_e} * \phi_h(\mathbf{m}_j^{k-1})},$	
$\widehat{W}_j^k = W_j^k \mathcal{N}(\mathbf{g}(\mathbf{m}_j^{k-1}), h_k^2 \widehat{\mathbf{C}}_{\text{dd}}^{k-1} + \mathbf{C}_d)$	
Normalizar os pesos mediante as Equações 2-41 e 2-42;	
$\widehat{W}_j^k = \delta \widehat{W}_j^k + (1 - \delta) \frac{1}{N_e}$	
Fim	

estimar, \mathbf{m}^a , é realizada mediante a

$$\begin{aligned}
 \mathbf{m}_j^a &= \mathbf{m}_j^f + \widehat{\mathbf{C}}_{\text{md}}^f (\widehat{\mathbf{C}}_{\text{dd}}^f + h^{-2} \alpha_k \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^f) - \mathbf{e}_j), \\
 \widehat{\mathbf{C}}_{\text{m}}^a &= \widehat{\mathbf{C}}_{\text{m}}^f - \widehat{\mathbf{C}}_{\text{md}}^f (\widehat{\mathbf{C}}_{\text{dd}}^f + h^{-2} \alpha_k \mathbf{C}_d) (\widehat{\mathbf{C}}_{\text{m}}^f)^T, \\
 W_j^a &\propto W_j^f \mathcal{N}(\mathbf{g}(\mathbf{m}_j^f), h^2 \widehat{\mathbf{C}}_{\text{dd}}^f + \alpha_k \mathbf{C}_d),
 \end{aligned} \tag{2-43}$$

onde o parâmetro α_k e o número de iterações, N_a , são os mesmos utilizados no método ES-MDA e h é a largura de banda do Kernel gaussiano similar ao método IAGS, mas aqui mantem-se constante em cada iteração. O valor do fator de inflação (α_k) também segue a restrição da Equação 2-30.

Para lidar com o colapso dos pesos durante o processo iterativo, os pesos são reduzidos uniformemente, como foi mostrado no método IAGS (Equação 2-41 e 2-42). O fluxo de trabalho completo do algoritmo GMMDA é apresentado na Tabela 3.

Tabela 3: Fluxo de Trabalho do GMMDA

<i>Gaussian Mixture with Multiple Data Assimilation</i>	
Definir $\mathbf{m}_j^0 = \mathbf{m}_{\text{pr},j}$	
Calcular o vetor de perturbações para cada membro do conjunto: $\mathbf{e}_j = \mathbf{C}_d^{1/2} \mathbf{z}_d$	
Onde $\mathbf{z}_d \sim \mathcal{N}(0, I_{N_d})$	
Definir os pesos iniciais para cada membro do conjunto: $\widehat{W}_j^0 = \frac{1}{N_e}$	
Para $k = 1$ a N_a	
$\mathbf{g}(\mathbf{m}^{k-1}) \leftarrow$ Executa \mathbf{m}^{k-1} desde o tempo zero.	
Calcular $\widehat{\mathbf{C}}_{\text{dd}}^{k-1}$ e $\widehat{\mathbf{C}}_{\text{md}}^{k-1}$, utilizando a Equação 2-39	
Atualizar cada membro do conjunto mediante:	
$\mathbf{m}_j^k = \mathbf{m}_j^{k-1} + \widehat{\mathbf{C}}_{\text{md}}^{k-1} (\widehat{\mathbf{C}}_{\text{dd}}^{k-1} + \alpha_k h^{-2} \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^{k-1}) + \mathbf{e}_j)$,	
Calcular os pesos:	
$W_j^k = W_j^{k-1} \mathcal{N}(\mathbf{g}(\mathbf{m}_j^{k-1}), h^2 \widehat{\mathbf{C}}_{\text{dd}}^{k-1} + \alpha_k \mathbf{C}_d)$	
Normalizar os pesos mediante as Equações 2-41 e 2-42:	
$W_j^k = \delta W_j^k + (1 - \delta) \frac{1}{N_e}$	
Fim	

2.2.8

Localização do Ganho de Kalman

Localização do Ganho de Kalman [58, 59, 60] é um procedimento bastante utilizado para reduzir as correlações espúrias que geralmente são geradas no processo de amostragem dos métodos baseados no EnKF e para incrementar o número de graus de liberdade disponível para assimilar os dados.

O procedimento comum de localização assume que o ganho de Kalman depende da distância entre os parâmetros e dados observados. Esse procedimento assume que somente as medições localizadas dentro de uma distância específica deveria afetar a análise. Nessa abordagem a localização do ganho de Kalman é feita mediante o produto *Schur* ou *Hadamard* entre uma matriz de correlação e o ganho de Kalman. Assim, é importante para essa abordagem a escolha do alcance de correlação para formar a matriz de correlação.

A escolha do alcance de correlação depende do problema e não existe um consenso na literatura de como deve ser feito. Emerick e Reynolds [60] argumentam que essa escolha deve levar em consideração as correlações *a priori* usados na modelagem geológica e também a área de drenagem do cada poço.

A equação do método método ES-MDA com localização do ganho de Kalman pode ser escrita como

$$\mathbf{m}_j^k = \mathbf{m}_j^{k-1} + \rho_{\mathbf{K}} \circ [\widehat{\mathbf{C}}_{\text{md}}^{k-1} (\widehat{\mathbf{C}}_{\text{dd}}^{k-1} + \alpha_k \mathbf{C}_d)^{-1}] (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^{k-1}) + \mathbf{e}_j), \quad (2-44)$$

onde $\rho_{\mathbf{K}}$ representa a matriz de correlação e o símbolo “o” denota o produto *Schur*, que é um produto elemento a elemento das matrizes. Para calcular a matriz de correlação $\rho_{\mathbf{K}}$ geralmente é utilizada uma função de correlação, como por exemplo a função de correlação compacta de quinta ordem proposta por Gaspari e Cohn [61]:

$$\rho_{\mathbf{K}} = \begin{cases} -\frac{1}{4}\left(\frac{h}{L}\right)^5 + \frac{1}{2}\left(\frac{h}{L}\right)^4 + \frac{5}{8}\left(\frac{h}{L}\right)^3 - \frac{5}{3}\left(\frac{h}{L}\right)^2 + 1 & 0 \leq h < L \\ \frac{1}{12}\left(\frac{h}{L}\right)^5 - \frac{1}{2}\left(\frac{h}{L}\right)^4 + \frac{5}{8}\left(\frac{h}{L}\right)^3 + \frac{5}{3}\left(\frac{h}{L}\right)^2 - 5\left(\frac{h}{L}\right) + 4 - \frac{2}{3}\left(\frac{h}{L}\right)^{-1} & L \leq h \leq 2L \\ 0 & h > 2L \end{cases} \quad (2-45)$$

onde L é o alcance e h é a distância euclidiana entre a célula do modelo e a localização da observação.

2.2.9

Aplicações dos Métodos Baseados no Filtro de Kalman no Ajuste de Histórico

A primeira aplicação reportada na literatura do EnKF na área de petróleo foi apresentada por Lorentzen et. al [62]. Já a primeira aplicação em ajuste de histórico foi apresentada por Nævdal et. al [62]. Após essa aplicação pioneira, o EnKF chamou muita atenção da academia e indústria como um método promissor para ajuste de histórico. A popularidade do EnKF e suas variantes como método de ajuste de histórico pode ser atribuída aos seguintes fatos:

- Permitem ajustar modelos com um número muito grande de parâmetros e a um custo computacional razoável.
- Evitam a linearização do modelo de reservatório, o que é desejável quando o gradiente do modelo não está disponível, permitindo vê-lo como uma “caixa preta” que pode ser facilmente acoplada a diferentes simuladores de reservatórios.
- Apresentam uma fácil implementação.
- Geram múltiplos modelos ajustados, o que é importante para estudos de quantificação de incertezas.

Na literatura é possível encontrar aplicações do EnKF tanto para modelos de reservatórios sintéticos como para casos reais. Gu et. al [64] aplicaram o EnKF para ajustar modelos de porosidade e permeabilidade do caso *benchmark* PUNQ-S3 [65, 66]. Os resultados apresentados pelos autores mostram uma maior eficiência em comparação com outras técnicas de ajuste de histórico até essa época. Posteriormente, Bianco et. al [67] apresentaram uma das primeiras aplicações do EnKF dentro de um reservatório real. Por outro lado, Skjervheim

et. al [68] apresentaram resultados da comparação de um tradicional método de ajuste de histórico [69] e o EnKF para um reservatório real mostrando uma significativa melhoria no ajuste obtido pelo EnKF.

Embora exista uma grande variedade de métodos baseados no Filtro de Kalman, o EnKF é de longe o mais pesquisado. Por outro lado, os métodos iterativos baseados no ES são menos investigados na literatura. Entretanto, os resultados reportados em [16, 70, 18, 17] mostram que a aplicação prática dos ES iterativos são mais adequados para o problema de ajuste de histórico, principalmente por não requerem a atualização das variáveis dinâmicas do reservatório (e.g., pressão e saturação de fluidos).

Emerick e Reynolds [53] realizaram um estudo comparativo entre vários métodos baseados no EnKF para ajuste de histórico aplicados a um modelo sintético simples, porém extremamente não linear. Os resultados apresentados nesse trabalho mostram o bom desempenho do método ES-MDA sobre os outros. Apesar disso, a literatura reconhece que os métodos baseados no filtro de Kalman ainda apresentam dificuldades quando são aplicados a problemas de reservatórios de maior complexidade. De fato, esses métodos não produzem uma correta estimativa da função de distribuição de probabilidade dos parâmetros utilizados em problemas fortemente não-gaussianos.

A fim de lidar com essa hipótese de gaussianidade, surgiram novas abordagens que tentam contornar a suposição de gaussianidade implícita, introduzindo uma ponte entre os métodos baseados no filtro de Kalman com o método de filtro de partículas. Baseado nisso, Stordal [18] apresentou uma forma iterativa do ES (IAGS) combinando o conceito de IS com o método de ES. Posteriormente, Stordal e Elsheikh [17] combinaram o filtro de partículas com o ES (GMMDA). Os autores compararam o GMMDA com o ES-MDA para um caso bastante simples de ajuste de histórico. Apesar dos resultados promissores do método, ainda não está claro se o GMMDA é superior ao ES-MDA para problemas mais complexos e com maiores dimensões.

Apesar dos métodos iterativos baseados no ES apresentarem resultados com relativo sucesso na indústria do petróleo para resolver o problema de ajuste de histórico, eles não conseguem conservar algumas características geológicas e geoestatísticas, particularmente quando os modelos do reservatório são altamente não lineares ou apresentam propriedades com uma distribuição de probabilidade não gaussiana (e.g. indicadores de fácies).

2.2.10

Atualização de Modelos de Fácies

A atualização de modelos de fácies geológicas é uma dos principais desafios no desenvolvimento de métodos de AHA. A maioria das pesquisas buscam reparametrizar as fácies em termos de variáveis que possam ser facilmente modificadas durante o processo de ajuste de histórico.

Dentre as parametrizações mais utilizadas, pode-se citar as baseadas em simulação plurigaussiana truncada [71, 72, 73]. Neste caso, as fácies são parametrizadas em termos de vetores cujos elementos seguem distribuições aproximadamente gaussianas e da uma regra de truncamento. Simulação plurigaussiana truncada foi utilizada em conjunto com os métodos baseados no filtro de Kalman por diversos autores [52, 74, 75, 76, 77]. Apesar do relativo sucesso reportado nesses trabalhos, a simulação plurigaussiana não parece adequado para descrever modelos de fácies em ambientes deposicionais com múltiplos canais sísmicos (e.g., canais fluviais), como os modelos gerados com técnicas de simulação por objetos [78] e geoestatística multiponto [79, 80, 81].

Outro tipo de parametrização para ajuste de histórico de fácies que tem sido usada em conjunto com métodos baseados no filtro de Kalman são as funções de *level-set* [82]. Contudo, esse tipo de parametrização tem um desempenho bastante semelhante à simulação plurigaussiana truncada [83]. Dessa forma, o uso de funções *level-set* também não é apropriada no caso de modelos com canais sísmicos..

Parametrizações baseadas em PCA também são bastante comuns na literatura [84, 85, 86, 87, 21]. PCA é uma técnica bastante eficiente para parametrização de modelos gaussianos, e.g., modelos gerados com variogramas gaussianos, PCA pode ser facilmente estendido para parametrizar modelos de fácies simples. No Capítulo 4 dessa tese é apresentada um procedimento para o ajuste de fácies com PCA. Contudo, PCA não é adequado para descrever modelos de fácies com canais mais complexos. Esse fato levou Sarma et. al [84] a propor o uso de Kernel-PCA como técnica de parametrização de modelos de fácies gerados com MPS. Embora os autores tenham reportados resultados bastante promissores, distorcem os limites entre diferentes propriedades e apresentam um custo computacional alto.

2.3

Análise de Componentes Principais

A ideia central do PCA é reduzir a dimensionalidade de um conjunto de dados de um grande número de variáveis correlacionadas, mantendo o máximo possível da variação do conjunto de dados. Isto é feito transformando

as variáveis em um novo conjunto de componentes principais, os quais são não correlacionadas e ordenadas de acordo com quem retém a maior variância do conjunto inicial [20]. A representação do PCA implica uma expansão em série com respeito aos termos de um conjunto de vetores base ortogonais. Os vetores são derivados de uma auto decomposição da matriz de covariância do campo aleatório.

A parametrização por PCA, também conhecida como expansão linear de Karhunen-Lòeve ou a *Proper Orthogonal Decomposition* (POD) [88], proporciona uma caracterização dos dois primeiros momentos (média e covariância) da distribuição das variáveis aleatórias em funções determinísticas. PCA permite um encapsulamento ótimo das informações contidas em um processo aleatório multi-gaussiano em um conjunto de variáveis aleatórias independentes.

O procedimento geral é usar uma transformação ortogonal para representar um conjunto de realizações (possivelmente correlacionadas) do vetor aleatório \mathbf{m} em termos do vetor não correlacionada ξ . PCA pode ser implementada através da decomposição em valores singulares (SVD), problema de autovalores, e os resultados dos vetores singulares proporcionam a matriz básica ortogonal \mathbf{C}_m . A equação geral para resolver o problema de autovalores é formulada mediante a

$$\lambda \mathbf{v} = \mathbf{C}_m \mathbf{v} , \quad (2-46)$$

onde λ é um autovalor da matriz \mathbf{C}_m , e \mathbf{v} é um autovetor da matriz \mathbf{C}_m . Como resolver esse problema diretamente com um algoritmo padrão de SVD é um processo computacionalmente custoso, pode-se utilizar um enfoque baseado na construção empírica da matriz de covariância desde um conjunto de realizações *a priori*.

A parametrização com PCA tem tido muito sucesso não só na área de processamento de imagens como também na caracterização de modelos de reservatórios, pelo fato de capturar as principais características das propriedades do reservatório com um número reduzido de parâmetros.

2.3.1

Representação de Modelos Geológicos Mediante a Análise de Componentes Principais

As propriedades geológicas como permeabilidade, porosidade e indicadores de fácies podem ser vistas como campos aleatórios que, por sua vez podem ser representadas através de expressões em série que envolvem um conjunto de funções determinísticas e seus respectivos coeficientes aleatórios. Assim, uma representação mediante PCA das propriedades geológicas pode ser atribuída a um conjunto discreto de células. Realizações de propriedades geológicas em células podem ser geradas mediante simulações geoestatísticas, utilizando fer-

ramentas de modelagem geológica. Com um modelo geológico que contém N_g células e um conjunto de N_e realizações geradas se constrói a matriz \mathbf{C}_m , que é de posto deficiente, similar à Equação 2-18, ou seja

$$\mathbf{C}_m = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\mathbf{m}_i - \bar{\mathbf{m}})(\mathbf{m}_i - \bar{\mathbf{m}})^T, \quad (2-47)$$

$$\Delta\mathbf{M} = [\mathbf{m}_1 - \bar{\mathbf{m}} \quad \mathbf{m}_2 - \bar{\mathbf{m}} \quad \dots \quad \mathbf{m}_{N_e} - \bar{\mathbf{m}}], \quad (2-48)$$

$$\mathbf{C}_m = \frac{1}{N_e - 1} \Delta\mathbf{M}\Delta\mathbf{M}^T. \quad (2-49)$$

Nas equações acima, \mathbf{m}_i é o vetor que representa uma realização da propriedade geológica do reservatório de dimensão $N_g \times 1$, $\bar{\mathbf{m}}$ a média das N_e realizações e \mathbf{C}_m a matriz de covariância empírica de dimensão $N_g \times N_g$. Os autovalores e autovetores de \mathbf{C}_m podem ser calculados aplicando o algoritmo de SVD para $\Delta\mathbf{M}$

$$\Delta\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^T, \quad (2-50)$$

onde \mathbf{U} representa um matriz ortogonal com vetores singulares na esquerda de dimensão $N_g \times N_g$, Σ é uma matriz de dimensão $N_g \times N_e$ com todas as entradas não-zero correspondentes a valores singulares e \mathbf{V} é a $N_e \times N_e$, uma matriz ortogonal com vetores singulares de $\Delta\mathbf{M}$ na direita. Assim, a matriz que corresponde à raiz quadrada de \mathbf{C}_m é representada por

$$\mathbf{C}_m^{1/2} = \frac{1}{\sqrt{N_e - 1}} \mathbf{U}\Sigma, \quad (2-51)$$

assim, usando análise de componentes principais, pode-se aproximar o modelo do reservatório \mathbf{m}_j por

$$\mathbf{m}_j = \bar{\mathbf{m}} + \mathbf{C}_m^{1/2}\xi, \quad (2-52)$$

onde ξ representa os coeficientes do PCA que apresentam uma média zero e matriz de covariância identidade, $\mathbf{C}_m^{1/2}$ é uma matriz de dimensão $N_\xi \times N_g$ que está composta por vetores base ortogonais correspondentes aos N_ξ maiores autovalores de \mathbf{C}_m e N_ξ é o número de coeficientes do PCA.

As duas principais características do PCA são:

- Capacidade de captar as mais fortes correlações espaciais entre os parâmetros do modelo com poucos componentes principais.
- Os novos parâmetros gerados são variáveis independentes ξ com funções de distribuição de probabilidade gaussiana.

Embora os modelos obtidos mediante PCA tendam a modificar as características geoestatísticas ou a geologia do modelo inicial, normalmente o PCA é considerada como uma operação de suavização, a qual reduz o ruído de

alta frequência dos modelos iniciais. Entretanto, esse efeito de suavização leva a alguns problemas, como é o caso de não conservar alguns detalhes estatísticos sutis, como função de distribuição de probabilidade e função acumulada de probabilidade dos modelos iniciais. No caso de indicadores de fácies discretas, são mapeados como valores reais que levam a confundir os limites entre diferentes fácies, além de distorcer a função de distribuição de probabilidade, especialmente quando os modelos originais apresentam muitos picos e são não gaussianos.

2.3.2 Parametrização OPCA

OPCA é uma parametrização proposta por Vo e Durlofsky [86] baseada no PCA, para a representação de modelos geológicos complexos em termos de um número reduzido de parâmetros. Em OPCA, o procedimento de mapeamento de um modelo geológico a um espaço de menor dimensão é visto como um problema de otimização. Essa otimização envolve o uso restrições e de funções de regulação, que agem para ampliar substancialmente a capacidade do algoritmo PCA, tudo isso para capturar as características geologias e modelar sistemas complexos (não gaussianos). A formulação do problema de otimização para OPCA é baseada na Equação principal de PCA 2-52, e pode ser escrita como

$$\mathbf{m} = \arg \min_{\widehat{\mathbf{m}}} \left\| \overline{\mathbf{m}} + \frac{1}{\sqrt{N_R - 1}} \mathbf{U} \Sigma \xi - \widehat{\mathbf{m}} \right\|_2^2 + \gamma \widehat{\mathbf{m}}^T (\mathbf{1} - \widehat{\mathbf{m}}), \quad (2-53)$$

$$s.t. \quad \widehat{\mathbf{m}}_i \in [0, 1] \quad \forall i = 1, \dots, N_{\mathbf{m}},$$

sendo γ um fator de regularização onde $\gamma \in [0, 1]$, e $\mathbf{1} = [1, 1, \dots, 1]^T$ é um vetor de igual dimensão de \mathbf{m} . Note-se que o regularizador $\gamma \widehat{\mathbf{m}}^T (\mathbf{1} - \widehat{\mathbf{m}})$ “empurra” os valores de \mathbf{m} para 0 ou 1. Vo e Durlofsky [86] apresentaram uma descrição mais detalhada da parametrização OPCA. Para facilitar a representação do problema de minimização da Equação 2-53, define-se

$$\mathbf{m}_{PCA} = \overline{\mathbf{m}} + \frac{1}{\sqrt{N_R - 1}} \mathbf{U} \Sigma \xi. \quad (2-54)$$

Equação 2-54 que representa a solução padrão do PCA associado ao problema, assim a solução do OPCA é dada mediante a

$$\begin{aligned} \widehat{\mathbf{m}}_i &= 0, \text{ se } \mathbf{m}_{PCA,i} \leq \frac{\gamma}{2}, \\ \widehat{\mathbf{m}}_i &= 1, \text{ se } \mathbf{m}_{PCA,i} \geq 1 - \frac{\gamma}{2}, \\ \widehat{\mathbf{m}}_i &= \frac{\mathbf{m}_{PCA,i} - \frac{\gamma}{2}}{1 - \gamma}, \text{ se } \frac{\gamma}{2} \leq \mathbf{m}_{PCA,i} < 1 - \frac{\gamma}{2}. \end{aligned} \quad (2-55)$$

A parametrização OPCA também fornece uma representação das derivadas para modelos de duas e três fácies. Isso foi desenvolvido por Vo e Durlflosky [85], assim como os detalhes do processo de minimização da Equação 2-55. Embora os métodos baseados no filtro de Kalman não precisem de derivadas, isso pode ser viável para outros algoritmos de ajuste de histórico.

Quando OPCA é utilizado conjuntamente com o método ES-MDA dentro de um processo de ajuste de histórico, a Equação 2-29 de atualização do ES-MDA é reescrita para atualizar o vetor ξ que representa a parametrização, e que é usada na representação do vetor pelo PCA \mathbf{m}_{PCA} . Posteriormente, aplicando a Equação 2-55 a cada elemento, é obtido o vetor de parâmetros atualizado \mathbf{m} . A nova equação de atualização é dada por

$$\xi_j^k = \xi_j^{k-1} + \hat{\mathbf{C}}_{\xi d}^{k-1} (\hat{\mathbf{C}}_{dd}^{k-1} + \alpha_k \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_j^{k-1}) + \mathbf{e}_j) . \quad (2-56)$$

2.4 Geoestatística

A geoestatística compreende um conjunto de ferramentas e técnicas que ajudam a analisar e prever os valores de uma variável que está distribuída no espaço ou no tempo de uma forma contínua. Devido ao fato de sua aplicação ser orientada para sistemas de informação geográfica, também seria possível definir como a estatística relacionada com os dados geográficos, sendo, portanto, conhecida como estatística espacial. A geoestatística está baseada na teoria das variáveis regionalizadas, que é definida como uma função numérica com distribuição espacial e certa continuidade aparente, que varia de um lugar a outro. Dessa forma, a geoestatística define uma função aleatória como um conjunto de variáveis aleatórias \mathbf{m} caracterizadas como uma média $\bar{\mathbf{m}}$, variância \mathbf{C}_m e uma função de distribuição acumulada [89]. Além disso, essas funções aleatórias apresentam certas características, tais como suporte, localização, continuidade e anisotropia.

O objetivo principal da geoestatística é inferir as propriedades do fenômeno espacial em estudo (reservatório de petróleo) por meio da interpolação ou estimativa. O procedimento geoestatístico tem que ser feito em três passos:

1. Análise exploratória dos dados: Estuda-se os dados amostrados sem ter em conta sua distribuição geográfica (estatística comum).
2. Análise estrutural: Estuda-se a continuidade espacial da variável. Calcula-se o variograma ou outra função que explique a variabilidade espacial e ajusta-se ao mesmo um variograma teórico.

3. Predições: Estimativas da variável nos pontos não amostrados, considerando a estrutura de correlação espacial selecionada e integrando a informação obtida de forma discreta com os pontos amostrados, assim como a conseguida indiretamente sob a forma de tendências conhecidas.

2.4.1

Simulações Geoestatísticas

Simulação geoestatística é o termo usado para o processo de geração de realizações estocásticas de um vetor aleatório representando uma propriedade física de interesse. Através de simulações geoestatísticas, os geocientistas podem gerar múltiplas realizações das propriedades de um reservatório, permitindo assim que se estime as incertezas associadas ao processo de exploração do campo [71]. Em geral simulações gaussianas, também conhecidos como geoestatística de dois pontos ou baseados em variogramas, são as técnicas mais utilizadas para gerar realizações de propriedades petrofísicas de um reservatório, como porosidade e permeabilidade. Simulação gaussiana também é usada para gerar realizações de fácies através de algoritmos de simulação gaussiana truncada e plurigaussiana truncada [90]. Quando as fácies apresentam distribuições mais complexas, como, por exemplo, canais fluviais, outras técnicas são utilizadas, como a simulação booleana ou por objetos [78] ou técnicas de geoestatística multiponto [79, 80, 81].

2.4.2

Simulação Gaussiana Truncada

O método de simulação gaussiana truncada é adequado quando os reservatórios apresentam litologias em ordem sequencial. Por exemplo, em um reservatório de três fácies, a fácies 1 tem limite com a fácies 2, mas não tem limites com a fácies 3, e esta tem como vizinho somente a fácies 2. Na Figura 4 ilustra o procedimento para obter uma propriedade discreta do reservatório (e.g., indicadores de fácies) a partir de uma simulação gaussiana, Figura 4(a), definidos dois limites de truncamento em -0.3 e 0.27 (na prática, os pontos são experimentalmente calculados baseado na proporção de cada fácies). Seja \mathbf{z} o campo aleatório gaussiano; a partir dos pontos de truncamento, os valores de $\mathbf{z} < -0.3$ são estabelecidos como fácies 1 e são coloridos de azul na Figura 4(b). Para valores entre $-0.3 \leq \mathbf{z} \leq 0.27$ são coloridos de verde (fácies 2) e para $\mathbf{z} > 0.27$ são coloridos de vermelho (fácies 3).

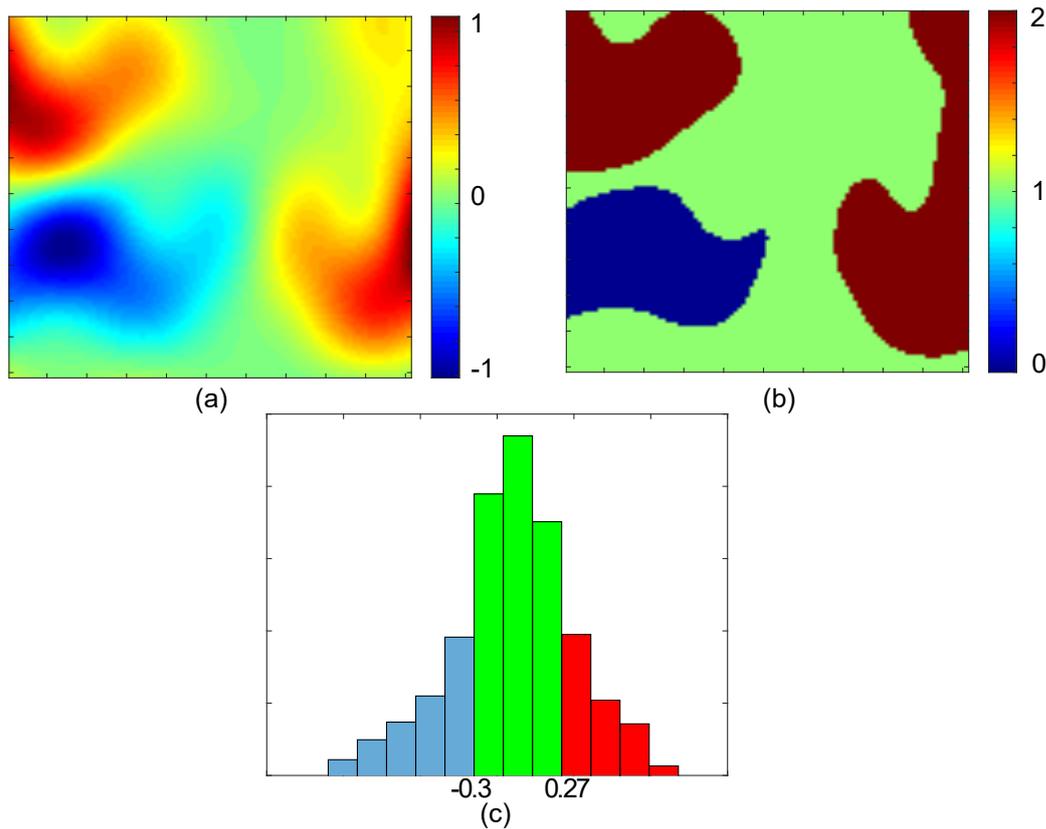


Figura 4: Procedimento de simulação gaussiana truncada para geração de fácies: na parte acima, à esquerda, o campo gaussiano; na direita, a correspondente representação da fácies para o campo gaussiano; na parte inferior, o histograma com os respectivos cortes para gerar as fácies.

2.4.3

Simulação Plurigaussiana Truncada

A simulação plurigaussiana truncada foi proposta por Galli et. al [90] como uma extensão da simulação gaussiana truncada, na qual executa os truncamentos, não sobre uma gaussiana simulada, mas sobre uma combinação de n gaussianas. Essa estratégia permite simular características geológicas mais complexas flexibilizando a restrição de ordenamento das fácies.

Uma ilustração do procedimento da simulação plurigaussiana truncada é mostrada na Figura 5, onde (a) e (b) representam as duas simulações gaussianas, \mathbf{z}_1 e \mathbf{z}_2 respectivamente, enquanto (c) apresenta o modelo de fácies gerado a partir da regra de truncamento (d). Sejam th_1 e th_2 os pontos de corte de cada gaussiana, que representam os limites do quadrilátero apresentada na Figura 5 (d) (th_1 para a horizontal e th_2 para a vertical). Assim, quando os valores das gaussianas cumprem a condição de $\mathbf{z}_1 \geq th_1 \wedge \mathbf{z}_2 \geq th_2$ é porque a posição no modelo corresponde à fácies 2 (fácies vermelha). Se $\mathbf{z}_1 \leq th_1 \cap \mathbf{z}_2 \geq th_2$ então tem-se a fácies 1 (cor verde) e, por último, se $\mathbf{z}_2 < th_2$ tem-se a fácies 0 (cor azul).

A regra de truncamento é similar ao apresentado no histograma da Figura 4(d), onde o truncamento foi feito só no eixo horizontal. Para a simulação plurigaussiana truncada, o truncamento é feito nos eixos horizontal e vertical.

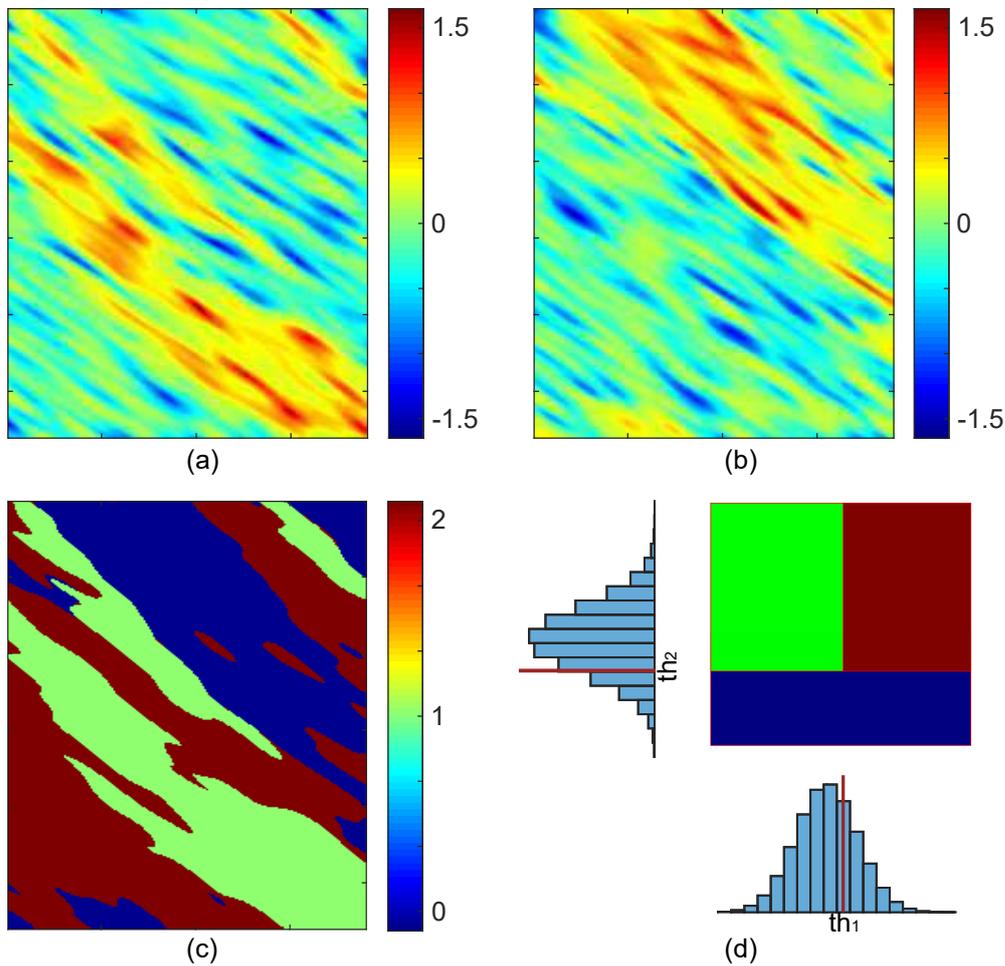


Figura 5: Procedimento de simulação gaussiana para geração de três fácies: na parte de cima os 2 campos gaussianos, na parte inferior esquerda a correspondente representação de fácies para dos campos gaussianos e na parte de baixo, à direita, o histograma com os respectivos cortes para gerar as fácies com a respectiva RTR.

2.4.4 Geoestatística Multiponto

A geoestatística é um ramo da estatística aplicada e da matemática que nasceu da necessidade de se ter uma ferramenta, mais poderosa que a estatística clássica, que permitisse a interpretação e a manipulação mais adequadas de dados provenientes de medições efetuadas na terra. Os dados dessa natureza apresentam uma correlação espacial, que acontece em maior ou menor grau, dependendo da distância que os separa, ou seja, a similaridade entre duas medições será tão maior quanto menor for a distância entre os

pontos de coleta [91]. Esse comportamento é típico de fenômenos ditos regionais e são representados pelas variáveis regionalizadas que são o objeto de estudo da geoestatística. Essa variabilidade leva em consideração tanto a heterogeneidade quanto a direcionalidade de um conjuntos de dados.

Na área de petróleo, a geoestatística só foi introduzida nos anos 70 no processo de caracterização de um reservatório, que envolve a construção de uma série de mapas que são modelos numéricos da distribuição espacial de propriedades do reservatório como, por exemplo, porosidade e permeabilidade. A qualidade dessa caracterização é fundamental para se obter previsões mais precisas a respeito do comportamento do reservatório, por isso, os mapas que compõem o modelo de um reservatório devem representar as suas heterogeneidades com a maior fidelidade possível. Para atingir esse objetivo, a geoestatística tem sido uma ferramenta amplamente aceita para a construção desses mapas, especialmente quando incorpora dados secundários como, por exemplo, dados provenientes de análise sísmica em 3D e, mais recentemente, em 4D [92].

Durante a caracterização de um reservatório, os dados reais disponíveis geralmente não são suficientes para permitir a inferência da estatística de múltiplos pontos. Guardiano [79] propôs o algoritmo de simulação sequencial por indicadores capaz de inferir as estatísticas de múltiplos pontos a partir de imagens de treinamento, em seguida, o modelo geoestatístico numérico é gerado utilizando estas estatísticas. A inferência das estatísticas a partir da imagem de treinamento elimina a necessidade de calcular o variograma para identificar a variabilidade espacial dos dados, além de descartar a necessidade da krigagem para derivar as probabilidades condicionais. Strebelle [93] propôs o algoritmo sequencial SNESIM. Esse algoritmo foi uma extensão do original proposto por Guardiano e proporcionou uma redução no custo computacional durante o cálculo da probabilidade condicional associada a cada ponto simulado. Para essa redução, Strebelle propôs um algoritmo que percorria apenas uma vez a imagem de treinamento durante todo o processo de simulação.

3

Estudo Comparativo de Métodos Baseados no Filtro de Kalman

Conforme discutido no Capítulo 2, dentre os métodos baseados no filtro de Kalman, os ES iterativos são os métodos mais adequados para o ajuste de histórico. Dentre esses métodos o ES-MDA tem sido bastante usado devido a sua robustez e simplicidade matemática na implementação computacional. Contudo, existem formulações mais recentes que buscam melhorar o desempenho do processo de assimilação de dados. Dentre essas técnicas o IAGS e o GMMDA utilizam ideias oriundos dos filtros de partícula com o objetivo de relaxar a hipótese de que a distribuição *a priori* deve ser gaussiana.

Neste capítulo é realizado um estudo comparativo entre os modelos ES-MDA, IAGS e GMMDA. Os métodos são comparados através de três casos *benchmark* disponíveis na literatura. O primeiro caso, aqui denominado “1D”, foi proposto por Emerick e Reynolds [53]. O segundo caso corresponde ao “PUNQ-S3” [65, 66] e o último caso é o “BRUGGE” [94]. Os métodos a serem comparados foram descritos no Capítulo 2 seções 2.2.5, 2.2.6 e 2.2.7 e foram implementados baseados na informação proporcionada na literatura e reescrevendo as equações, de forma que a formulação seja similar para os três métodos. Isso pode ser visto nas Tabelas 1, 2 e 3, onde foi apresentado o algoritmo completo para cada metodologia.

A Tabela 4 mostra um resumo dos modelos utilizados para os experimentos realizados. O primeiro caso de estudo (“1D”), diferentemente dos outros dois, também é comparado com os resultados obtidos pelo algoritmo MCMC, e que está disponível na literatura [53].

Na Tabela 4, κ representa o vetor de permeabilidade do reservatório e os índices I e K referem à permeabilidade horizontal e vertical, respectivamente. ϕ é o vetor de porosidade do reservatório e NTG e S_{wc} representam os vetores de *net-to-gross* e saturação de água do reservatório, respectivamente.

Tabela 4: Casos de Estudo para os Métodos Baseados no ES

	1D	PUNQ-S3	BRUGGE
Dimensão do modelo	31×1	$19 \times 28 \times 5$	$139 \times 48 \times 9$
Células ativas	31	1761	44550
Número de observações	12	240	1407
Número de conjuntos	10	10	1
Realizações por conjunto	100	100	104
Número de iterações	10 e 15	4	4
Dimensão de cada conjunto	31×100	5283×100	222750×104
Parâmetros incertos	κ	κ_I	κ_I
		κ_K	κ_K
		ϕ	ϕ
			NTG
			SW

3.1

Caso de Estudo 1D

O primeiro caso de estudo testado é um modelo com fluxo horizontal de 31×1 células, mas altamente não linear (Figura 6). Neste modelo, os parâmetros a serem estimados são o logaritmo natural das permeabilidades em cada célula. Assim, o vetor de parâmetros é formado por

$$\mathbf{m} = \ln(\kappa). \tag{3-1}$$

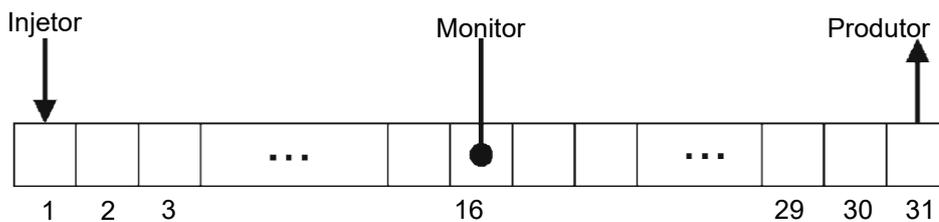


Figura 6: Modolo do caso “1D” [53].

A informação *a priori* (o conjunto inicial) foi gerada de uma distribuição de média $\kappa_{m_{pr}} = 5$ e matriz de covariância \mathbf{C}_m , a qual contém valores de 1 na diagonal e para valores fora da diagonal, foram computados mediante a função de covariância exponencial com alcance de 10 células. As medições utilizadas para o processo de ajuste de histórico correspondem à pressão na célula 16, a qual foi coletada durante um período de 360 dias em um intervalo de 30 dias e com uma matriz de erro de medição gaussiana de média 0 e variância 1 psi. Na Figura 7 mostra-se a distribuição *a posteriori* do logaritmo

natural da permeabilidade obtida a partir do algoritmo MCMC de Metropolis-Hasting com uma cadeia extremamente longa construída a partir de 20 milhões de simulações de reservatório [53]. O resultado dessa cadeia é assumido como referência para as comparações apresentadas aqui. As linhas pretas na Figura 7 corresponde aos percentis P2, P25, P75 e P98, enquanto que as linhas vermelha e verde representam a mediana e o valor verdadeiro do logaritmo natural da permeabilidade.

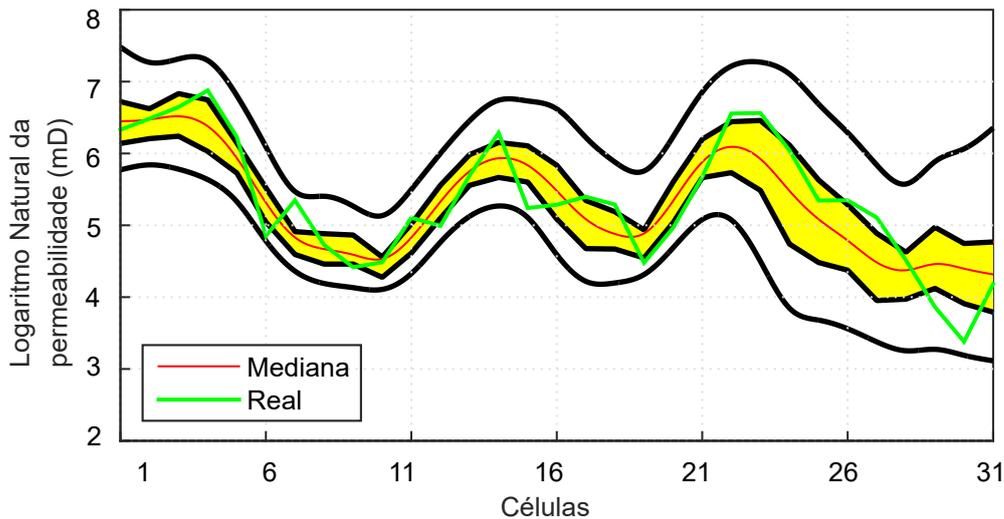


Figura 7: Distribuição *a posteriori* do logaritmo natural da permeabilidade do caso de estudo “1D” obtido por MCMC [53].

3.1.1 Resultados

A comparação da distribuição *a posteriori* do logaritmo natural da permeabilidade obtida pelos métodos ES-MDA, IAGS e GMMDA é apresentada na Figura 8. Para os resultados obtidos mediante esses algoritmos foram utilizados o mesmo número de iterações. No caso de ES-MDA e GMMDA, o número de iterações é igual ao número de assimilações múltiplas que realiza (parâmetro N_a). Para uma comparação do método IAGS foram utilizados valores de h equivalentes aos valores de α dos métodos de assimilações múltiplas. Assim, foram realizados testes de 10 e 15 iterações para 10 de conjuntos onde cada um deles apresenta 100 modelos iniciais, e os resultados apresentados são baseados na média dos 10 conjuntos analisados. Na Figura 8 mostra-se a distribuição *a posteriori* do logaritmo natural da permeabilidade, e, como pode ser visto, os resultados apresentados são muito semelhantes para os três métodos analisados e bastante próximos ao obtido com o método MCMC (referência para comparação).

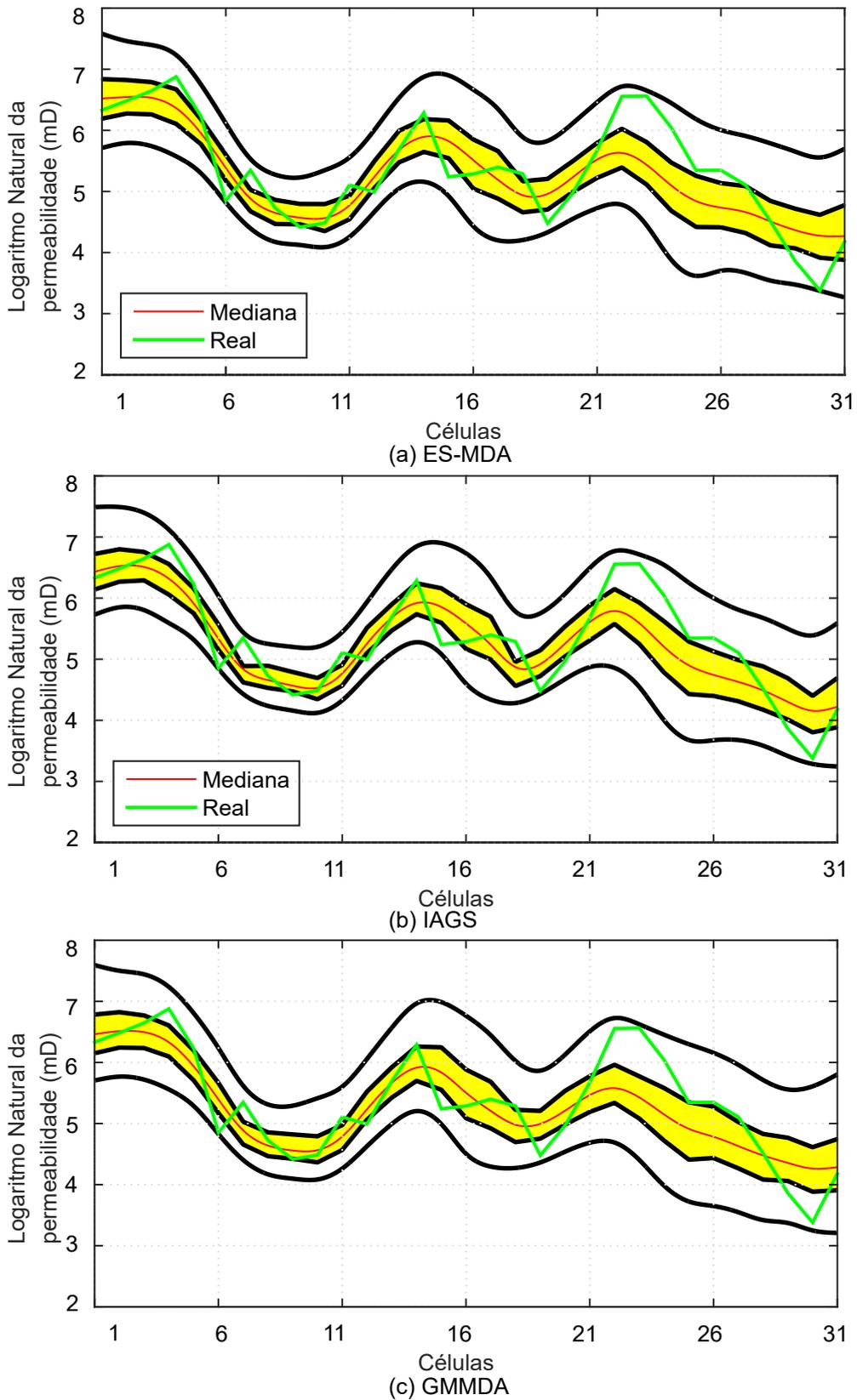


Figura 8: Distribuição *a posteriori* do logaritmo natural da permeabilidade após o procedimento de ajuste de histórico com 10 iterações, para o caso de estudo “1D”.

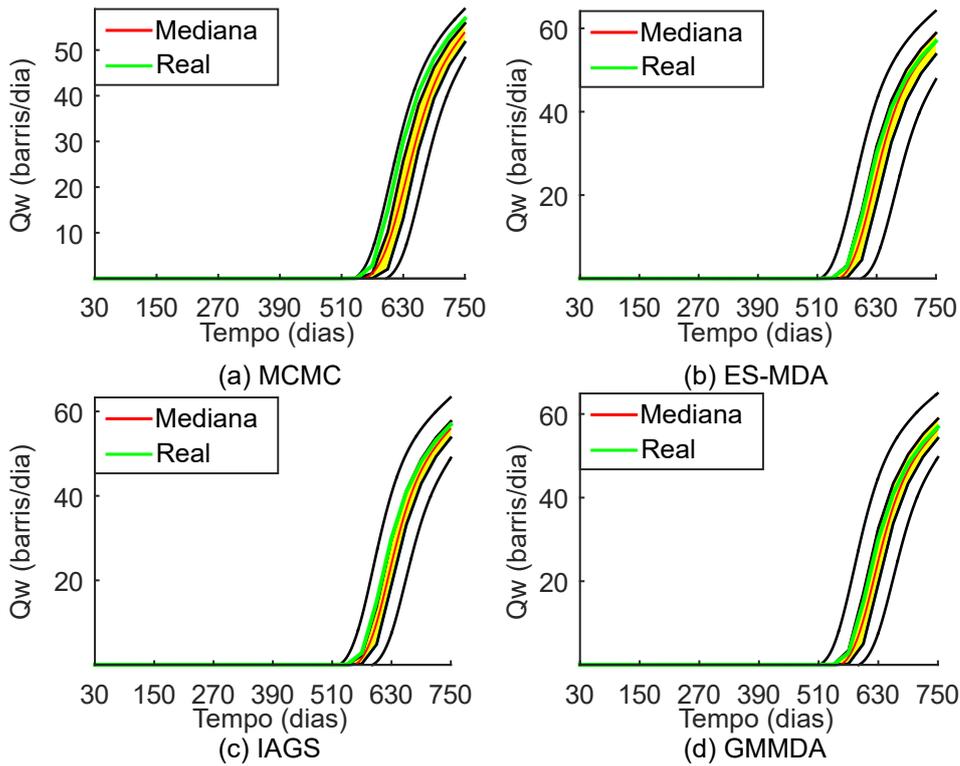


Figura 9: Distribuição *a posteriori* da vazão de produção de água (Q_W) ao final do ajuste de histórico com 10 iterações, para o caso de estudo “1D”.

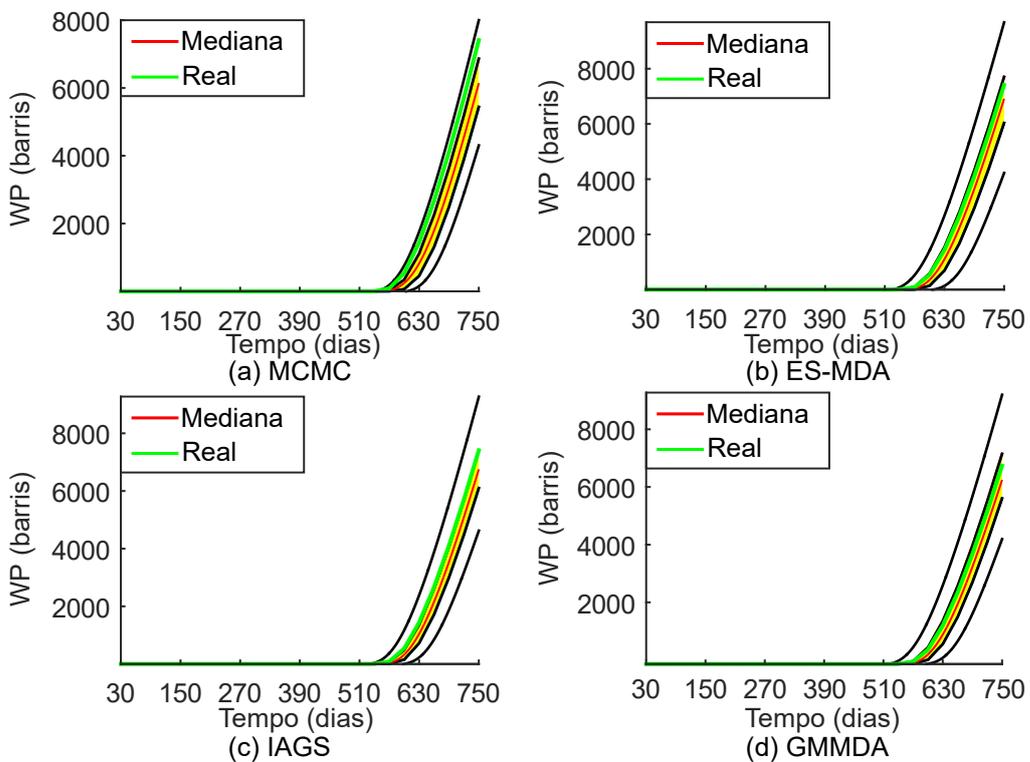


Figura 10: Distribuição *a posteriori* da produção de água acumulada (W_P) ao final do ajuste de histórico com 10 iterações, para o caso de estudo “1D”.

Nas Figuras 9 e 10 são mostradas as distribuições da vazão de produção de água (Q_W) e a produção acumulada de água (W_P), para os três métodos analisados. Além disso, é apresentado o resultado tomado como referência (MCMC).

Fora os resultados apresentados na Figura 8, 9 e 10, foi realizada uma análise prévia para determinar o parâmetro h do método GMMDA. O algoritmo GMMDA requer a definição de dois parâmetros de controle α_k e h . Para o estudo comparativo foi utilizado o α_k igual ao número de iterações, e para determinar o valor do parâmetro h foram feitas simulações variando ele de $h = 0.1$ até $h = 0.9$ em intervalos de 0.1. Na Tabela 5 pode-se observar que o melhor resultado foi obtido para $h = 0.5$, segundo o valor da função objetivo \bar{O}_{N10} calculada com a média das simulações dos 10 conjuntos.

Tabela 5: Variação de h em GMMDA para o Caso de Estudo “1D”

h	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.9
\bar{O}_{N10}	32.80	9.29	6.03	4.80	4.43	5.46	5.95	7.04

Esta função é representada mediante a estimativa máxima *a posteriori* denotada como \mathbf{m}_{MAP} , que é o modelo que maximiza a função de distribuição *a posteriori*, o que equivale minimizar a função $O(\mathbf{m})$ [50]

$$O(\mathbf{m}) = \frac{1}{2}(\mathbf{m} - \mathbf{m}_{pr})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{pr}) + \frac{1}{2}(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}))^T \mathbf{C}_d^{-1}(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m})) , \quad (3-2)$$

onde \mathbf{C}_m e \mathbf{C}_d representam a matriz de covariância do modelo *a priori* e a matriz de covariância dos erros de medições, $\mathbf{g}(\mathbf{m})$ é o vetor de dados simulados, \mathbf{d}_{obs} é o vetor de dados observados e \mathbf{m}_{pr} a média *a priori*. De acordo com Oliver et. al [50], a função objetivo normalizada da Equação 3-2 é calculada mediante o número de medições, pela equação

$$O_N \equiv \frac{O(\mathbf{m})}{N_d} . \quad (3-3)$$

onde \mathbf{m} é o vetor de parâmetros do modelo e N_d é o número de dados de produção observados.

Como pode ser notado nos resultados, os três métodos estudados apresentaram uma favorável estimativa da função de distribuição do logaritmo natural da permeabilidade (Figura 8), estando esses resultados muito perto dos obtidos pelo método MCMC apresentado por Emerick e Reynolds [53], o qual é tomada como referência (Figura 7). O método IAGS apresentou uma melhor aproximação, principalmente no intervalo nas células de 22 a 25, onde

os outros dois métodos apresentam um afastamento maior do valor verdadeiro (linha verde nas figuras). Uma tendência similar está presente nas Figuras 9 e 10, em que, para o método de IAGS, nota-se uma menor dispersão na função de distribuição.

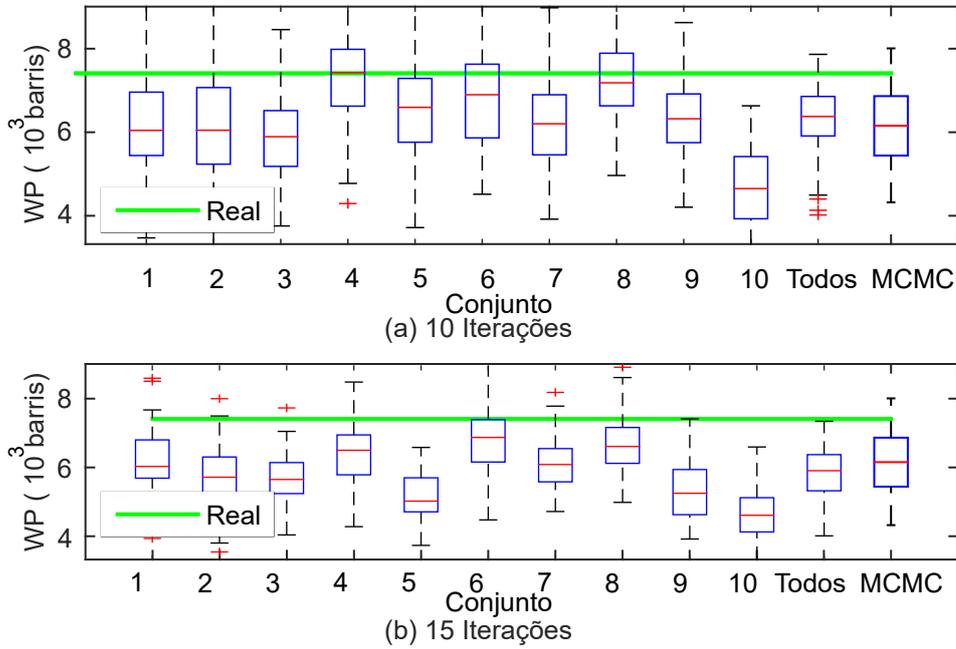


Figura 11: Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de ES-MDA, para o caso de estudo “1D”.

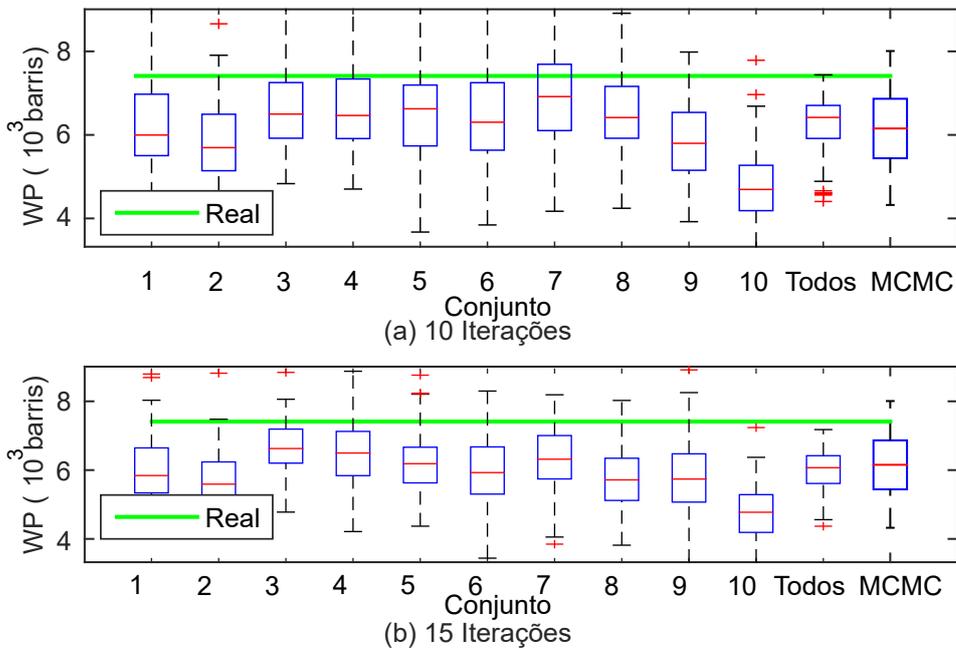


Figura 12: Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de GMMDA, para o caso de estudo “1D”.

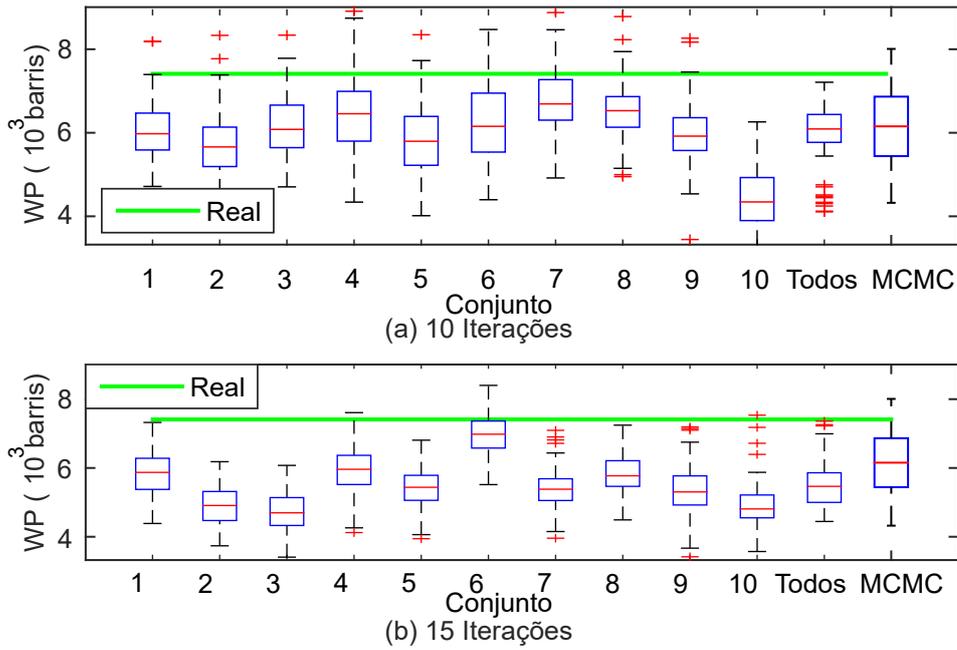


Figura 13: Diagrama de caixas da produção acumulada de água (W_P) ao final do ajuste de histórico mediante o método de IAGS, para o caso de estudo “1D”.

Nas Figuras 11, 12 e 13 apresentam os diagramas de caixa de W_p para cada método considerando 10 e 15 iterações. Os resultados nessas figuras indicam que aumentando o número de iterações melhora os resultados comparados com o MCMC para todos os métodos. Contudo, novamente o IAGS apresentou os melhores resultados. O GMMDA foi apenas um pouco superior ao ES-MDA padrão.

A Figura 14 compara os métodos em termos da função objetivo final. Nessa figura observa-se que aumentando o número de iterações reduz o valor de O_N para todos os métodos. Novamente os resultados para os métodos IAGS e GMMDA foram superiores ao ES-MDA.

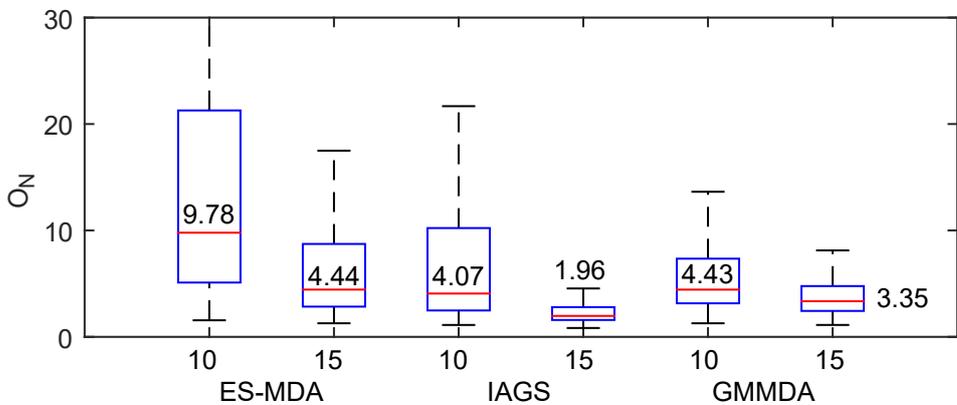


Figura 14: Diagrama de caixas da função objetivo das experiências realizadas no caso de estudo “1D”.

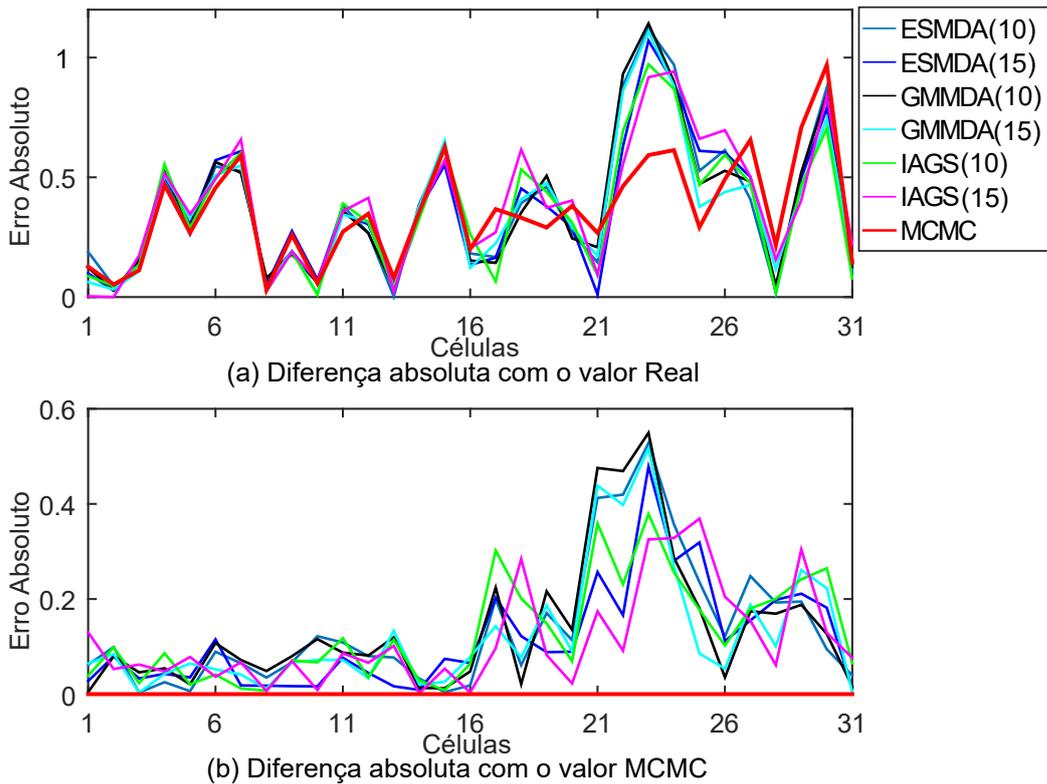


Figura 15: Diferença absoluta do valor do logaritmo natural da permeabilidade com o valor real e com o valor de referência (MCMC), para todas as experiências realizadas no caso de estudo “1D”.

Considerando a diferença absoluta da média dos resultados obtidos com referência ao valor real e ao valor obtido mediante MCMC, o algoritmo IAGS com 10 e 15 iterações obteve melhor desempenho, podendo ser observado na Figura 15. Apesar do fato que todas as experiências obtiveram praticamente o mesmo valor nas células 1 até 16, nota-se um melhor desempenho para as experiências com IAGS nas células 21 até 26.

3.2

Caso de Estudo PUNQ-S3

O segundo caso de estudo onde foram testados os métodos é o modelo PUNQ-S3 [66], sendo um caso de teste muito utilizado na indústria de petróleo. As dimensões do reservatório são de $19 \times 28 \times 5$ células, das quais só 1761 são células ativas. Os parâmetros a ser estimados neste teste são a porosidade (ϕ), permeabilidade horizontal e vertical (κ_I e κ_K , respectivamente) em cada célula. Sendo assim, o vetor de parâmetros a estimar de dimensão 5283 é formado por

$$\mathbf{m} = [\ln(\kappa_I)^T, \ln(\kappa_K)^T, \phi^T]^T. \quad (3-4)$$

3.2.1 Resultados

Analogamente ao caso de estudo “1D”, os testes foram realizados para o mesmo número de iterações e os valores do parâmetro h do IAGS são os correspondentes para o valor de inflação dos métodos de múltiplas assimilações ($\alpha_k = N_a$). Aqui também foi feita uma pré análise para determinar o valor ótimo de $h = 0.9$ no método GMMDA, o qual é apresentado na Tabela 6.

O conjunto inicial de realizações foi construído mediante a simulação sequencial gaussiana usando as informações descritas em [65]. Como neste caso não está facilmente disponível a matriz \mathbf{C}_m , decidiu se comparar os resultados utilizando apenas a parte referente à verossimilhança da função objetivo, ou seja,

$$O(\mathbf{m}) = \frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}})^T \mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}}). \quad (3-5)$$

Assim, o valor de O_N considerando só os dados observados é calculado mediante a Equação 3-3, com o valor da função $O(\mathbf{m})$ modificada.

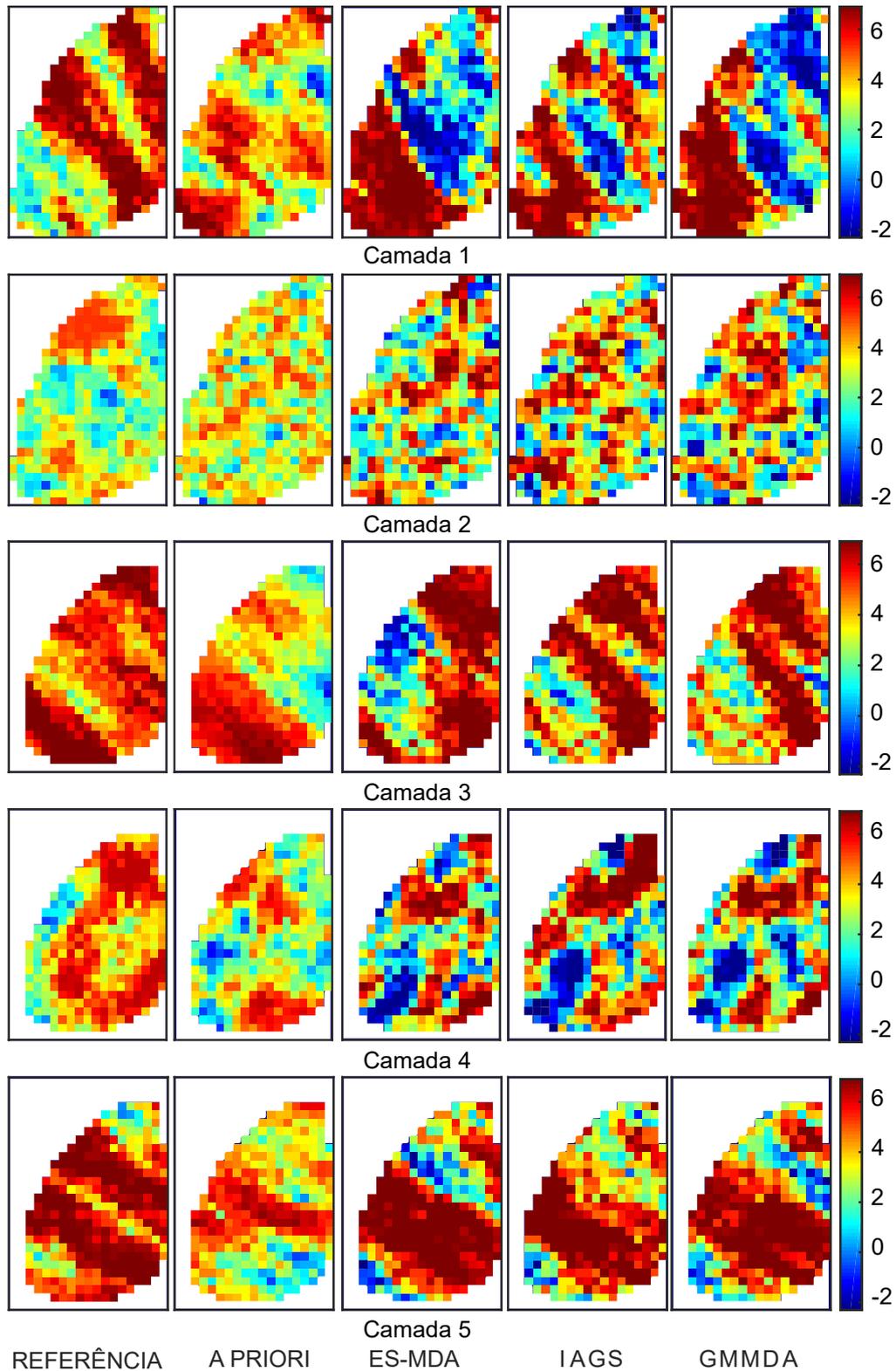
Tabela 6: Variação de h em GMMDA para o Caso de Estudo PUNQ-S3

h	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
\bar{O}_{N10}	1.46	1.22	1.06	1.00	0.94	0.90	0.85	0.84	0.87

Os resultados apresentados foram obtidos utilizando um número de quatro iterações, onde cada teste foi realizado com 10 conjuntos iniciais e 100 amostras em cada conjunto ($N_e = 100$).

As Figuras 16 e 17 mostram as realizações 1 e 33 do logaritmo natural de permeabilidade horizontal do primeiro conjunto, ao final do ajuste de histórico para os três métodos estudados neste capítulo. Além disso, é apresentado na primeira coluna o modelo de referência e na segunda coluna o modelo inicial (*a priori*), de cima para baixo estão os modelos da camada 1 até 5. Como pode-se observar, os modelos de permeabilidade horizontal dos três algoritmos apresentam resultados muito semelhantes, embora tenha uma pequena diferença pelo método IAGS.

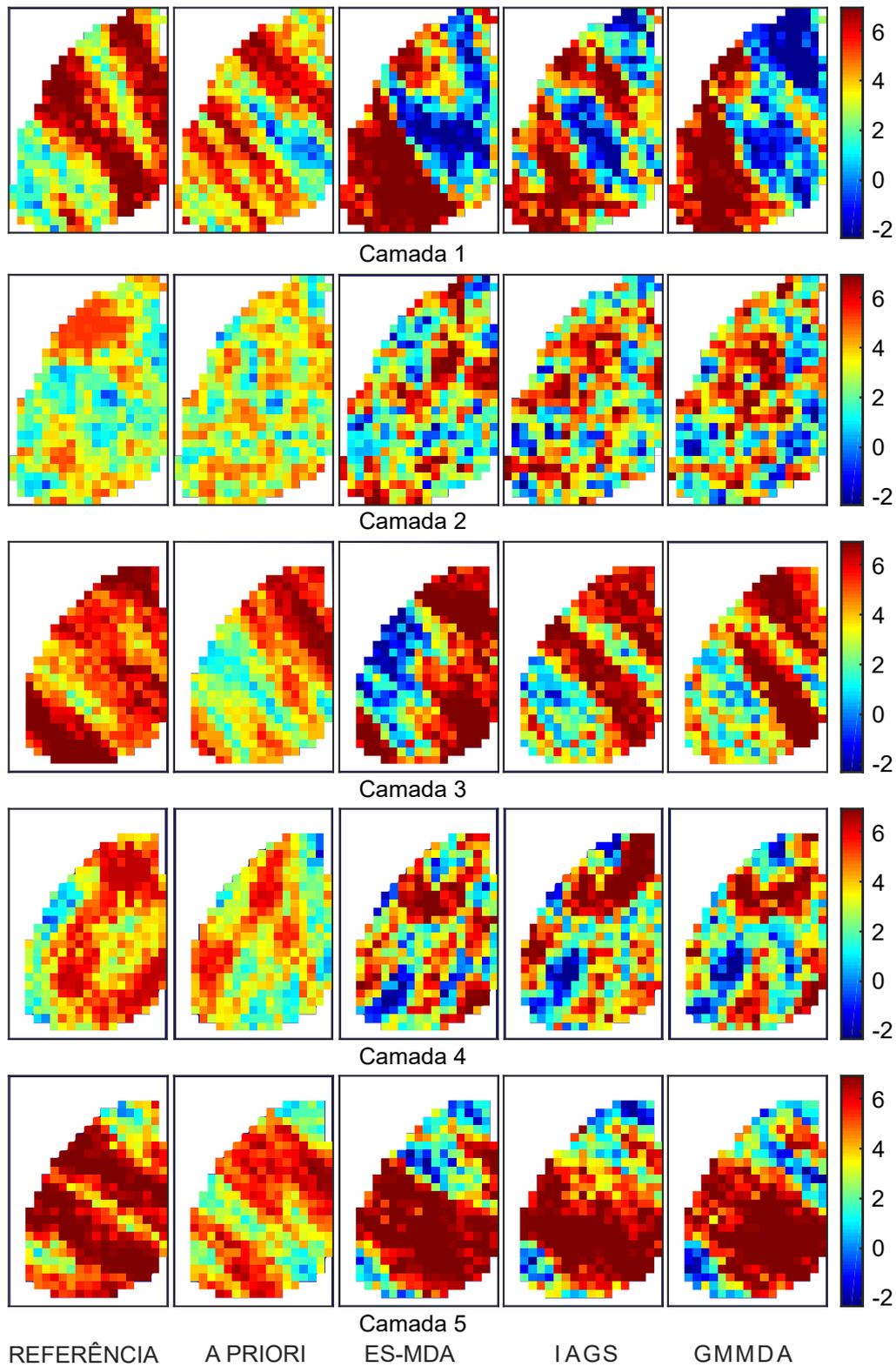
A seguir será apresentado o diagrama de caixas da função objetivo baseada nas observações do acoplamento dos 10 conjuntos para cada método. Na Figura 18 pode-se ver que segundo o valor de \bar{O}_{N10} , o algoritmo de IAGS obteve melhor desempenho, mas em termos gerais os três algoritmos proporcionam resultados muito pertos, quase sem diferenças. A distribuição dos dados (diagrama de caixas) é muito semelhante com valores da mediana de \bar{O}_{N10} em 0.8 para IAGS e aproximadamente 0.85 para os algoritmos restantes.



PUC-Rio - Certificação Digital N° 1313464/CA

Figura 16: Modelos do logaritmo natural da permeabilidade horizontal da primeira realização do primeiro conjunto, para o caso de estudo PUNQ-S3.

A Figura 19 apresenta as curvas de produção durante todo o período de coleta de dados para dois indicadores, pressão no poço produtor $P1$ na parte esquerda e a razão gás-óleo (RGO) do poço $P5$ no lado direito. A parte de



PUC-Rio - Certificação Digital Nº 1313464/CA

Figura 17: Modelos do logaritmo natural da permeabilidade horizontal da realização 33 do primeiro conjunto, para o caso de estudo PUNQ-S3.

cima da Figura 19 corresponde aos resultados obtidos mediante ES-MDA, no centro ao algoritmo IAGS e na parte de baixo para GMMDA. Como pode ser observado, os três métodos conseguiram ajustar os dados de produção

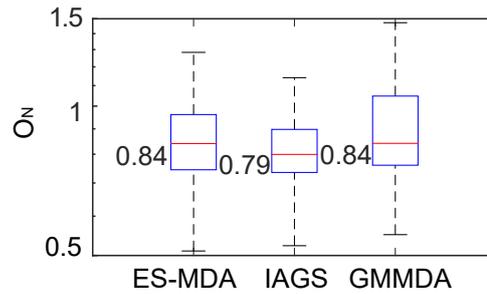


Figura 18: Diagrama de caixas da função objetivo do caso de estudo PUNQ-S3.

observados (pontos vermelhos). As curvas azuis correspondem aos modelos finais e as curvas cinzas correspondem aos modelos *a priori*, e a média das curvas ajustadas estão indicadas mediante a linha verde. Na curva de *RGO* pode ser percebido que todos os modelos ao final do ajuste tendem a uma única curva.

3.3

Caso de Estudo BRUGGE

O último caso de estudo corresponde ao modelo de reservatório BRUGGE, com 20 produtores e 10 injetores. Esse caso de estudo foi desenvolvido como uma referência na indústria do petróleo para testar métodos de ajuste de histórico e otimização. Uma descrição completa do caso de estudo BRUGGE é apresentada por Peters et. al [94]. As dimensões do reservatório são de $139 \times 48 \times 9$ células, dos quais 44550 células estão ativas.

Os parâmetros a serem estimados neste caso de estudo são o logaritmo natural de permeabilidade horizontal e vertical, ($\ln(\kappa_I)$ e $\ln(\kappa_K)$), a porosidade (ϕ), a razão *Net-to-Gross* (NTG) e a saturação da água (S_{wc}) em cada célula. Sendo assim, o vetor a estimar está formado por

$$\mathbf{m} = [\ln(\kappa_I)^T, \ln(\kappa_K)^T, \phi^T, NTG^T, S_{wc}^T]^T, \quad (3-6)$$

onde \mathbf{m} é de dimensão 222750. Para este caso de estudo foi utilizado apenas conjunto com 104 realizações. Nesse caso de estudo foi utilizada a técnica de localização do ganho de Kalman [135].

3.3.1

Resultados

Para os testes de ajuste de histórico foram utilizadas medições a cada 120 dias pelo período de 10 anos. Neste caso de estudo foram realizadas quatro iterações para comparação dos métodos estudados. Primeiro foram feitos testes no algoritmo GMMDA para obter o ótimo valor de h . A Figura 20 apresenta

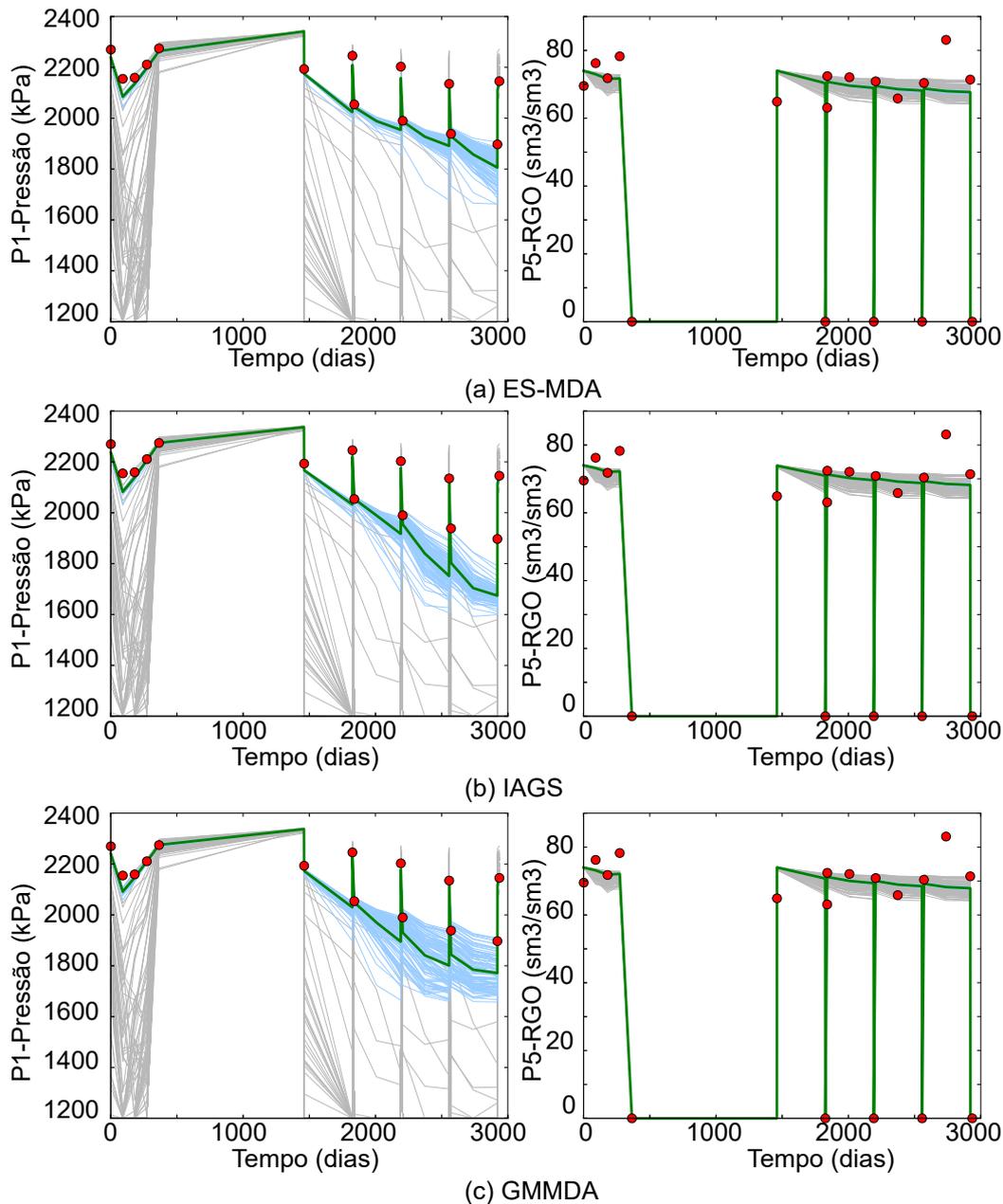


Figura 19: Curvas de produção dos modelos iniciais (cinza) e finais (azul). A coluna da esquerda apresenta a pressão no poço produtor P1. A coluna da direita apresenta a razão gás-óleo (RGO) do poço produtor P5, para o caso de estudo PUNQ-S3.

o diagrama de caixas para o valor da função objetivo variando o valor de h , obtendo o valor ótimo para $h = 0.7$.

Alguns modelos ao final do ajuste de histórico são apresentados nas Figuras 21 e 22. Esses modelos correspondem ao logaritmo da permeabilidade horizontal na primeira camada, além de mostrar o modelo *a priori* do qual foi iniciado na parte superior esquerda. Como pode ser visto aqui, os modelos finais são semelhantes. A Figura 23 apresenta o desvio padrão do conjunto final, e aqui pode-se observar que na parte central dos modelos o desvio é

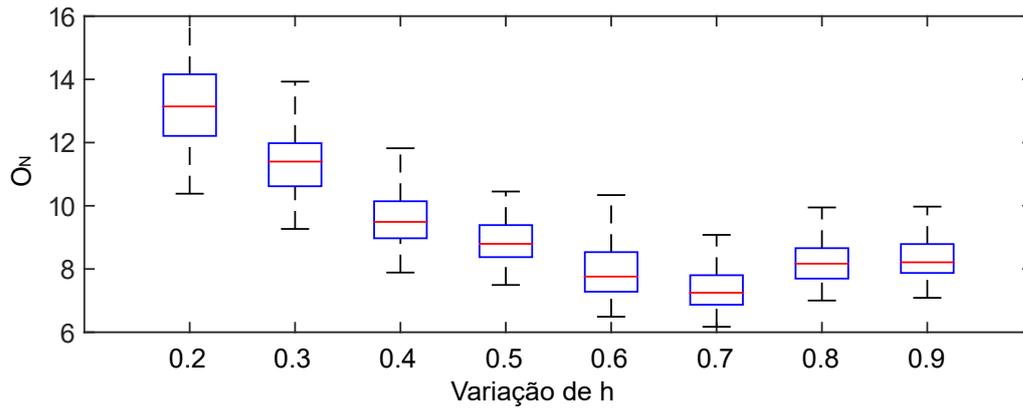


Figura 20: Diagrama de caixas da função objetivo para o caso de estudo BRUGGE mediante o método GMMDA variando o parâmetro de controle h .

muito baixo, devido ao fato de que nessa área estão encontrados os poços de produção onde foram coletados os dados.

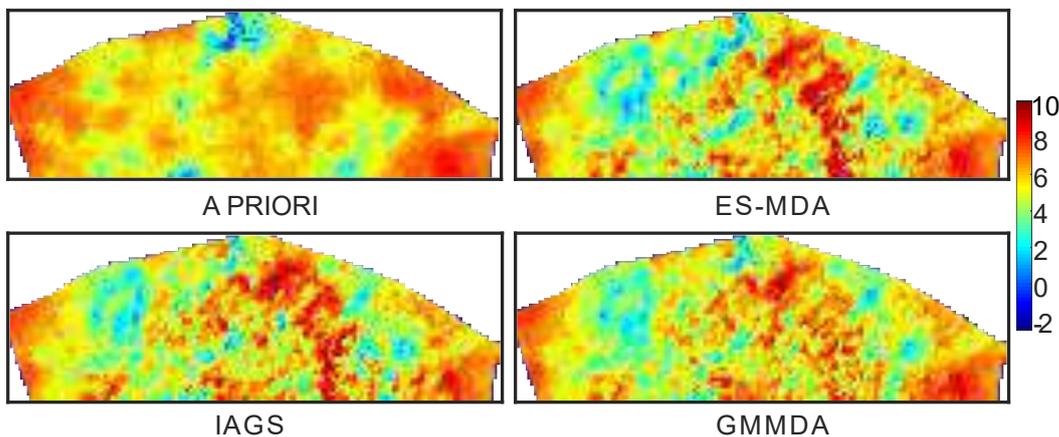


Figura 21: Modelo de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE, correspondente ao membro número 1 do conjunto.

A comparação das distribuições finais da função objetivo dos três métodos é apresentada na Figura 24, onde pode-se ver que os três algoritmos apresentam distribuições muito semelhantes, e os valores médios obtidos são aproximadamente 7.2 para todos. Embora apresenta-se uma pequena diferença no método de ES-MDA, que apresenta menor dispersão no diagrama de caixas da função objetivo. Por último são apresentadas as curvas de produção (Figura 25 e 26), no início do ajuste de histórico e no final. Cada produção gerada pelo conjunto inicial está apresentada nas figuras pelas linhas cinzas, e as azuis são a produção do modelo ajustado. Aqui também pode-se ver os dados observados mediante os pontos vermelhos e a média dos modelos ajustados está apresentada pela linha verde. Em todas as curvas pode-se observar que o

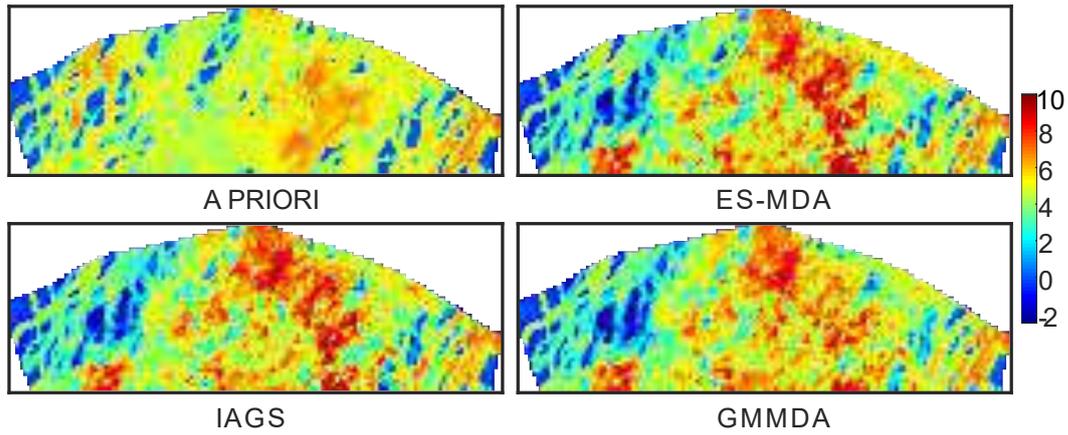


Figura 22: Modelo de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE, correspondente ao membro número 33 do conjunto.

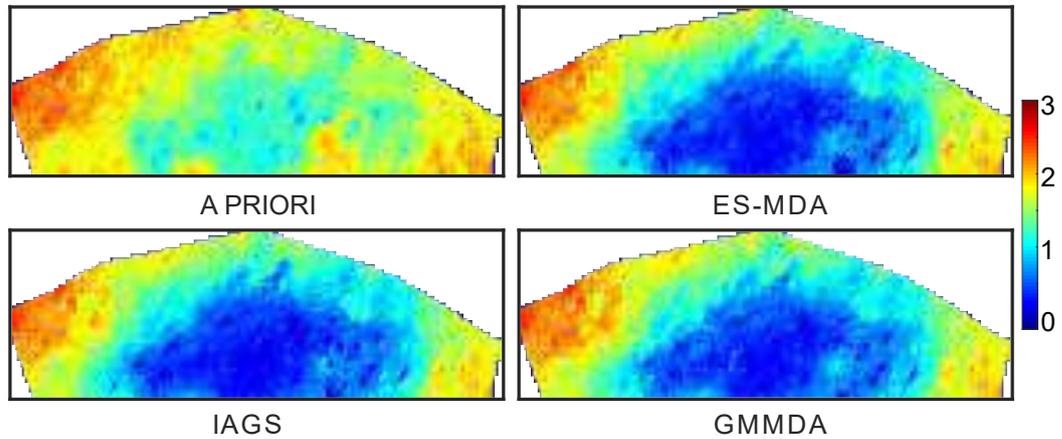


Figura 23: Modelos de desvio padrão de logaritmo natural da permeabilidade horizontal na primeira camada do caso de estudo BRUGGE.

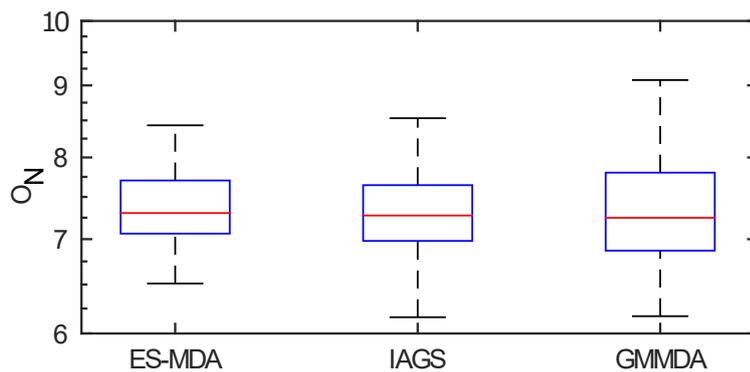


Figura 24: Diagrama de caixas da função objetivo dos três métodos estudados, para o caso de estudo BRUGGE.

ajuste foi de qualidade bastante satisfatória, com uma pequena dispersão na distribuição maior no método GMMDA que está na parte inferior das Figuras 25 e 26. No centro se encontram os resultados mediante o método de IAGS, e na parte superior do ES-MDA.

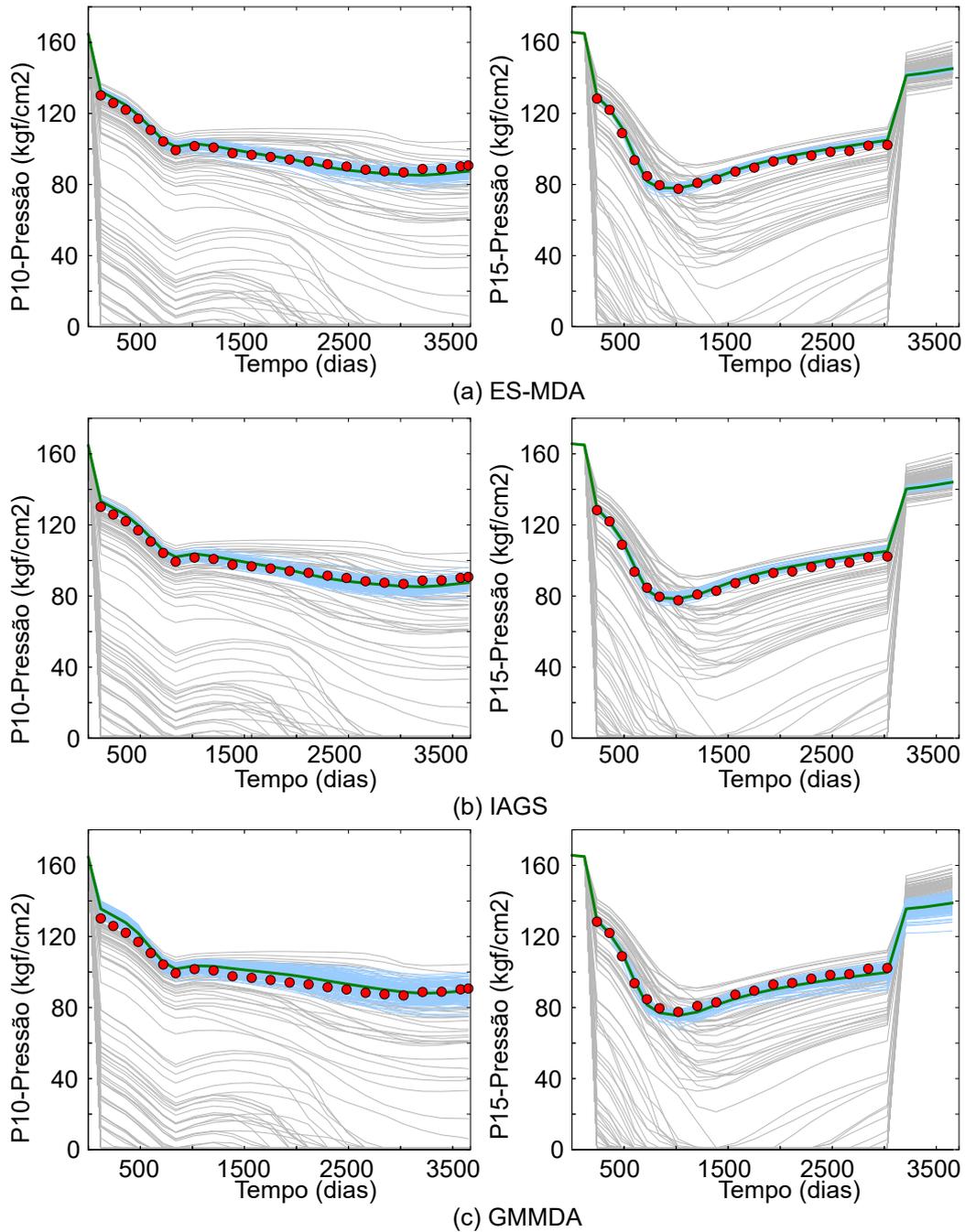


Figura 25: Curvas de pressão dos modelos iniciais (cinza) e finais (azul) para os poços P10 e P15, para o caso de estudo BRUGGE.

O estudo comparativo entre o ES-MDA, IAGS e GMMDA mostrou que os métodos IAGS e GMMDA tem um desempenho superior ao ES-MDA quando o modelo de reservatório tem dimensões reduzidas. De fato, para o caso de estudo “1D”, os resultados obtidos pelo ES-MDA apresentaram valores de função objetivo maiores. Note que a influência dos pesos atribuídos nas metodologias IAGS e GMMDA, para lidar com a hipótese de gaussianidade dos métodos baseados no filtro de Kalman, não faz muita diferença em modelos de maior dimensão.

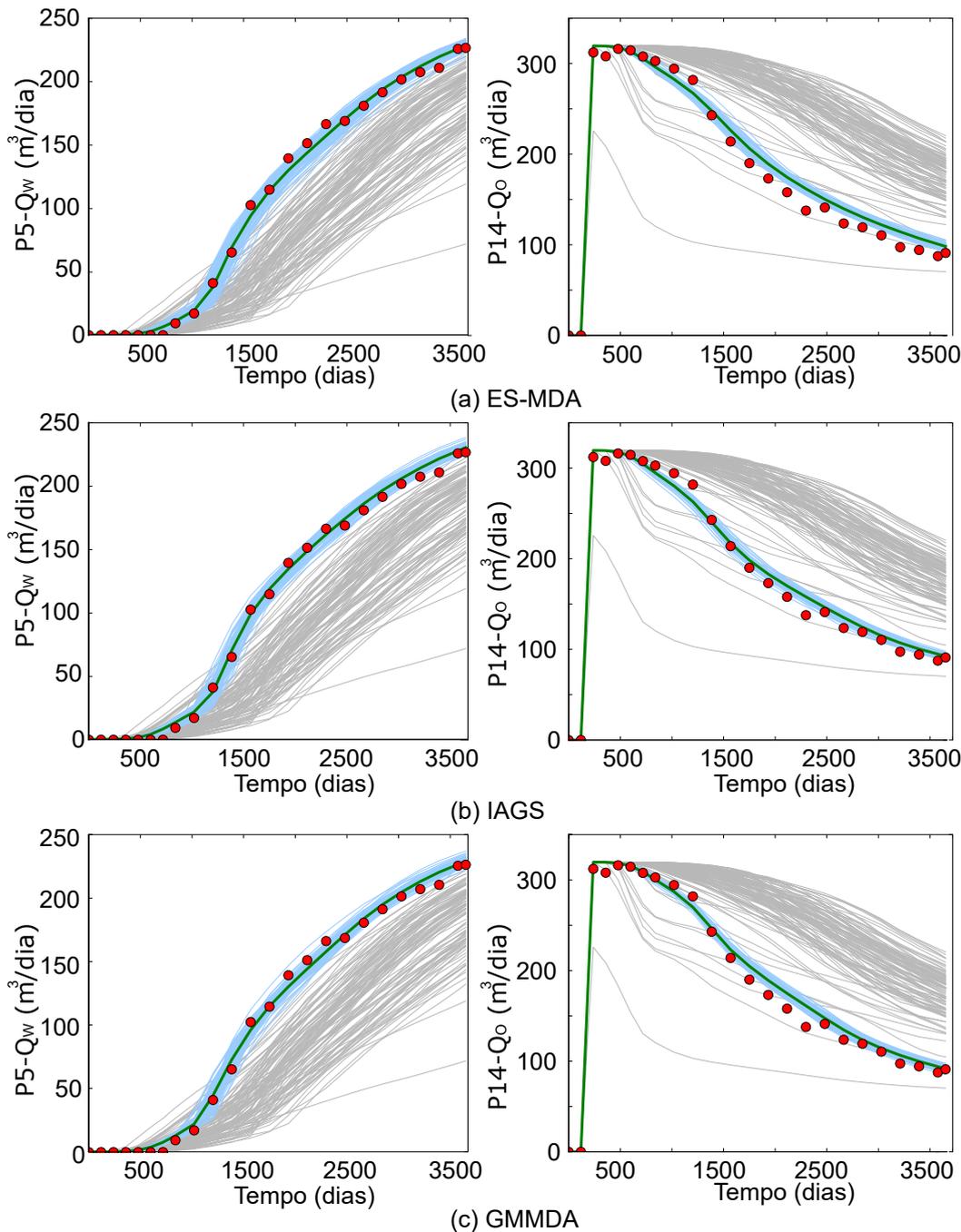


Figura 26: Curvas de produção dos modelos iniciais (cinza) e finais (azul). A coluna da esquerda apresenta a vazão de água (Q_w) do poço P5. A coluna da direita apresenta a vazão de óleo (Q_o) do poço P14, para o caso de estudo BRUGGE.

Além disso, as metodologias IAGS e GMMDA requerem a definição do parâmetro h . Os resultados mostram que a escolha adequada desse parâmetro varia de acordo com cada modelo de reservatório, o que dificulta sua aplicação direta. Nos testes apresentados neste capítulo, foram escolhidos os valores ótimos para cada caso de estudo. Contudo, isso não seria adequado para casos reais, uma vez que o cálculo de h ótimo requer a repetição da assimilação

diversas vezes.

Por outro lado, a medida que a dimensão do modelo aumenta, o efeito dos pesos atribuídos a cada membro do conjunto dos métodos que utilizaram a IS começam a diminuir ao ponto que para os outros dois casos de estudo tratados, PUNQ-S3 e BRUGGE, os três métodos tiveram resultados bastante semelhantes.

A formulação simples e a aplicação prática do ES-MDA motivou a escolha desse método, conjuntamente com técnicas de parametrização para resolver problemas de maior complexidade, como é o caso de ajuste de histórico de modelos de fácies geológicas. Esta metodologia será integrada com duas técnicas de parametrização propostas nesta tese, para atualizar modelos de fácies geológicas conjuntamente com as propriedades petrofísicas do reservatório, conservando as características geológicas e geoestatísticas do reservatório.

4

Metodologia de Atualização de Fácies Baseado em Análise de Componentes Principais

Os métodos baseados no filtro de Kalman assumem que a distribuição dos modelos *a priori* é gaussiana. Isso torna particularmente desafiador o ajuste de modelos de fácies, uma vez que as fácies assumem valores discretos, normalmente em termos de categorias, por exemplo “areia”, “interlaminado” ou “folhelho”.

Este capítulo introduz uma parametrização baseada em PCA e simulações plurigaussiana truncada, desenvolvida para permitir o ajuste de histórico de modelos de fácies usando o método ES-MDA. No restante deste texto esse método será denominado ES-MDA-PCA.

4.1

Procedimento

A ideia básica do método proposto é transformar as fácies em parâmetros contínuos seguindo distribuições aproximadamente gaussianas e atualizar esses parâmetros com o método ES-MDA. Após cada ciclo de atualização os parâmetros contínuos são convertidos em fácies usando a transformação inversa. A transformação é feita utilizando dois elementos principais: simulação gaussiana truncada [71] e PCA. O procedimento completo é composto por dez passos ilustrados na Figura 27 e descrito a seguir:

1. Inicialização:

O processo inicia com a geração de um conjunto inicial com N_r realizações de modelos de fácies, aqui representada pelo conjunto $\{\mathbf{f}_j\}_{j=1}^{N_r}$, onde \mathbf{f}_j é o vetor contendo o valor das fácies para cada célula do modelo j . Essas realizações de fácies serão usadas para construir a transformação de fácies para parâmetros contínuos e a transformação inversa. Portanto, é importante usar um número relativamente grande de realizações, por exemplo, $N_r = 5000$. O processo de geração dessas realizações iniciais é feito com procedimentos usuais de geoestatística.

Cada realização \mathbf{f}_j é subdividida em N_f realizações binárias, onde N_f é o número de fácies do modelo. Essas realizações binárias \mathbf{b}_i , para

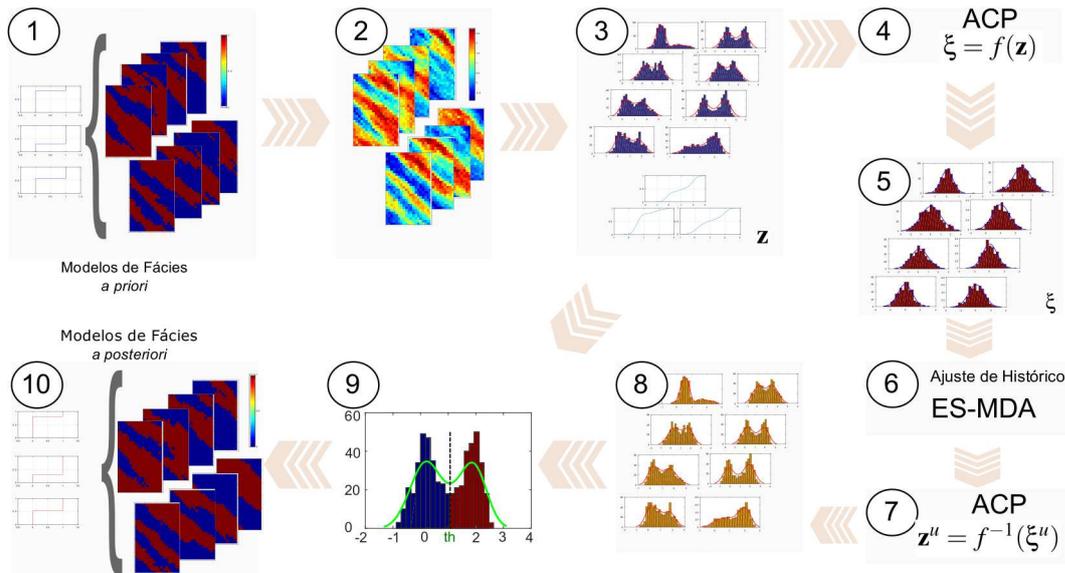


Figura 27: Procedimento do método ES-MDA-PCA.

$i = 1, 2, \dots, N_f$, contém o valor de 1 se a fácies for i e zero se a fácies for diferente de i . A Figura 28 ilustra esse procedimento para uma realização de três fácies.

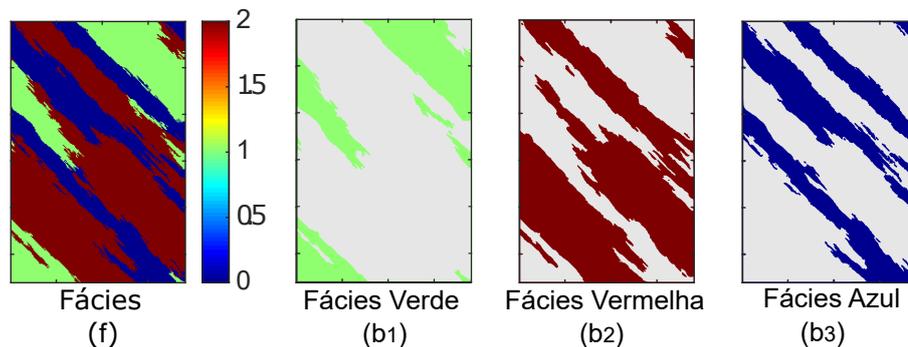


Figura 28: Partição de um modelo com três fácies.

2. Transformação Usando Simulação Gaussiana:

A seguinte etapa do processo (passo 2 na Figura 27) corresponde a transformar cada realização binária em uma realização contínua. Para isso utilizou-se o algoritmo de simulação sequencial gaussiana condicionada [96] com média zero e variância igual a um usando as realizações binárias como variáveis secundárias.

A Figura 29 mostra um exemplo onde uma realização binária \mathbf{b}_i (esquerda) é transformada em uma realização contínua \mathbf{z}_i (centro) e o histograma resultante (direita). Note que o histograma resultante ainda preserva uma característica bi-modal, porém possui uma variação contí-

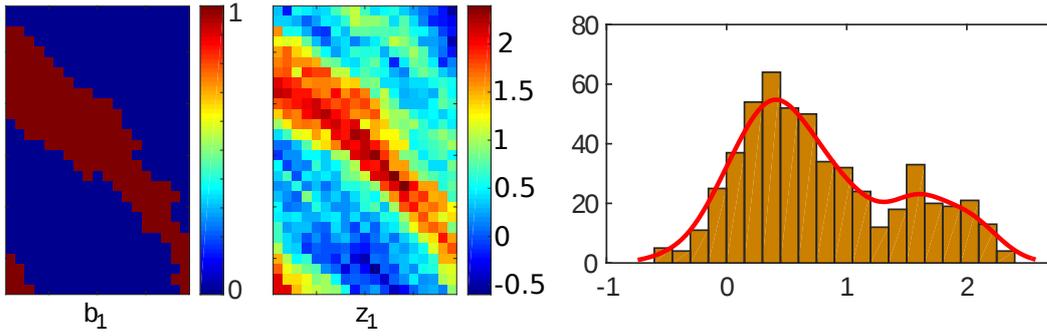


Figura 29: Simulação sequencial gaussiana condicionada à fácies binárias.

nua de valores. Observa-se que ao final do passo 2 tem-se um total de $N_r \times N_f$ realizações \mathbf{z}_{ij} com $i = 1, \dots, N_f$ e $j = 1, \dots, N_r$.

3. Definição de Regras de Truncamento:

A terceira etapa é definir de forma automática as regras de truncamento que serão usados na transformação inversa das realizações contínuas \mathbf{z}_{ij} para realizações binárias \mathbf{b}_{ij} , que posteriormente são transformadas em fácies \mathbf{f}_j . As regras de truncamento são calculadas a partir de um processo de otimização usando a distribuição acumulada das variáveis \mathbf{z}_{ij} . Esse processo de otimização busca minimizar a seguinte função objetivo

$$FO(t) = \sum_{j=1}^{N_r} \sum_{k=1}^{N_g} |b_{kj} - \delta_{kj}(t)|, \quad (4-1)$$

onde b_{kj} corresponde ao valor na k -ésima célula da realização \mathbf{b}_j e $\delta_{kj}(t)$ é dado por

$$\delta_{kj}(t) = \begin{cases} 1 & \text{se } z_{kj} > [FDC_{\mathbf{z}_j}(t)]^{-1} \\ 0 & \text{caso contrário,} \end{cases} \quad (4-2)$$

sendo N_r o número de realizações e N_g o número de células ativas do modelo de reservatório. $FDC_{\mathbf{z}_j}(t)$ corresponde à função de densidade acumulada do vetor \mathbf{z}_j , onde $t \in [0, 1]$.

A minimização da Equação 4-1 é repetida para cada fácies. Ao final do processo, é obtido um valor de t_i^* que minimize a Equação 4-1 para $i = 1, \dots, N_f$. A otimização pode ser feita com qualquer algoritmo, nesta tese foi utilizada o algoritmo de otimização por enxame de partículas da ferramenta de otimização do Matlab.

4. Análise de Componentes Principais:

A quarta etapa do processo é transformar as variáveis contínuas \mathbf{z}_{ij} em variáveis com distribuição normal padrão, ou seja, $\xi_{ij} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Isso é feito aplicando o algoritmo de PCA. Para cada tipo de fácies é calculada

a média e a covariância usando

$$\bar{\mathbf{z}}_i = \frac{1}{N_r} \sum_{j=1}^{N_r} \mathbf{z}_{ij} , \quad (4-3)$$

$$\mathbf{C}_{\mathbf{z}_i} = \frac{1}{N_r - 1} \sum_{j=1}^{N_r} (\mathbf{z}_{ij} - \bar{\mathbf{z}}_i)(\mathbf{z}_{ij} - \bar{\mathbf{z}}_i)^T . \quad (4-4)$$

O PCA é aplicada à matriz $\mathbf{C}_{\mathbf{z}_i}$ usando decomposição em valores singulares de tal forma que

$$\mathbf{C}_{\mathbf{z}_i} = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i , \quad (4-5)$$

onde \mathbf{U}_i é a matriz que contém os vetores singulares de $\mathbf{C}_{\mathbf{z}_i}$ e $\mathbf{\Sigma}_i$ é a matriz que contém os valores singulares na sua diagonal. A partir da decomposição de $\mathbf{C}_{\mathbf{z}_i}$ pode se escrever cada realização \mathbf{z}_{ij} como

$$\mathbf{z}_{ij} = \bar{\mathbf{z}}_i + \mathbf{C}_{\mathbf{z}_i}^{1/2} \xi_i , \quad (4-6)$$

onde $\mathbf{C}_{\mathbf{z}_i}^{1/2}$ pode ser definido em termos de \mathbf{U}_i e $\mathbf{\Sigma}_i$, aqui se escolheu definir uma $\mathbf{C}_{\mathbf{z}_i}^{1/2}$ simétrica como

$$\mathbf{C}_{\mathbf{z}_i}^{1/2} = \mathbf{U}_i \mathbf{\Sigma}_i^{-1/2} \mathbf{U}_i^T . \quad (4-7)$$

Essa escolha faz com que os vetores ξ_i na Equação 4-6 tenham a mesma dimensão de \mathbf{z}_{ij} . Isso é importante para permitir o uso de técnicas de localização no ajuste com ES-MDA [21]. A partir da Equação 4-6 calcula-se o vetor ξ_i como

$$\begin{aligned} \xi_i &= \mathbf{C}_{\mathbf{z}_i}^{-1/2} (\mathbf{z}_{ij} - \bar{\mathbf{z}}_i) \\ &= \mathbf{U}_i \mathbf{\Sigma}_i^{-1/2} \mathbf{U}_i^T (\mathbf{z}_{ij} - \bar{\mathbf{z}}_i) . \end{aligned} \quad (4-8)$$

É importante observar que o cálculo da decomposição de $\mathbf{C}_{\mathbf{z}_i}$ pode ser feita de forma mais eficiente, levando em consideração que $\mathbf{C}_{\mathbf{z}_i}$ foi calculada usando a Equação 4-3 conforme foi descrito no Capítulo 2 Seção 2.3.1.

A transformação da Equação 4-8 gera vetores ξ_i que seguem aproximadamente uma distribuição normal padrão, ou seja, $\xi_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Essa transformação faz com que ξ_i seja apropriado para o uso no método ES-MDA.

5. Atualização do Modelo com ES-MDA:

O quinto passo de processo corresponde à aplicação do método ES-MDA

para a incorporação das informações dinâmicas (dados de produção). O vetor de parâmetros do modelo usado no ES-MDA, \mathbf{m}_j contém os vetores ξ_i , com $i = 1, \dots, N_f$ além de outras propriedades petrofísicas como porosidade e permeabilidade, ou seja o vetor de parâmetros é formado por

$$\mathbf{m}_j = [\xi_1^T, \dots, \xi_{N_f}^T, \phi_1^T, \dots, \phi_{N_f}^T, \ln \kappa_1^T, \dots, \ln \kappa_{N_f}^T]^T . \quad (4-9)$$

No método ES-MDA utilizou-se um conjunto menor de realizações, $N_e < N_r$, uma vez que será necessário simular o período de produção para cada realização. Além disso, utilizou-se a técnica de localização para reduzir o efeito deletério das correlações espúrias.

6. Transformação Inversa:

O sexto passo corresponde à transformação dos vetores ξ_{ij} em \mathbf{z}_{ij}

$$\mathbf{z}_i = f^{-1}(\xi_i) . \quad (4-10)$$

Está transformação é feita utilizando a Equação 4-6.

7. Truncamento e Reconstrução de Fácies:

O último passo corresponde à reconstrução dos modelos de fácies a partir dos vetores \mathbf{z}_{ij} e as regras de truncamento definidas no passo 3. Assim, são recalculadas as realizações binárias \mathbf{b}_{ij} que por sua vez dão origem às fácies \mathbf{f}_j .

4.2

Aplicação no Modelo PUNQ-S3

Para realizar o teste de atualizar as fácies, foi criado um caso de estudo baseado no modelo de reservatório PUNQ-S3, onde utilizando a ferramenta de geoestatísticas, *Stanford Geostatistical Modeling Software* (SGeMS) [96], com a qual baseado nos variogramas disponíveis na literatura para o reservatório PUNQ-S3 [66], foram geradas duas fácies por cada camada, e baseadas nestas fácies foram gerados modelos de porosidade e permeabilidade horizontal e vertical, de acordo com a Tabela 7. As propriedades nas camadas 1,3 e 5 apresentam as mesmas características.

De acordo com os valores da Tabela 7 foram geradas simulações gaussianas para gerar propriedades de porosidade, permeabilidade horizontal e vertical. Posteriormente as propriedades foram agrupadas para gerar as realizações petrofísicas. Este procedimento de agrupar as propriedades é baseado na distribuição espacial da fácies. Em total foram geradas 5000 realizações para cada camada e dessas foi selecionada uma para ser a referência de um

Tabela 7: PUNQ-S3: A Média e o Desvio Padrão das Propriedades Petrofísicas por Tipo de Fácies

		Porosidade		Log-Perm. Horizontal		Log-Perm. Vertical	
		Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
Camada 1	Mudstone	0.1	0.031	1.8	0.224	1.5	0.224
	Channel	0.23	0.031	2.5	0.224	2.2	0.224
	Marine	–	–	–	–	–	–
	Mouthbar	–	–	–	–	–	–
		Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
Camada 2	Mudstone	0.035	0.031	0.8	0.283	0.4	0.173
	Channel	–	–	–	–	–	–
	Marine	0.1	0.045	1.7	0.224	1.1	0.173
	Mouthbar	–	–	–	–	–	–
		Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
Camada 4	Mudstone	0.06	0.071	1.4	0.224	1.1	0.173
	Channel	–	–	–	–	–	–
	Marine	–	–	–	–	–	–
	Mouthbar	0.15	0.031	2.2	0.224	1.6	0.173

PUC-Rio - Certificação Digital Nº 1313464/CA

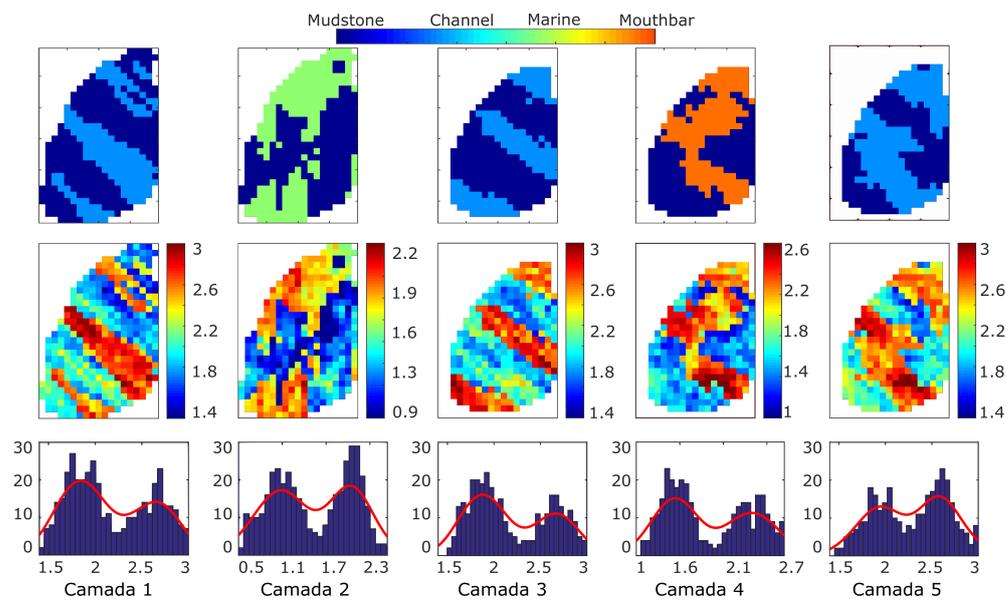
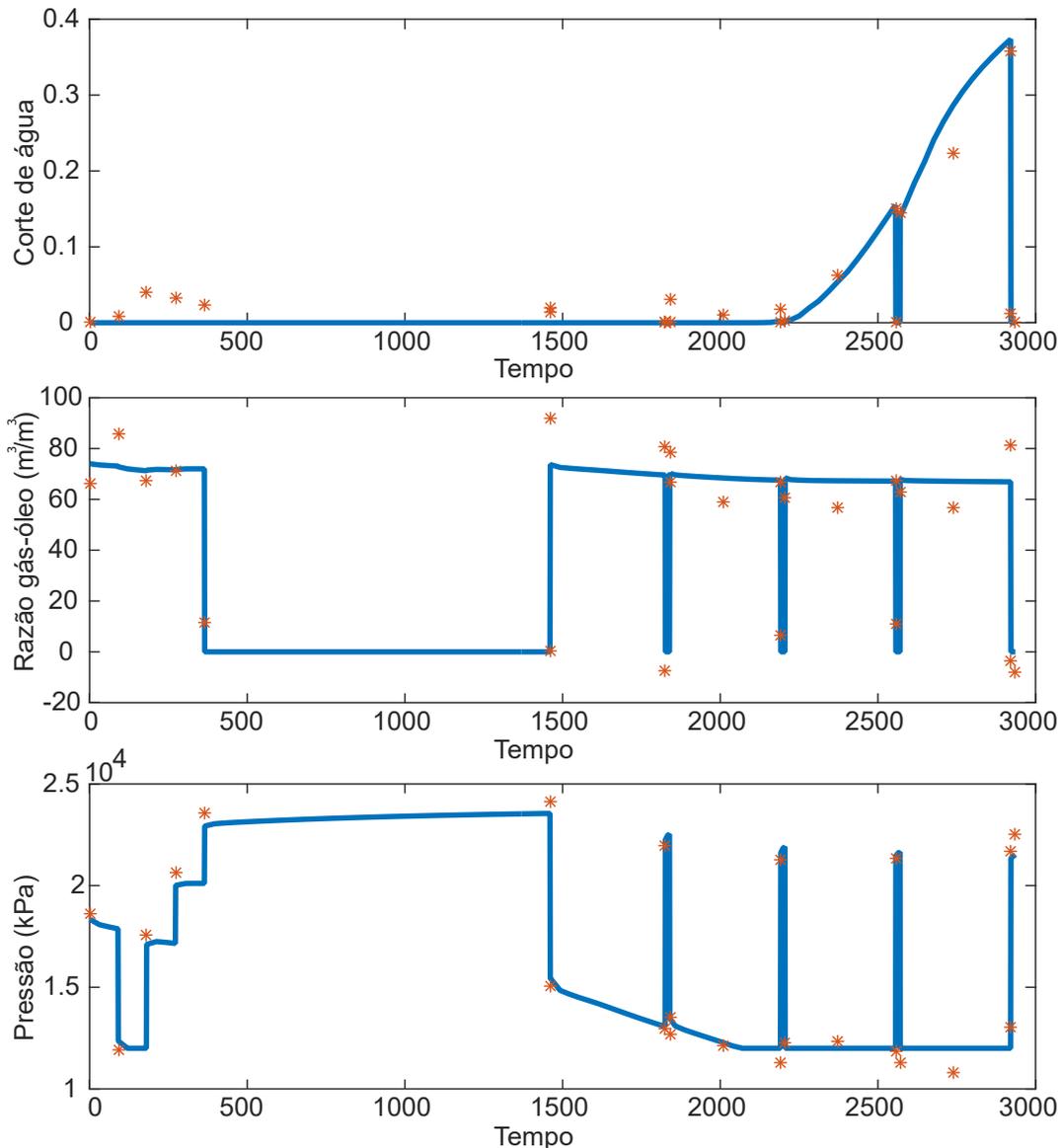


Figura 30: Modelo de fácies de referência na linha superior, a linha central o logaritmo da permeabilidade e a linha inferior os histogramas correspondentes.

modelo real, este modelo escolhido é apresentado na Figura 30, onde pode ser visto a imagem de fácies escolhida como referência para cada camada (na parte superior), no centro é mostrado o correspondente logaritmo natural da permeabilidade horizontal de cada camada e seus respectivos histogramas na parte inferior. Note que a distribuição do logaritmo natural da permeabilidade horizontal é bi-modal. Além da permeabilidade horizontal, foram geradas realizações de porosidade e permeabilidade vertical.



PUC-Rio - Certificação Digital N° 1313464/CA

Figura 31: Curvas de produção do produtor 11 (curvas azuis). Histórico de produção (asteriscos vermelhos).

Com os modelos de porosidade e permeabilidade horizontal e vertical, o reservatório foi simulado utilizando o simulador da *Computer Modelling Group* (CMG), IMEX. Na Figura 31 a linha azul apresenta a curva de corte de água na linha superior da figura, razão gás-óleo (RGO) no centro e a curva

Tabela 8: Ruído para as Observações

	Pressão	RGO	Corte de água
σ	500 kPa	10	0.02

de pressão na parte de baixo, todos eles do produtor 11. Para a geração das medições, foi acrescentado um ruído gaussiano de média zero e desvio padrão, σ , de acordo com a Tabela 8. Assim na Figura 31 os asteriscos vermelhos são o resultado dos valores da curva real, acrescentado o ruído respectivo.

Com o modelo de teste apresentado, serão realizados dois experimentos. Primeiramente será utilizado o método ES-MDA para atualizar somente as propriedades petrofísicas do caso de estudo, seguidamente será atualizado os modelos de fácies conjuntamente com as propriedades utilizando a método proposto, ES-MDA-PCA. A Figura 32 mostra as três primeiras realizações do conjunto inicial a ser ajustado. Essas realizações pertencem aos modelos de fácies geológicas (parte superior), as realizações do logaritmo natural de permeabilidade horizontal no centro e seu respectivo histograma na parte inferior.

4.2.1 ES-MDA

Primeiramente foi feito um experimento utilizando o ES-MDA padrão, utilizando quatro assimilações, $N_a = 4$. O teste é feito para comparação dos resultados com a metodologia proposta neste capítulo (ES-MDA-PCA), que são apresentados mais adiante.

As propriedades da rocha a atualizar neste caso de estudo são os modelos de porosidade (ϕ) e permeabilidade horizontal e vertical (κ_h e κ_v). Assim, o vetor de parâmetros \mathbf{m} , para atualizar mediante o método ES-MDA (Equação 2-29) é formado por

$$\mathbf{m} = [\phi^T, \log_{10}(\kappa_h)^T, \log_{10}(\kappa_v)^T]^T . \quad (4-11)$$

Para atualizar esses modelos mediante ES-MDA foram utilizadas as 200 primeiras realizações ($N_e = 200$) geradas para este caso de estudo. Além disso, para o experimento não foi utilizado a abordagem de localização do ganho de Kalman, pois o caso de estudo é relativamente simples. O resultados são ilustrados na Figura 33 e mostram as 3 primeiras realizações da permeabilidade horizontal da primeira camada, depois do ajuste de histórico. Também são apresentados seus histogramas correspondentes. Note-se que algumas das

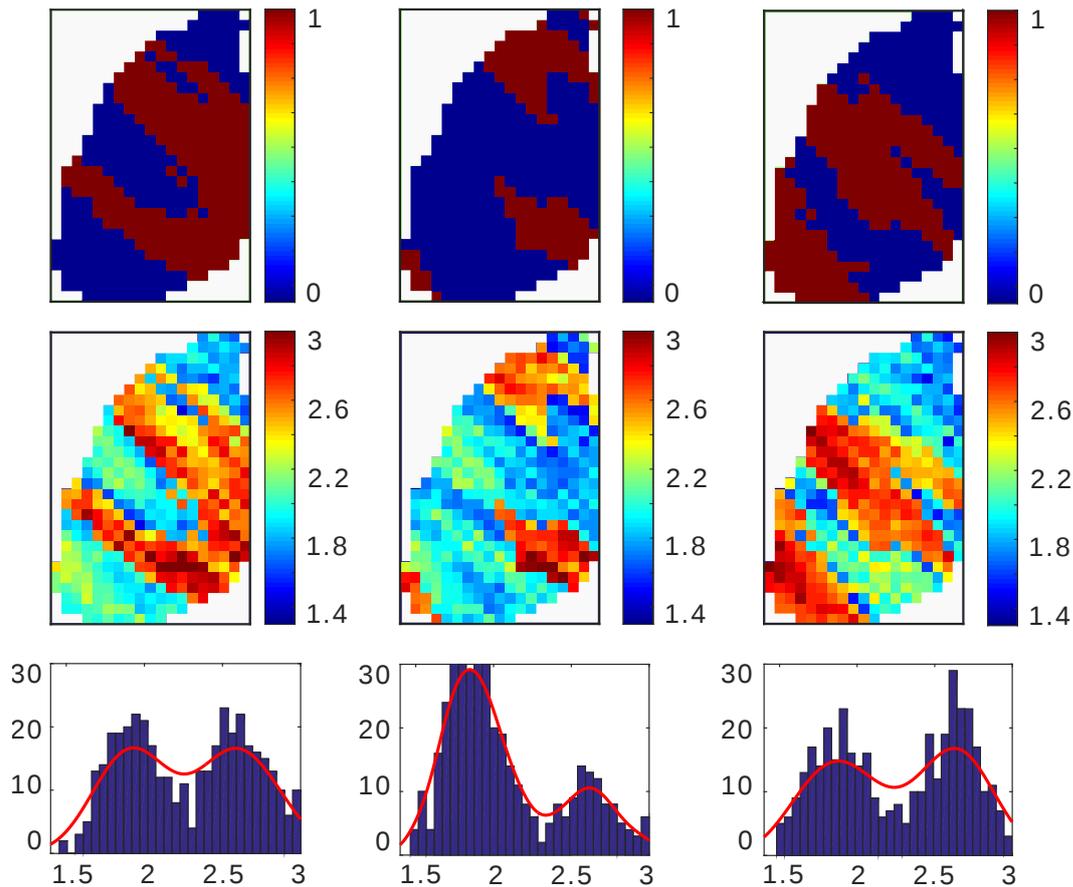


Figura 32: Primeiras três realizações de modelos de fácies, logaritmo da permeabilidade horizontal e seus histogramas do conjunto *a priori*, para o caso de estudo PUNQ-S3 modificado.

realizações finais tendem a perder as características iniciais, por exemplo, os modelos finais não apresentam distribuições bi-modais e nas duas primeiras realizações (parte esquerda da Figura 33) não pode ser visível claramente as fácies do modelo.

As curvas de produção resultantes ao aplicar o modelo ES-MDA são mostradas na Figura 34. Embora as realizações *a posteriori* tenham perdido suas características iniciais, o ajuste dos dados observados resultantes foram bem sucedidos. Isso pode ser visto nas curvas de produção, onde durante o período de histórico as curvas finais apresentam uma pequena dispersão (curvas azuis).

4.2.2 ES-MDA-PCA

Para testar o modelo ES-MDA-PCA no ajuste de histórico do caso de estudo (PUNQ-S3), utilizou-se um número de 4 assimilações, ou seja, $N_a = 4$, e as 200 primeiras realizações do conjunto de inicialização criado (primeiro passo do procedimento ES-MDA-PCA), tanto dos modelos de fácies como das

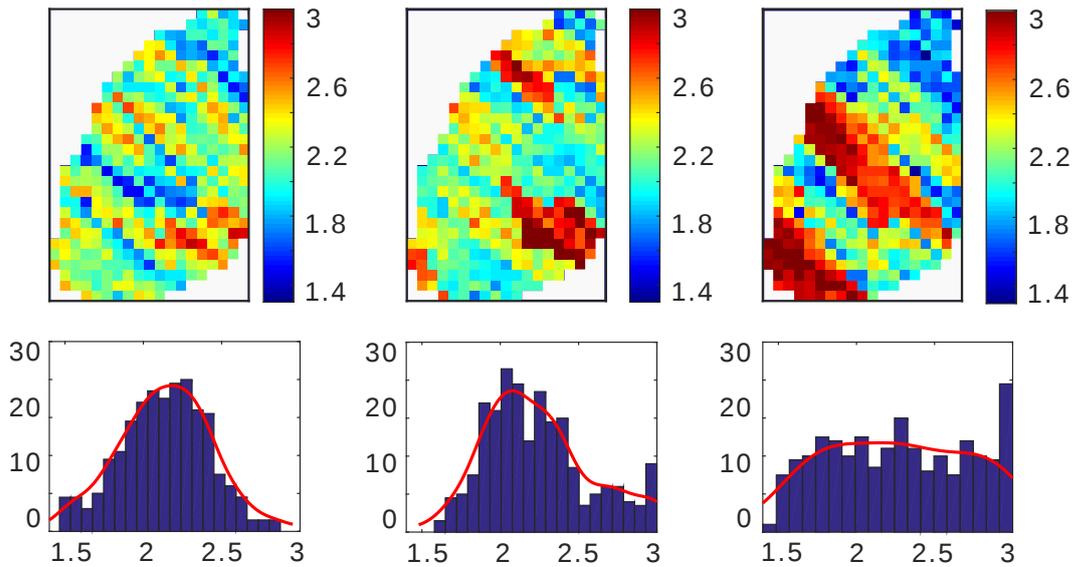


Figura 33: Primeiras três realizações do logaritmo da permeabilidade horizontal do conjunto final (*a posteriori*) utilizando ES-MDA, para o caso de estudo PUNQ-S3 modificado.

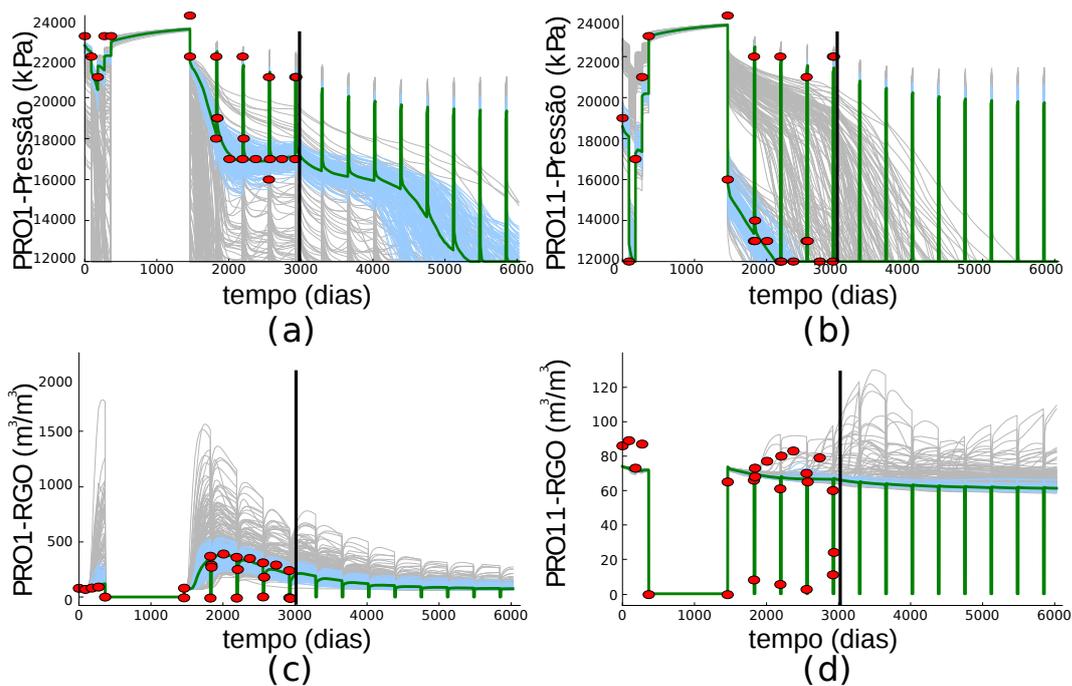


Figura 34: Curvas de produção resultantes do ES-MDA aplicado ao caso de estudo PUNQ-S3 modificado. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11.

realizações de propriedades da rocha.

Na etapa de inicialização foram criadas 5000 realizações para computar os 4 primeiros passos da metodologia ES-MDA-PCA. Assim, ao final do quarto passo, é obtido a representação dos modelos de fácies, ou seja, o vetor ξ_1 e ξ_2 . Assim o vetor de parâmetros a ser atualizado \mathbf{m} para este caso de estudo é

formado por

$$\mathbf{m} = [\xi_i^T, \xi_2^T, \phi_1^T, \phi_2^T, \log_{10}(\kappa_{h,1})^T, \log_{10}(\kappa_{h,2})^T, \log_{10}(\kappa_{v,1})^T, \log_{10}(\kappa_{v,2})^T]^T .$$

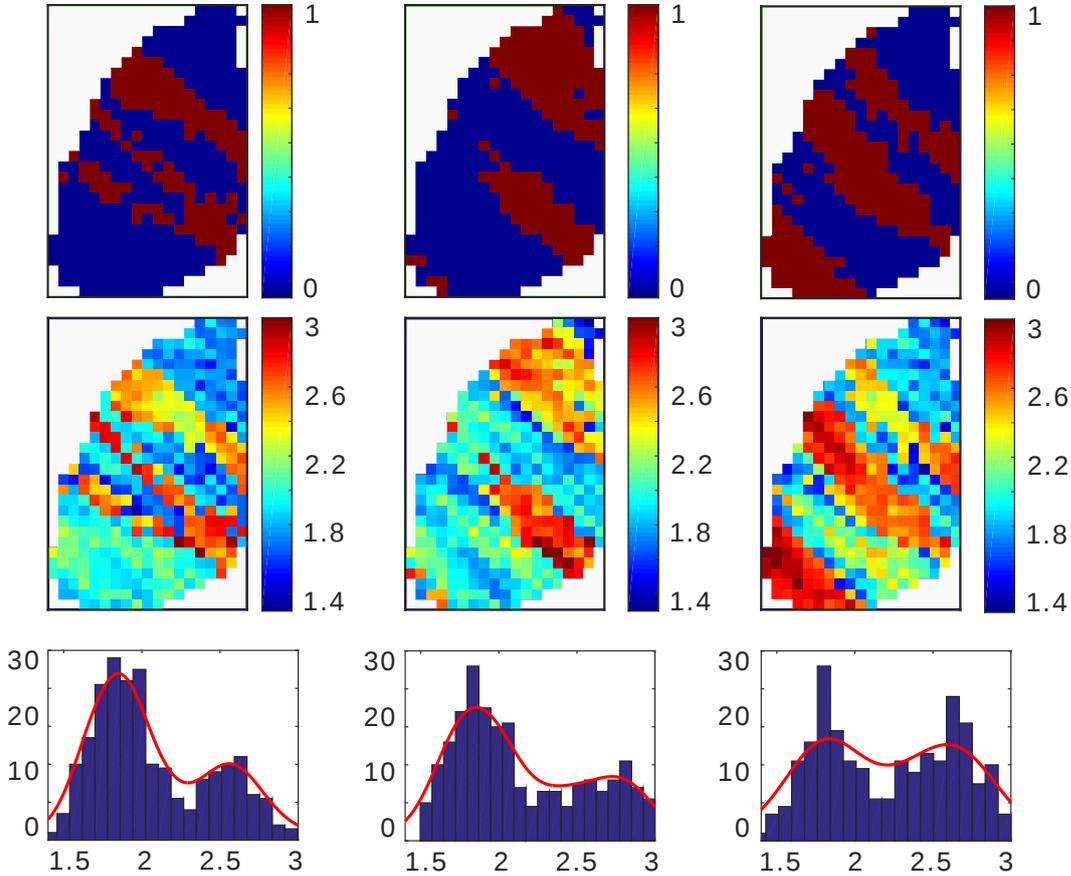


Figura 35: Primeiras três realizações de modelos de fácies, logaritmo da permeabilidade horizontal e seus histogramas do conjunto *a posteriori* utilizando ES-MDA-PCA, para o caso de estudo PUNQ-S3 modificado.

A Figura 35 ilustra as três primeiras realizações *a posteriori* da permeabilidade horizontal obtidas por ES-MDA-PCA. Note-se que aqui os mapas de logaritmo da permeabilidade horizontal obtidos conservam as características das realizações iniciais, por exemplo, os modelos finais de permeabilidade apresentam uma distribuição bi-modal (histogramas da Figura 35), características que as realizações *a priori* apresentam (histogramas da Figura 32). Além disso, nos mapas do logaritmo da permeabilidade resultantes pode-se observar claramente as fácies do modelo (linha superior da figura).

A Figura 36 apresenta as curvas de produção, tanto no início do ajuste de histórico como no final. A produção gerada pelo conjunto inicial está representada pelas curvas de cor cinza e as curvas azuis são a produção do modelo ajustado. Também é possível ver os dados observados mediante os pontos vermelhos. Comparando-se estas curvas de produção com as obtidas

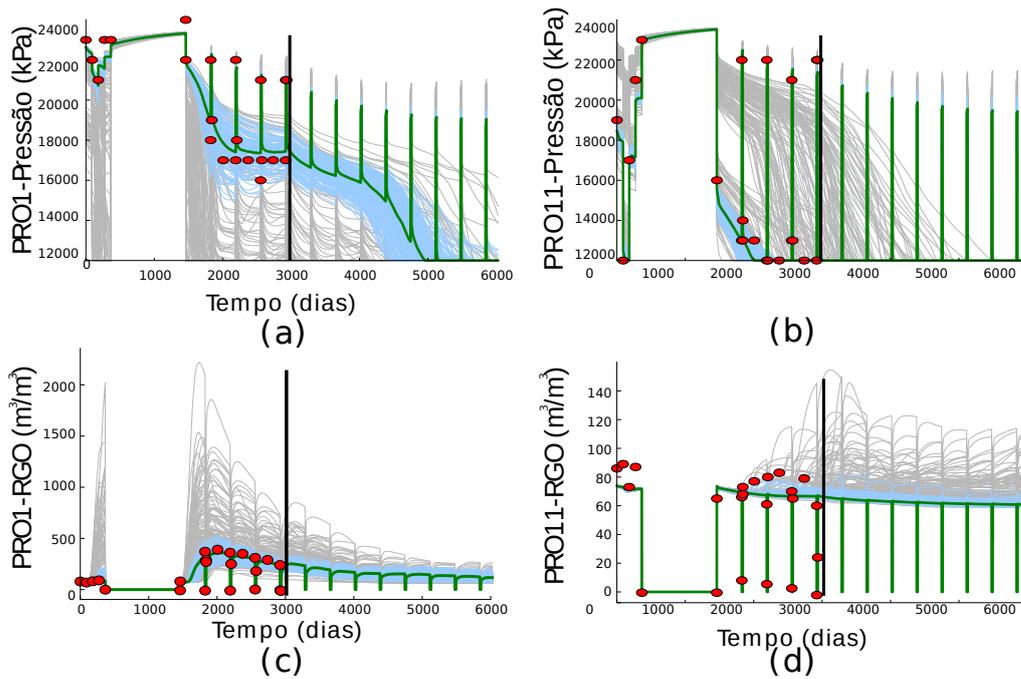


Figura 36: Curvas de produção resultantes do ES-MDA-PCA aplicado ao caso de estudo PUNQ-S3 modificado. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11.

pelo método ES-MDA, nota-se que os resultados anteriores apresentam menor dispersão, ou seja, o método ES-MDA fornece um melhor ajuste dos dados observados, embora tenda a distorcer os modelos iniciais.

Os resultados apresentados neste capítulo mostraram bom desempenho para o método ES-MDA-PCA, em comparação com os resultados obtidos pelo método ES-MDA padrão. Os resultados mostram que a aplicação direta de ES-MDA sem uma parametrização específica para tratamento das fácies não preserva alguns canais que estavam presentes nas realizações *a priori*, além de obter realizações *a posteriori* de permeabilidade que não apresentam a distribuição bi-modal - característica geostatística de cada membro do conjunto inicial. Embora a metodologia proposta baseada na PCA permita preservar apenas a média e covariância, isso foi suficiente para o caso de estudo PUNQ-S3, já que os modelos de fácies apresentadas nesse modelo não apresentam características geológicas muito complexas, apenas canais inclinados. Entretanto, provavelmente uma parametrização baseada em PCA não seja suficiente para distribuições de fácies mais complexas, como por exemplo, canais fluviais ou canais que contenham curvas na sua estrutura. À medida que aumenta a complexidade, são necessários algoritmos mais robustos para poder contornar o problema de ajuste de histórico satisfatoriamente.

Motivados pela ideia de obter uma melhor metodologia para ajustar modelos de maior complexidade, os próximos capítulos introduzem uma nova

parametrização de modelos de fácies baseada em uma possível interpretação para DL como uma generalização não-linear de PCA. De fato, DL surgiu como uma técnica inovadora nos últimos anos, com impressionantes resultados em diferentes áreas.

O próximo capítulo apresenta os principais conceitos de DL, que são posteriormente utilizados conjuntamente com ES-MDA para o ajuste de histórico.

5 Deep Learning

Neste capítulo é apresentada uma revisão sobre os principais conceitos e arquiteturas de DL que são utilizados para construir uma parametrização para a atualização de modelos de fácies geológicas em conjunto como o método ES-MDA.

As redes neurais artificiais foram alvo de uma intensa pesquisa e desenvolvimento na década de 80. Contudo, durante a década seguinte o interesse nas redes neurais diminuiu em favor de modelos mais simples e cujo treinamento era mais eficiente [29] (e.g., os campos aleatórios condicionais e as máquinas de vetores de suporte [100]). Ao longo dos últimos anos começou-se a explorar a possibilidade da utilização de técnicas poderosas para a representação de características. *Deep Learning* (DL), surgiu como uma série de técnicas que permitem representar dados com alto nível de abstração. O primeiro uso bem sucedido de DL foi reportado por Hinton et. al [97], que introduziram uma rede DL denominada redes de crenças profundas ou *Deep Belief Networks* (DBN). Com esse método a dificuldade no treinamento foi consideravelmente aliviada, trazendo assim novamente o interesse nas redes neurais para a comunidade científica.

O princípio geral de operação de uma arquitetura profunda é orientar o treinamento das camadas intermediárias, utilizando uma aprendizagem não supervisionada especificamente para cada camada. Para isso, diferentes técnicas são utilizadas, dentre as quais destacam-se, o *Autoencoder* (AE) [98], a *Restricted Boltzmann Machine* (RBM) [99] e DBN. O sucesso de DL é devido ao fato de cumprir os principais requisitos que um algoritmo de inteligência artificial precisa para ser bem-sucedido [101]:

- Habilidade de aprender funções complexas de alta variabilidade.
- Capacidade de aprender de forma não supervisionada.
- Habilidade de aprender a partir de um vasto número de exemplos.
- Capacidade de explorar a sinergia do fenômeno, ou seja, ter um aprendizado multiobjetivo.

- Habilidade de atingir uma boa generalização no caso de aprendizado não supervisionado.

5.1 Modelos Generativos Profundos

DL se refere a uma vasta quantidade de algoritmos e arquiteturas capazes de modelar tipos de dados estruturados complexos, como imagens, sons e texto. Estas arquiteturas profundas têm sido focadas em algoritmos de aprendizagem supervisionados [102], que buscam treinar um modelo estatístico capaz de estimar uma função de densidade de probabilidade (FDP) *a posteriori*, $p(y|x)$, de uma amostra de entrada x para uma amostra de saída y , que comumente é usada para problemas de classificação. Além disso, existem algoritmos de DL não supervisionados que aprendem uma representação de características extraíndo *patches* aleatórios de dados de treinamento não marcados, adotando um estágio de pré-processamento para os *patches* e aprendendo um mapeamento de recursos usando um algoritmo de aprendizado não supervisionado [103]. Esses tipos específicos de técnica são conhecidos como modelos generativos profundos ou *deep generative models* [104], e podem ser treinados e construídos usando algoritmos baseados em AE. Todos esses modelos representam distribuições de probabilidade em múltiplas variáveis de alguma forma. Alguns permitem que a FDP seja avaliada explicitamente e outras não permitem a avaliação da FDP, mas dão suporte a operações que implicitamente requerem conhecimento dela, como amostragem de dados dessa distribuição [105].

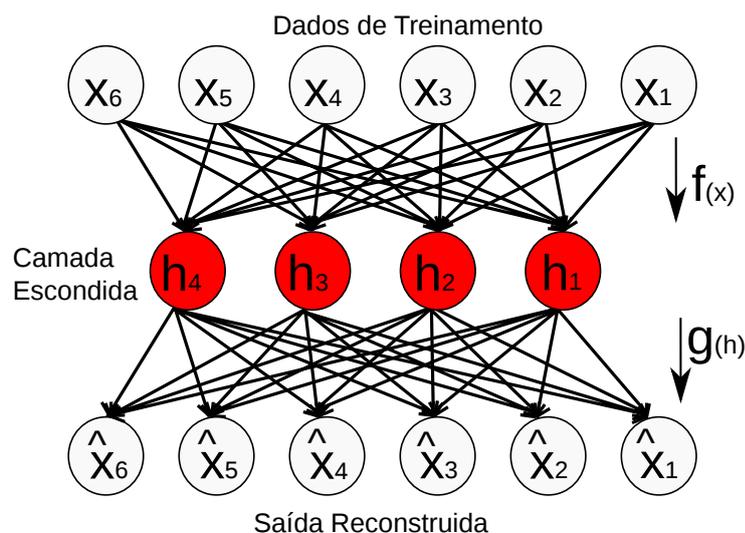


Figura 37: Estrutura geral de uma rede AE, mapeando uma entrada \mathbf{x} a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} . A rede AE apresenta duas componentes: o *encoder* $f(\mathbf{x})$ e o *decoder* $g(\mathbf{h})$.

5.1.1 Autoencoder

AE é uma rede neural não supervisionada, no sentido que o processo de treinamento não requer qualquer informação de classe. Ela é treinada para tentar copiar sua entrada para a saída. Internamente na camada oculta, \mathbf{h} , descreve-se um código usado para representar a entrada.

Uma rede AE padrão, ilustrado na Figura 37, realiza transformações na entrada de acordo com dois tipos de funções, conforme descrito a seguir.

1. **Encoder:** Função de extração de características. Modelo que mapeia a entrada $\mathbf{x} = [x_1, x_2, \dots, x_M]^T$ em outro espaço (espaço latente), usualmente nomeado como camada escondida e representado por $\mathbf{h} = [h_1, h_2, \dots, h_N]^T$, ou seja

$$\mathbf{h} = f(\mathbf{x}) = \phi^e(\mathbf{W}^e \mathbf{x} + \mathbf{b}^e). \quad (5-1)$$

2. **Decoder:** Função de reconstrução. Modelo que aprende o mapeamento inverso, leva a latente representação da camada escondida à camada visível ou saída (dada reconstruída), $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_M]$, ou seja

$$\hat{\mathbf{x}} = g(\mathbf{h}) = \phi^d(\mathbf{W}^d \mathbf{h} + \mathbf{b}^d), \quad (5-2)$$

onde \mathbf{W}^e e \mathbf{W}^d representam as matrizes de pesos do *encoder* e *decoder* respectivamente, \mathbf{b}^e e \mathbf{b}^d são os vetores de viés para o *encoder* e *decoder*, respectivamente, $\phi^e(*)$ e $\phi^d(*)$ são chamadas funções de ativação, que usualmente é não-linear e N, M representam o número de unidades na camada oculta e visível, respectivamente. Há diferentes alternativas para a função de ativação, dependendo da tarefa em questão. Três dessas funções são a *sigmoide logística* (Equação. 5-3), a *sigmoide hiperbólica tangente* (Equação. 5-4) e a *retificada linear* ou *ReLU* (Equação. 5-5)

$$\phi(x) = \frac{1}{1 + \exp(-x)}, \quad (5-3)$$

$$\phi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}, \quad (5-4)$$

$$\phi(x) = \max(0, x). \quad (5-5)$$

A Figura 38 apresenta os gráficos das funções (Equações 5-3, 5-4 e 5-5). Note-se que, com exceção da ReLU em que $x = 0$, todas são diferenciáveis e estritamente monotônicas. Note-se também que as funções sigmóides possuem assíntotas que as limitam inferior e superiormente. Outra decisão de projeto

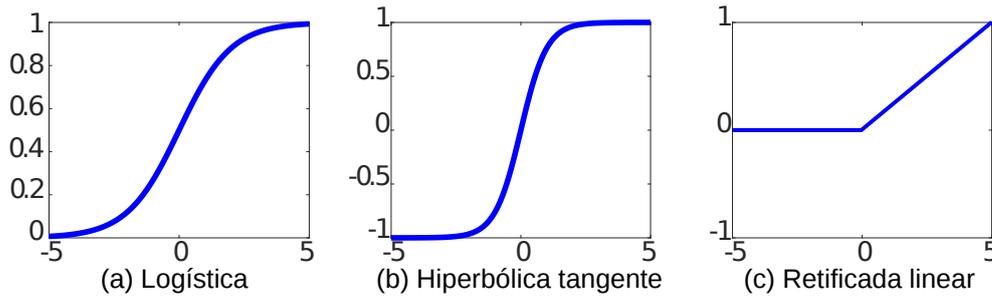


Figura 38: Funções de ativação comumente utilizadas, $\phi(x)$.

comum quando da definição da arquitetura de uma rede AE é usar o mecanismo de pesos atados ou *tied weights*, o que significa que $\mathbf{W}^e = (\mathbf{W}^d)^T$ (i.e., uma matriz é a transposta da outra).

Durante o processo de treinamento, o problema principal é aprender as matrizes e vetores viés do *encoder* e *decoder*. Assim o processo de treinamento ainda é baseado em uma otimização da função de custo Fc , definida como

$$Fc = L(\mathbf{x}, \hat{\mathbf{x}}), \tag{5-6}$$

onde L representa uma função que penaliza a diferença entre \mathbf{x} e $g(f(\mathbf{x}))$. Tradicionalmente Fc assume a forma

$$Fc = \|\mathbf{x} - g(f(\mathbf{x}))\|^2. \tag{5-7}$$

O objetivo de uma rede AE é minimizar a Equação 5-7, que claramente é um problema não convexo. No entanto, pode ser resolvido facilmente por técnicas de descida gradiente, sempre quando a função de ativação é suave e continuamente diferenciável [106]. Quando a função *decoder* é linear e Fc é o erro quadrático médio, o AE padrão aprende a projetar a entrada em um espaço onde cada parâmetro (camada escondida) são as componentes principais dos dados de entrada, que é o mesmo espaço de aprendizado do algoritmo de ACP [22]. Assim AE com funções não lineares no *encoder*, $f(\mathbf{x})$, e *decoder*, $g(\mathbf{h})$, podem ser interpretados como uma generalização não-linear da ACP [107, 105].

O AE padrão é treinado com apenas uma camada no *encoder* e outra no *decoder*. No entanto, esses dois componentes podem ser conectados em arquiteturas profundas para criar um modelo treinável de ponta a ponta para aprender representações de recursos de alto nível a partir de uma grande quantidade de entradas de dados. Esta arquitetura profunda é apresentada na Figura 39, onde à esquerda está a rede do *encoder* e à direita o *decoder*. Em muitas aplicações são utilizados os parâmetros na camada oculta.

Por outro lado, modernos AE têm generalizado a ideia de *encoder* e *decoder* além de funções determinísticas para mapeamentos estocásticos, ou seja, construir FDP [105], onde há *encoder distribution*, $p_{\text{encoder}}(h|x)$, e *decoder distribution*, $p_{\text{decoder}}(x|h)$.

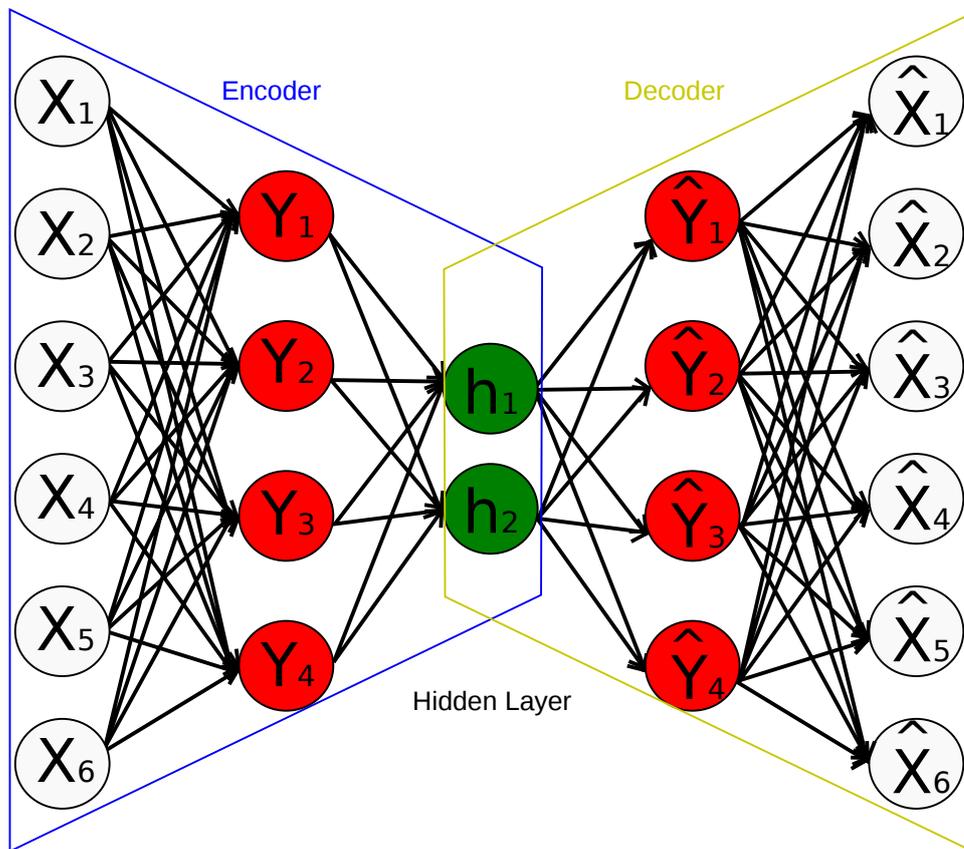


Figura 39: Estrutura de uma arquitetura AE com 3 camadas ocultas, mapeando uma entrada \mathbf{x} a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} (representação latente).

5.1.2

Variações do Autoencoder

Idealmente, os parâmetros aprendidos durante o treinamento de uma rede AE devem levar à identificação de uma representação latente dos dados de entrada por meio de uma redução de dimensionalidade, ao mesmo tempo em que mantém o máximo de informação acerca da distribuição de entrada $M > N$. Neste caso é chamado um *undercomplete autoencoder* e os parâmetros aprendidos são uma representação compacta da entrada, ou seja, a rede aprende as características mais ressaltantes do conjunto de treinamento. Por outro lado, em uma rede *overcomplete autoencoder*, onde $M \leq N$ fornece uma arquitetura que pode aprender características mais específicas do conjunto de treinamento. Embora a escolha do número de unidades na camada oculta da

complexidade da distribuição de dados a ser treinados. Assim, a escolha do valor de N , que determina a quantidade de unidades da camada intermediária, tem que ser baseada no problema onde está sendo aplicado.

Um AE pode aprender acerca da distribuição de entrada dependendo da problemática e dos dados de treinamento. Mas, é preciso tomar certas precauções para assegurar que a representação identificada seja útil. Assim, variações do AE padrão descrito acima foram propostas com o fim de evitar a produção de soluções indesejadas. Três dessas variações, descritas a seguir, são AE com filtragem de ruído ou *denoising autoencoder* (DAE), AE contrativa ou *contractive autoencoder* (CAE) e AE esparsa *sparse autoencoder* (SAE).

5.1.2.1 Denoising Autoencoder

O objetivo principal de uma rede DAE é aprender uma representação na camada escondida mais robusta a ruídos nos dados de entrada [108]. Isso pode ser realizado por meio da aplicação de um processo probabilístico em cada conjunto de treinamento antes de apresentá-lo à rede. Assim uma rede DAE é uma rede AE que recebe um dado corrompido como entrada e é treinada para prever o dado original. A saída da rede é um dado não corrompido.

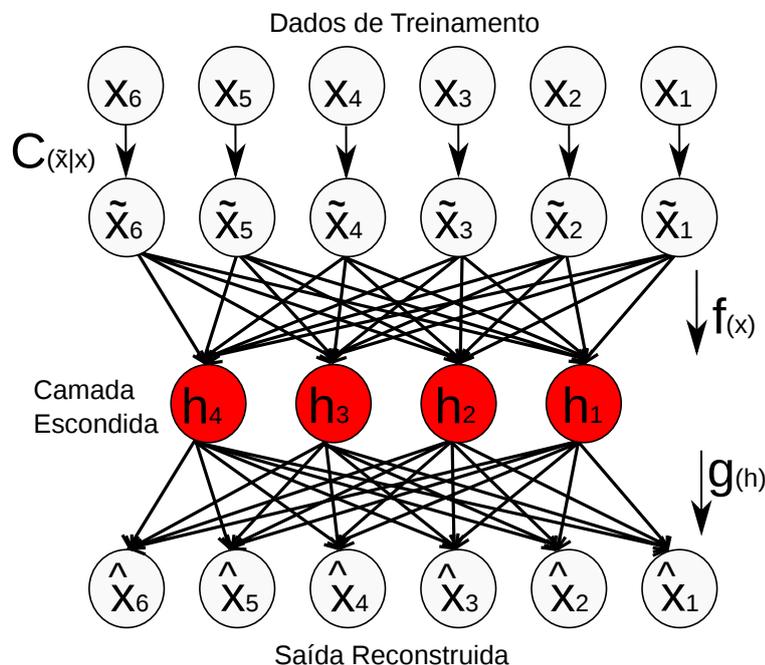


Figura 40: Estrutura geral de uma rede DAE, onde é acrescentado um ruído na entrada \mathbf{x} e mapeando a entrada ruidosa $\tilde{\mathbf{x}}$ a uma saída (nomeada reconstrução) $\hat{\mathbf{x}}$, através de uma representação interna, \mathbf{h} .

O procedimento de um rede DAE é ilustrado na Figura 40. A entrada

\mathbf{x} é corrompida mediante o processo $C(\tilde{\mathbf{x}}|\mathbf{x})$, que representa a distribuição condicional sobre as amostras corrompidas $\tilde{\mathbf{x}}$, dada uma amostra de dados \mathbf{x} .

A nova variável corrompida, $\tilde{\mathbf{x}}$, torna-se a entrada de um AE padrão. Como discutido anteriormente, a rede AE padrão durante o processo de treinamento tenta minimizar a função de custo (Equação 5-6). Contudo, na rede DAE a função de custo é modificada para as novas entradas ruidosas, onde

$$Fc = L(\mathbf{x}, g(f(\tilde{\mathbf{x}}))). \quad (5-8)$$

A introdução do ruído na camada visível faz que a rede DAE use as funções $f(\mathbf{x})$ e $g(\mathbf{h})$ para aprender a estrutura da distribuição de probabilidade do dado $p_{dada}(\mathbf{x})$ [109, 110].

5.1.2.2

Sparse Autoencoder

A rede SAE é uma alternativa que busca evitar que o algoritmo de treinamento aprenda uma representação inadequada (mesmo com o uso de mais unidades ocultas do que as existentes na camada de entrada) [111]. Esse tipo de AE procura fazer com que apenas uma pequena quantidade de unidades da camada oculta seja ativada para cada padrão de entrada. Nesse tipo de rede, a esparsidade pode ser obtida por meio da adição de termos adicionais na função de custo durante o treinamento. Assim à função de custo é acrescentada o termo de esparsidade, $\Omega(\mathbf{h})$, que depende somente dos parâmetros da camada oculta, ou seja

$$Fc = L(\mathbf{x}, \hat{\mathbf{x}}) + \Omega(\mathbf{h}). \quad (5-9)$$

A função $\Omega(\mathbf{h})$ pode ser interpretada como um termo de regularização. Existem diversas funções de regularização reportadas na literatura [112]. Uma função bastante utilizada é

$$\Omega(\mathbf{h}) = \lambda \sum_i^N |h_i|, \quad (5-10)$$

onde λ é um hiper-parâmetro conhecido como fator de esparsidade. A esparsidade também pode ser obtida mantendo apenas as k unidades mais ativas e tornando todas as demais unidades manualmente iguais a zero. Este último caso corresponde às redes denominadas k -SAE, redes com funções de ativação lineares e pesos atados [113]. Neste caso o valor de k corresponde à quantidade de unidades ocultas que devem ser mantidas.

5.1.2.3

Contractive Autoencoder

Similar à rede SAE, a rede CAE é uma abordagem para coibir a produção de representações indesejadas durante o treinamento de uma rede AE. Nessa abordagem, um termo é adicionado à função de custo para penalizar as representações indesejadas do espaço latente [114]. O regulador explícito no código $\mathbf{h} = f(\mathbf{x})$, encoraja as derivadas de f a serem tão pequenas quanto possível:

$$\Omega(\mathbf{h}) = \lambda \left\| \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right\|_F^2 \quad (5-11)$$

onde o termo de regularização é dado pela norma quadrada de Frobenius (soma dos elementos quadrados) da matriz jacobiana das derivadas parciais associadas com a função *encoder*.

Guillaume e Bengio [109] mostraram que há uma conexão entre o DAE e o CAE, para uma entrada pequena de ruído gaussiano. O erro das duas redes são equivalentes, sobre a função de reconstrução que mapeia \mathbf{x} para $\tilde{\mathbf{x}}$.

5.1.3

Restricted Boltzmann Machine

A RBM é uma rede neural estocástica capaz de reproduzir uma distribuição de probabilidade. Ela é um tipo particular do campo aleatório de Markov que apresenta uma estrutura de 2 camadas [115, 117]. Foram originalmente introduzidas por Hinton e Brown [116] como uma abordagem “conexionista” geral para aprender distribuições de probabilidade arbitrárias para vetores binários. A comunidade científica passou a dar maior atenção ao seu estudo anos mais tarde, quando o mesmo Hinton [118] em 2002 apresentou um algoritmo rápido para seu treinamento, algoritmo nomeado Divergência Contrastiva (DC), com uma facilidade para realizar inferências condicionais entre as variáveis do modelo.

Atualmente existem variantes das RBMs que incluem outros tipos de variáveis e há muito superaram a popularidade do original. Além disso, as estruturas RBM têm sido aplicadas de forma bem sucedida como blocos construtivos de arquiteturas com múltiplas camadas, como as DBN [97].

5.1.3.1

Estrutura de RBM

Uma RBM geralmente é formada por um conjunto de unidades visíveis $\mathbf{v} \in \{0, 1\}^D$ e outro de unidades ocultas $\mathbf{h} \in \{0, 1\}^F$, onde D e F respectivamente representam as dimensões do vetor de unidades visíveis e ocultas. A

estrutura da rede simples RBM é apresentada na Figura 41, onde os círculos grandes representam as unidades e os círculos pequenos verdes os parâmetros do modelo.

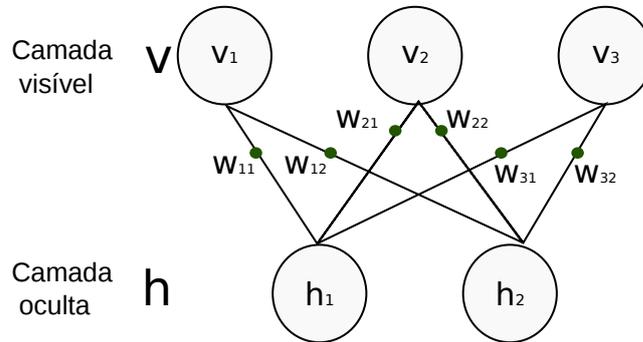


Figura 41: Representação gráfica de uma RBM com três unidades visíveis e duas ocultas.

Na Figura 41, há $F = 2$ unidades ocultas denotadas por $h_j \forall j = 1, 2$ e $D = 3$ e unidades visíveis denotadas por $v_i \forall i = 1, 2, 3$. As linhas ligando parâmetros a unidades indicam que esses estão associados. Cada parâmetro w_{ij} está simultaneamente associado às unidades h_j e v_i , sendo denominado peso da conexão entre as mesmas. Além dos pesos pode encontrar outro parâmetro na camada visível e oculta denominado viés. O viés a_i , que está associado à unidade v_i , é denominado viés dessa unidade visível. Similarmente, b_j é o viés da unidade oculta h_j .

A diferença principal entre a Máquina de Boltzmann e as RBM, é que na estrutura de uma RBM as unidades são separadas em camadas, de modo que o modelo pode ser representado por um grafo bipartido. Em outras palavras, não existem ligações entre unidades pertencentes à mesma camada, o que justifica a palavra “restrita” [119]. Em segundo lugar destaca-se que o modelo é simétrico, muito embora sejam dados nomes diferentes às unidades das duas camadas. Vale, portanto, dizer que as unidades visíveis são assim chamadas pelo fato de serem, por definição, aquelas que interagem diretamente com o meio externo, em contraste com as ocultas, cujas interações com o meio são indiretas e probabilísticas, uma nomenclatura oriunda dos Modelos Ocultos de Markov (*HMMs – Hidden Markov Models*) [120].

As RBMs fazem parte de uma vasta classe de modelos probabilísticos que são baseados em energia [105]. Nesses modelos, atribui-se um escalar, chamado de energia (total ou global), a cada estado do sistema. Para RBMs, tal estado é especificado pelo par $\{\mathbf{v}, \mathbf{h}\}$ e é mapeado em um escalar por meio de uma função de energia, definida por Desjardins et al. [121] como

$$\begin{aligned}
 E(\mathbf{v}, \mathbf{h}; \theta) &= -\mathbf{v}^T \mathbf{W} \mathbf{h} - \mathbf{a}^T \mathbf{v} - \mathbf{b}^T \mathbf{h} \\
 &= -\sum_{i=1}^D \sum_{j=1}^F w_{ij} v_i h_j - \sum_{i=1}^D a_i v_i - \sum_{j=1}^F b_j h_j,
 \end{aligned} \tag{5-12}$$

onde $\theta = \{\mathbf{W}, \mathbf{a}, \mathbf{b}\}$ são os parâmetros do modelo, os coeficientes w_{ij} representam os termos de interação simétrica entre as unidades visível e oculta e a_j e b_i são os termos viés da camada oculta e visível, respectivamente. A equação 5-12 é chamada de energia por analogia a sistemas físicos. Hopfield [122] mostrou que as unidades são simetricamente conectadas (o que é sempre verdade nas RBMs) e são atualizadas uma por vez, cada atualização reduz (ou pelo menos não aumenta) o valor de uma função perda que, quando particularizada para o caso das RBMs, fornece a expressão para $E(\mathbf{v}, \mathbf{h}; \theta)$. Consequentemente, reiteradas atualizações das unidades encontrariam, garantidamente, um mínimo de energia já que a atualização de cada unidade é realizada para minimizar a energia global do sistema.

Com base na sua energia, $E(\mathbf{v}, \mathbf{h}; \theta)$, um dado estado do sistema $\{\mathbf{v}, \mathbf{h}\}$, pode-se atribuir uma certa probabilidade, dada pela distribuição de probabilidade conjunta

$$P(\mathbf{v}, \mathbf{h}; \theta) = \frac{1}{Z(\theta)} \exp(-E(\mathbf{v}, \mathbf{h}; \theta)), \tag{5-13}$$

onde

$$Z(\theta) = \sum_{\mathbf{v}} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h}; \theta)). \tag{5-14}$$

Nas equações acima $Z(\theta)$ é chamada de função de partição e tem o objetivo de garantir que

$$\sum P(\mathbf{v}, \mathbf{h}; \theta) = 1. \tag{5-15}$$

Destaca-se que a Equação 5-13 têm a forma de uma distribuição de Boltzmann [123], justificando novamente o nome das RBMs. Pode-se notar, entretanto, que a $P(\mathbf{v}, \mathbf{h}; \theta)$ depende exclusivamente de $E(\mathbf{v}, \mathbf{h}; \theta)$, de modo que todas as informações sobre o modelo devem estar presentes na sua função de energia. Assim a probabilidade que atribui o modelo à unidade visível \mathbf{v} é representada de forma independente por

$$P(\mathbf{v}; \theta) = \frac{1}{Z(\theta)} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h}; \theta)) \tag{5-16}$$

Devido à estrutura bipartida especial das RBM, as unidades escondidas

podem ser explicitamente marginalizadas

$$\begin{aligned}
 P(\mathbf{v}; \theta) &= \frac{1}{Z(\theta)} \sum_{\mathbf{h}} \exp(\mathbf{v}^T \mathbf{W} \mathbf{h} + \mathbf{a}^T \mathbf{v} + \mathbf{b}^T \mathbf{h}) \\
 &= \frac{1}{Z(\theta)} \exp(\mathbf{b}^T \mathbf{v}) \prod_{j=1}^F \sum_{h_j \in \{0,1\}} \exp(\mathbf{b}_j h_j + \sum_{i=1}^D w_{ij} v_i h_j) \\
 &= \frac{1}{Z(\theta)} \exp(\mathbf{b}^T \mathbf{v}) \prod_{j=1}^F (1 + \exp(\mathbf{b}_j + \sum_{i=1}^D w_{ij} v_i)). \tag{5-17}
 \end{aligned}$$

As distribuições condicionais sobre o vetor de unidades ocultas \mathbf{h} e visíveis \mathbf{v} podem ser derivadas da equação 5-13 e são fornecidas por funções logísticas

$$P(\mathbf{h}|\mathbf{v}; \theta) = \prod_j^F p(h_j|\mathbf{v}), \tag{5-18}$$

$$P(\mathbf{v}|\mathbf{h}; \theta) = \prod_i^D p(v_i|\mathbf{h}), \tag{5-19}$$

$$p(h_j = 1|\mathbf{v}) = \varphi\left(\sum_i^D w_{ij} v_i + b_j\right), \tag{5-20}$$

$$p(v_i = 1|\mathbf{h}) = \varphi\left(\sum_j^F w_{ij} h_j + a_i\right), \tag{5-21}$$

onde $\varphi(x)$ é a função logística (Equação 5-3). Deve-se destacar que as Equações 5-19 a 5-21 formam o elo de ligação entre as RBMs e as redes neurais clássicas.

5.1.3.2 Divergência Contrastiva

Hinton [118] propôs o algoritmo DC visando o treinamento mais eficiente das RBM. Inicia-se a derivação do algoritmo DC obtendo-se uma expressão tratável para o gradiente da função perda (ou custo)

$$\partial \mathcal{L}(\mathbf{v}^{(t)}; \theta) = -\log P(\mathbf{v}^{(t)}; \theta), \tag{5-22}$$

sendo $\mathbf{v}^{(t)}$ uma amostra extraída da distribuição empírica dos dados de treinamento, e objetiva-se determinar os parâmetros do modelo que minimizem a função de perda (Equação 5-22), que usualmente é minimizada via o método do gradiente descendente estocástico (*Stochastic Gradient Descent*-SGD) [124]. Nesse método, o gradiente da função perda é determinado usando-se uma única amostra de treinamento e os parâmetros do modelo são ajustados no sentido oposto ao do vetor gradiente, isto é, no sentido de minimizar a perda. Esse processo é repetido até que algum critério de parada seja atingido. A derivada da log-verossimilhança em relação aos parâmetros do modelo θ pode ser obtida

da equação 5-16:

$$\frac{\partial \log P(\mathbf{v}; \theta)}{\partial \mathbf{W}} = E_{P_{\text{dado}}}[\mathbf{v}\mathbf{h}^T] - E_{P_{\text{modelo}}}[\mathbf{v}\mathbf{h}^T], \quad (5-23)$$

$$\frac{\partial \log P(\mathbf{v}; \theta)}{\partial \mathbf{b}} = E_{P_{\text{dado}}}[\mathbf{h}] - E_{P_{\text{modelo}}}[\mathbf{h}], \quad (5-24)$$

$$\frac{\partial \log P(\mathbf{v}; \theta)}{\partial \mathbf{a}} = E_{P_{\text{dado}}}[\mathbf{v}] - E_{P_{\text{modelo}}}[\mathbf{v}]. \quad (5-25)$$

onde $E_{P_{\text{data}}}[\cdot]$ denota a expectativa com respeito à distribuição de dados $P_{\text{dado}}(\mathbf{h}, \mathbf{v}; \theta) = P(\mathbf{h}|\mathbf{v}; \theta)P_{\text{dado}}(\mathbf{v})$, com $P_{\text{dado}}(\mathbf{v}) = \frac{1}{N} \sum_n \delta(\mathbf{v} - \mathbf{v}_n)$ que representa a empírica distribuição, e $E_{P_{\text{modelo}}}[\cdot]$ é a expectativa com respeito à distribuição definida pelo modelo.

Na prática, o aprendizado é feito mediante o algoritmo DC. Porque a aprendizagem exata de máxima verossimilhança deste modelo é intratável porque a computação exata da expectativa $E_{P_{\text{Modelo}}}[\cdot]$, leva um tempo que é exponencial em $\min\{D, F\}$ [118],

$$\Delta \mathbf{W} = \alpha (E_{P_{\text{dado}}}[\mathbf{v}\mathbf{h}^T] - E_{P_T}[\mathbf{v}\mathbf{h}^T]), \quad (5-26)$$

$$\Delta \mathbf{b} = \alpha (E_{P_{\text{dado}}}[\mathbf{h}] - E_{P_T}[\mathbf{h}]), \quad (5-27)$$

$$\Delta \mathbf{a} = \alpha (E_{P_{\text{dado}}}[\mathbf{v}] - E_{P_T}[\mathbf{v}]), \quad (5-28)$$

onde α é a taxa de aprendizado e P_T representa a distribuição definida mediante a cadeia de Gibbs, inicializada nos dados, para T passos completos. A estrutura bipartida especial das RBMs permite uma amostra bastante eficiente de Gibbs que alterna entre a amostragem dos estados das unidades escondidas de forma independente, dados os estados das unidades visíveis e vice-versa. Configurando $T = \infty$ recupera o aprendizado da máxima verossimilhança. No entanto, o aprendizado de DC com $T = 1$ mostrou-se funcionar razoavelmente bem [118, 125].

5.1.3.3

Generalizações do RBM

Nos últimos anos diversos progressos têm sido apresentados na definição de generalizações do RBM que melhor capturam dados de valor real, em particular dados de imagem de valor real, melhorando a modelagem da covariância condicional dos pixels de entrada [126]. O RBM padrão, conforme discutido acima, é definido com as ambas variáveis binárias visíveis $\mathbf{v} \in \{0, 1\}$ e variáveis binárias ocultas $\mathbf{h} \in \{0, 1\}$. A facilidade da inferência e do aprendizado das RBMs inspirou muitos autores a estendê-lo, por meio de modificações de sua função de energia, para modelar outros tipos de distribuição de dados. Em

particular, houve múltiplas tentativas de desenvolver modelos baseados na RBM de dados de valor real. A abordagem mais direta para a modelagem de observações de valor real nas redes RBMs é o chamado RBM gaussiano (GRBM) [127], onde uma única mudança na função de energia da rede RBM é para os viés das unidades visíveis, adicionando um termo de viés que apresenta uma forma quadrática nas unidades visíveis.

A formulação para a rede GRBM define uma nova função de energia que pode ser escrita como

$$E_G(\mathbf{v}, \mathbf{h}; \theta) = - \sum_{i=1}^D \sum_{j=1}^F h_j w_{ij} \frac{v_i}{\sigma_i^2} - \sum_{i=1}^D \frac{(v_i - a_i)^2}{2\sigma_i^2} - \sum_{j=1}^F b_j h_j, \quad (5-29)$$

onde os termos comuns à função de energia global da rede RBM (Equação 5-12) mantêm os seus significados. Surge agora um novo conjunto de parâmetros, o desvio padrão σ_i das unidades visíveis. Assim, o modelo fica totalmente especificado pelo conjunto de parâmetros $\theta = \{\mathbf{W}, \mathbf{a}, \mathbf{b}, \sigma^2\}$.

Similar a uma simples rede de RBM, a probabilidade associada a um dado estado das variáveis, $\{\mathbf{v}, \mathbf{h}\}$, decai exponencialmente com a energia global, de modo que a distribuição conjunta do modelo é definida como

$$P_G(\mathbf{v}, \mathbf{h}; \theta) = \frac{1}{Z_G(\theta)} \exp(-E_G(\mathbf{v}, \mathbf{h}; \theta)), \quad (5-30)$$

$$Z_G(\theta) = \int_{\mathbf{v}} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h}; \theta)). \quad (5-31)$$

Nota-se que essa definição tem a mesma forma daquela fornecida na Equação 5-13, similarmente as distribuições condicionais sobre o vetor de unidades escondidas \mathbf{h} e visíveis \mathbf{v} , ou seja

$$P_G(\mathbf{h}|\mathbf{v}; \theta) = \prod_j^F p_G(h_j|\mathbf{v}), \quad (5-32)$$

$$P_G(\mathbf{v}|\mathbf{h}; \theta) = \prod_i^D p_G(v_i|\mathbf{h}). \quad (5-33)$$

Fixando-se \mathbf{v} , cada variável oculta tem distribuição de Bernoulli (como nas RBMs) com probabilidade de ativação (ou sucesso), dado por

$$p_G(h_j = 1|\mathbf{v}) = \varphi\left(\sum_i^D w_{ij} \frac{v_i}{\sigma_i^2} + b_j\right). \quad (5-34)$$

Além disso, fixado \mathbf{h} , cada uma das variáveis visíveis tem distribuição gaussiana dada por

$$p_G(v_i = 1|\mathbf{h}) = \mathcal{N}(v_i | \sum_j^F w_{ij}h_j + a_i, \sigma_i^2), \quad (5-35)$$

em que a notação $\mathcal{N}(\cdot|\mu, \sigma^2)$ representa a função de densidade de probabilidade gaussiana com média μ e variância σ^2 .

5.1.4 Deep Belief Networks

As DBNs são redes generativas probabilísticas que contém múltiplas camadas de variáveis ocultas, na qual cada camada oculta captura correlações de alto nível entre as atividades das características ocultas na camada inferior. As redes DBN foram criadas por Hinton et. al [97], motivado pela limitação de trabalhar com RBM, devido ao fato de as redes RBM, apesar de ter reduzido o tempo de treinamento com o algoritmo de DC e seus variantes, apresentam dificuldades para modelar distribuições complexas, porque utilizam somente uma única camada oculta, tornando a rede rígida para adaptar-se a não linearidades de alta ordem. Assim, Hinton et. al [97] consideraram uma rede neural profunda (com várias camadas ocultas) como um empilhamento de RBM, na qual a primeira é treinada com os dados a serem modelados, enquanto que as subsequentes redes RBM pegam como dados de entrada a probabilidade de ativação da camada oculta da RBM anterior.

A convergência das redes DBN é demonstrada ao serem convergentes as redes RBM que a compõem, mas aplicando o algoritmo DC, e sendo ele uma aproximação, um refinamento da rede é necessário em alguns casos, usando métodos tradicionais tais como *backpropagation* [128], ou outro especificamente concebido para DBN como *Wake-Sleep* [129] ou seu variante *Up-Down*.

5.1.4.1 Treinamento de DBN

O processo de treinamento de uma rede DBN consiste principalmente de duas etapas, das quais a segunda etapa é opcional, mas recomendável.

1. Etapa 1: A rede DBN é treinada por camadas, agrupando as camadas de duas em duas formando assim redes RBMs. Estas redes podem ser treinadas separadamente utilizando o algoritmo DC. Aqui os pesos de reconhecimento (pesos que geram a distribuição de probabilidade na camada superior a partir de dados numa camada inferior) entre \mathbf{v} e \mathbf{h} , são mantidas iguais aos de geração (os que realizam o processo contrário, inferindo um estado para as camadas inferiores a partir do estado das camadas superiores).

2. Etapa 2: Mediante um treinamento tradicional e mais custoso computacionalmente (*backpropagation*). Este processo é conhecido como *fine-tuning* ou refinamento.

Por exemplos considere o treinamento de uma DBN com duas camadas de unidades escondidas $\{\mathbf{h}_1, \mathbf{h}_2\}$, além de supor que o número de unidades na segunda camada escondida é o mesmo que o número de unidades visíveis (ver Figura 42, na parte da direita). As duas camadas superiores da DBN formam uma RBM e as camadas inferiores formam uma rede crença sigmoide dirigida.

A distribuição conjunta sobre $\{\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2\}$ é definida pelo seguinte modelo

$$P(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta) = P(\mathbf{v}|\mathbf{h}^1; \mathbf{W}^1)P(\mathbf{h}^1, \mathbf{h}^2; \mathbf{W}^2), \quad (5-36)$$

onde $\theta = \{\mathbf{W}^1, \mathbf{W}^2\}$ são os parâmetros do modelo a ser ajustado durante o treinamento, $P(\mathbf{v}|\mathbf{h}^1; \mathbf{W}^1)$ é a rede crença sigmoide dirigida, e o termo $P(\mathbf{h}^1, \mathbf{h}^2; \mathbf{W}^2)$ é a distribuição conjunta definida pela segunda camada da rede RBM.

$$P(\mathbf{v}|\mathbf{h}^1; \mathbf{W}^1) = \prod_i p(v_i|\mathbf{h}^1; W^1), \quad (5-37)$$

$$P(\mathbf{h}^1, \mathbf{h}^2; \mathbf{W}^2) = \frac{1}{Z(\mathbf{W}^2)} \exp(\mathbf{h}^{1T} \mathbf{W}^2 \mathbf{h}^2), \quad (5-38)$$

$$p(v_i|\mathbf{h}^1; W^1) = \phi\left(\sum_j \mathbf{W}_{ij}^1 h_j^1\right). \quad (5-39)$$

Essa estratégia é conhecida na literatura como *greedy* (ambiciosa). Aqui as duas camadas da rede DBN cumprem que os parâmetros estão vinculados ($\mathbf{W}^2 = \mathbf{W}^{1T}$). Logo esta rede apresenta uma distribuição conjunta $P(\mathbf{v}, \mathbf{h}^1; \theta) = \sum_{\mathbf{h}^2} P(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta)$ idêntica à distribuição conjunta da rede RBM [130]. O algoritmo de aprendizado *greedy* usa uma pilha de redes RBM e procede da seguinte maneira. Primeiro é treinada a RBM inferior com os parâmetros \mathbf{W}^1 , conforme descrito acima. Em seguida, os pesos da segunda camada para $\mathbf{W}^2 = \mathbf{W}^{1T}$ são inicializados, o que garante o DBN de duas camadas ocultas ser pelo menos tão bom quanto nosso RBM original. Finalmente pode-se melhorar o aprendizado da rede DBN mediante a etapa de refinamento.

Assim, essa ideia pode ser estendida para o treinamento de rede DBN com mais camadas como pode-se observar na Figura 43, uma rede DBN formada de 4 redes RBM.

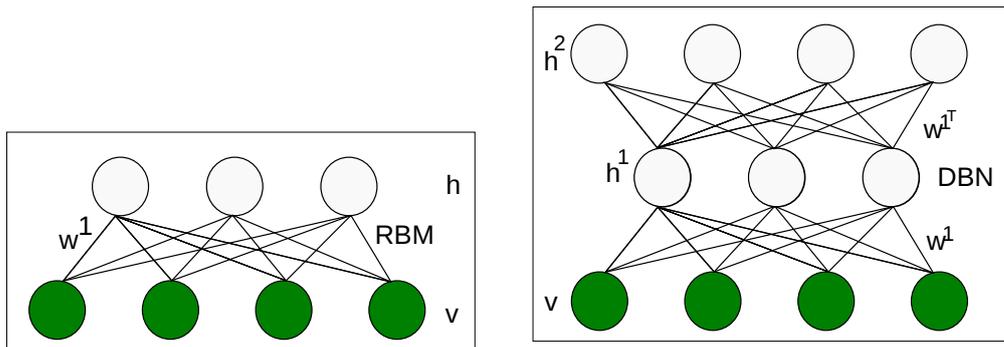


Figura 42: Esquerda: Uma rede RBM. Direita: Uma rede DBN com 2 camadas ocultas, com pesos atados $\mathbf{W}^2 = \mathbf{W}^{1T}$.

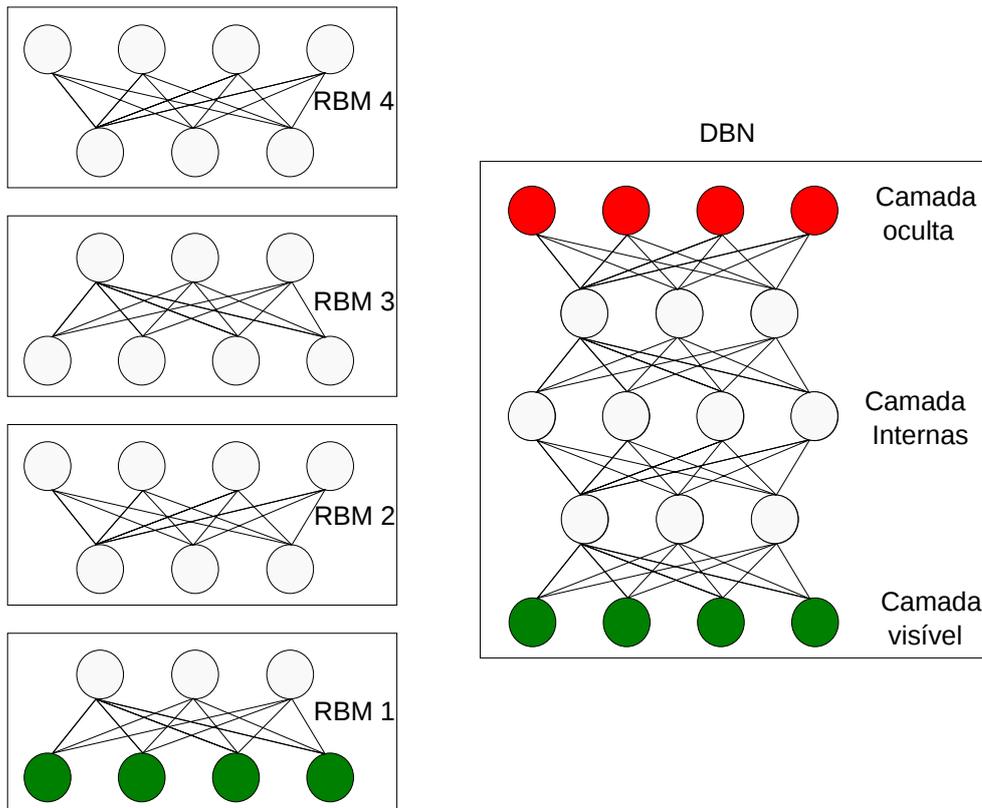


Figura 43: Representação gráfica de uma rede DBN (direita) de 4 camadas escondidas, acoplando 4 redes RBMs (esquerda).

5.1.4.2 DBN como Autoencoder

A rede DBN é um modelo gerador composto de múltiplas camadas de redes RBMs. Uma rede DBN pode ser usada como AE ou classificador. Como AE pode ser usada para criar um modelo gerador e pode ser usada em muitas aplicações, como a extração de recursos. A Figura 44 mostra uma rede DBN utilizada como AE com duas camadas RBM. As unidades ocultas da última camada podem ser usadas como um vetor de características que foi extraído da

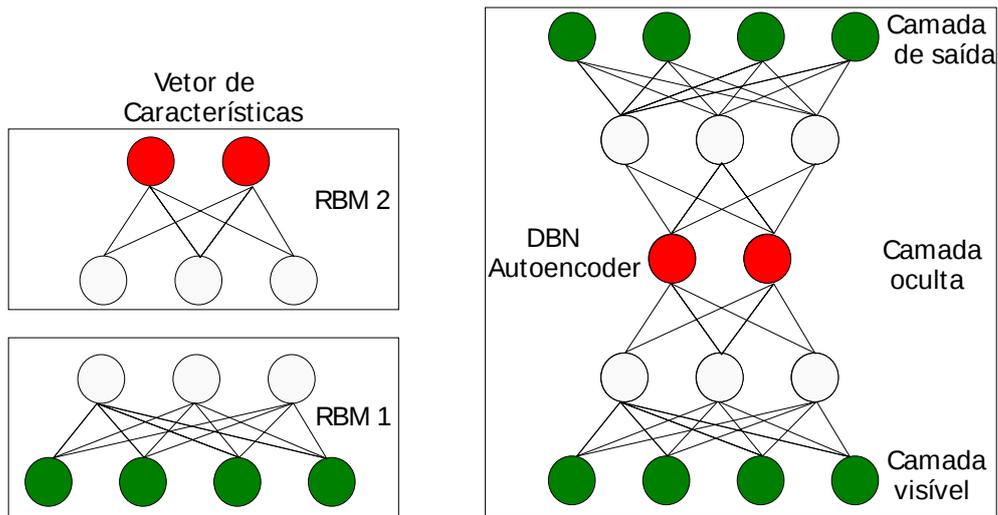


Figura 44: Representação gráfica de uma RBM com três unidades visíveis e duas ocultas.

entrada de dados visíveis. Sabe-se que o objetivo de uma rede AE é reconstruir o dado de entrada, mediante duas funções, *encoder* e *decoder*. A conversão de uma rede DBN para ser uma rede AE, consiste em replicar a estrutura da rede de forma simétrica (ver Figura 44 na direita). Aqui pode-se identificar na parte superior a estrutura referente ao *decoder* e na parte inferior a estrutura da função *encoder*.

5.2 Treinamento de Modelos de Fácies

Nesta seção são apresentados os resultados obtidos para diferentes conjuntos de imagens, que representam modelos de fácies geológicas. O objetivo destes experimentos é mostrar como as redes generativas profundas conseguem representar uma imagem (retratada em uma matriz com valores discretos $x \in \{0, 1, 2, \dots\}$), em um espaço contínuo, especificamente um espaço onde a variável parametrizada, \mathbf{h} , apresenta uma distribuição gaussiana. Isto é importante, pois esses parâmetros serão integrados com o método ES-MDA que é ótimo para modelos gaussianos.

Cada matriz que representa a imagem de fácies, é estendida em um vetor que foi definida para todos os testes como \mathbf{x} . A ideia é treinar uma rede que possa reproduzir uma imagem de entrada (representada pelo vetor \mathbf{x}), a partir de uma representação dessa imagem. O processo consiste em levar o vetor \mathbf{x} para um novo espaço onde o vetor \mathbf{x} é representado por um vetor \mathbf{h} . Posteriormente essa representação \mathbf{h} é transformada para um vetor $\hat{\mathbf{x}}$, que corresponde à saída da rede, ou seja, uma imagem reproduzida (representada pelo vetor $\hat{\mathbf{x}}$). Para isso, a rede deve retornar ao final do treinamento duas

funções: *encoder* e *decoder*. Uma vez treinada a rede, o sistema pode receber um vetor qualquer \mathbf{x} (imagem de fácies geológica) e representá-la em um novo espaço contínuo, \mathbf{h} .

Tabela 9: Resumo das Bases de Treinamento

Base	Dimensão	N_t	N_{teste}	\mathbf{x}	\mathbf{h}	Fácies
PUNQ-S3	19×28	5000	1000	532	532	2
MOD-3F	19×28	5000	1000	532	532	3
MOD45	45×45	5000	1000	2045	2045	2
MOD100	100×100	5000	1000	10000	10000	2
BRUGGE	139×48	2000	–	6672	6672	2

Tabela 9 mostra o resumo das diferentes bases de dados utilizadas nesta tese para o treinamento, o nome atribuído, número de amostras para o treinamento e teste (N_t e N_{teste}), a dimensão do vetor \mathbf{x} e \mathbf{h} , e o número de fácies das imagens da base de dados. Em todos os testes realizados considerou-se que o número de parâmetros na camada visível \mathbf{x} e oculta \mathbf{h} , sejam de igual dimensão. Devido ao fato de que este aprendizado é utilizado no próximo capítulo para realizar o procedimento de ajuste de histórico, e dado que em determinados casos o ajuste de histórico é realizado utilizando a abordagem de localização, este procedimento precisa que os parâmetros a ser atualizados sejam da mesma dimensão das propriedades geológicas (permeabilidade, porosidade).

A base de treinamento PUNQ-S3 corresponde ao modelo criado no capítulo anterior, a segunda base MOD-3F corresponde a um modelo de 3 fácies e criada utilizando co-simulação gaussiana, mediante a ferramenta SGeMS [96] e as três últimas bases de dados, correspondem a imagens de fácies geológicas criadas por algoritmos da MPS e que foram utilizados em uma pesquisa de métodos de parametrização baseadas na ACP com o método ES-MDA e que foi desenvolvida por Emerick [21].

5.2.1

Arquitetura das Redes Utilizadas

Para os conjuntos de treinamento da Tabela 9 foram utilizadas redes AE, RBM e DBN como AE. As especificações das redes são mostradas na Tabela 10, onde é apresentada o nome da rede, a base de dados ao qual foi aplicada a rede e o número de camadas ocultas que apresenta a rede, $N_{camadas}$. Além, disso todas as redes foram treinadas utilizando como critério de parada, o número de épocas $N_{epocas} = 2000$ (diferentes experimentos mostraram que para esse

valor a rede conseguiu um aprendizado ótimo). Os prefixos AE, RBM e DBN correspondem a uma rede AE, RBM e DBN, respectivamente, e os números entre parênteses correspondem ao número de unidades na camada oculta. As diferenças das redes DBN testadas está no número de parâmetros nas camadas visível e oculta.

Tabela 10: Configuração das Redes AEs para Modelos de Fácies

Rede	Base	x	h	$N_{camadas}$
AE(532)	PUNQ-S3	532	532	1
RBM(532)	PUNQ-S3	532	532	1
DBN(532-256-532)	PUNQ-S3, MOD-3F	532	532	3
DBN(2045-1024-2045)	MOD45	2045	2045	3
DBN(10000-4096-10000)	MOD100	10000	10000	3
DBN(6672-2048-6672)	BRUGGE	6672	6672	3

Figura 45 é apresentado o grafo da rede AE(532). Como pode ser observado, nesta rede somente existe uma camada oculta (círculos vermelhos), internamente a rede AE foi configurada com a função de ativação logística na saída do *decoder* e uma ReLU na saída do *encoder* e um fator de esparsidade de $\lambda = 0.05$ no termo de regularização (Equação 5-10). Similarmente, na Figura 46 é ilustrada a rede RBM utilizada para o treinamento da base PUNQ-S3, nota-se que a rede também apresenta o mesmo números de unidades na camada oculta e visível. Na Figura 47, a rede DBN(532-256-532) é ilustrada, onde é mostrada que têm três camadas ocultas. O grafo dessa rede é o mesmo para as três últimas redes da Tabela 10, somente trocando o número de parâmetros na camadas visível e oculta. Tendo em vista que cada caso de estudo corresponde a imagens de realizações de fácies de diferente dimensão.

5.2.2 Resultados

O processo de treinamento dos modelos de fácies para cada caso de estudo é realizado como etapa prévia à metodologia proposta no próximo capítulo. As arquiteturas já treinadas obtidas nesta seção são utilizadas posteriormente em integração com o método ES-MDA para realizar o ajuste de histórico. Ao final de cada treinamento é salva as duas funções treinadas, (*encoder* e *decoder*).

Em continuação são apresentados os resultados do treinamento para todos os casos de estudo realizados nesta tese.

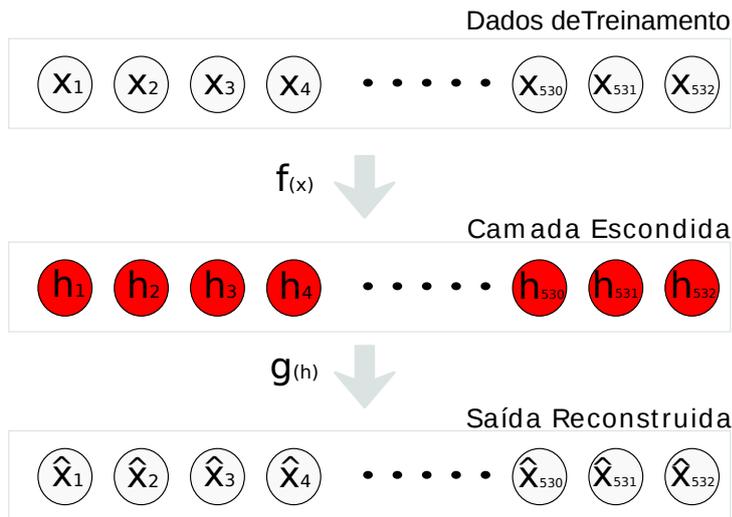


Figura 45: Arquitetura da rede AE para o treinamento do caso de estudo PUNQ-S3.

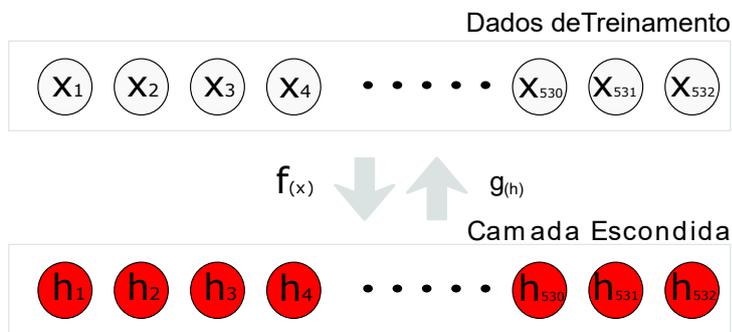


Figura 46: Arquitetura da Rede RBM para o treinamento do caso de estudo PUNQ-S3.

5.2.2.1 PUNQ-S3

Segundo a Tabela 10, três experimentos foram aplicados para este caso de estudo. A base PUNQ-S3 é simples em comparação com as outras, apresentando duas fácies canalizadas, geralmente canais inclinados, mas sem curvas. Na Figura 48 ilustra-se as 30 primeiras realizações de fácies que pertencem à base de treinamento. Note-se a característica principal dessas imagens, que contem canais em uma só direção.

O primeiro experimento corresponde a uma rede AE(532). Depois do processo de treinamento, o modelo foi testado com a base de teste. Na Figura 49 são apresentados os resultados obtidos e como pode ser visto nesta, na linha superior estão as imagens de fácies geológicas de teste, que ao passar pela extração de características (função *encoder*) obtém-se uma representação da imagem em um espaço contínuo, onde a parametrização \mathbf{h} segue uma distribuição gaussiana (histograma no centro da figura). Na linha inferior são

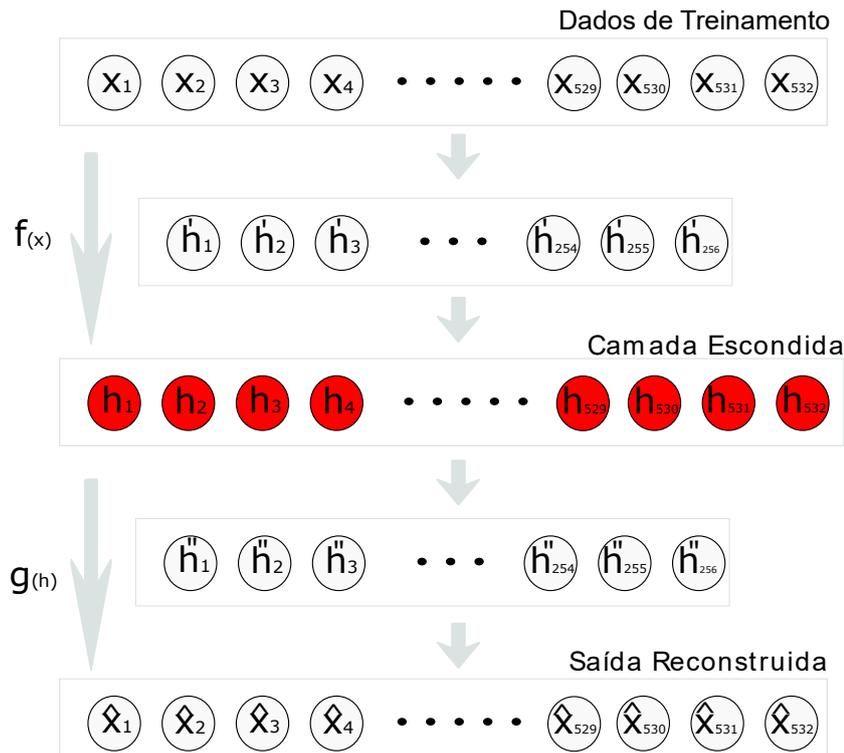


Figura 47: Arquitetura da rede DBN como AE para o treinamento do caso de estudo PUNQ-S3.

PUC-Rio - Certificação Digital N° 1313464/CA

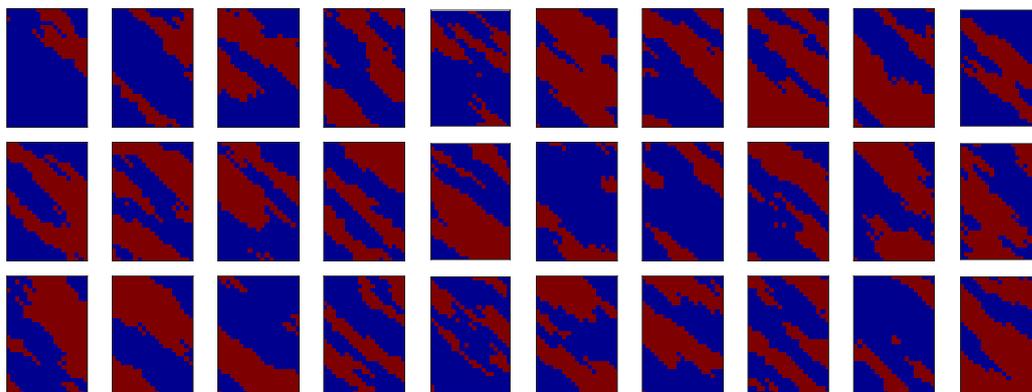


Figura 48: Imagens de fâcies de treinamento do caso de estudo PUNQ-S3.

ilustradas as imagens reconstruídas das fâcies geológicas pela rede mediante a função *decoder*, previamente ajustada com os parâmetros finais de treinamento.

Apesar das imagens reconstruídas mostrarem uma clara reprodução das imagens de entrada, existem pixels que não estão claramente definidos. Embora os resultados obtidos por esta arquitetura possam ser considerados satisfatórios, quando essa rede é aplicada a modelos com maior complexidade apresenta dificuldades para reconstruir as imagens de fâcies.

A segunda arquitetura testada, RBM(532), para o caso de estudo PUNQ-S3, apresenta uma melhor reconstrução, como pode ser observado na Figura 50 (linha inferior). Os resultados na reconstrução em comparação com a

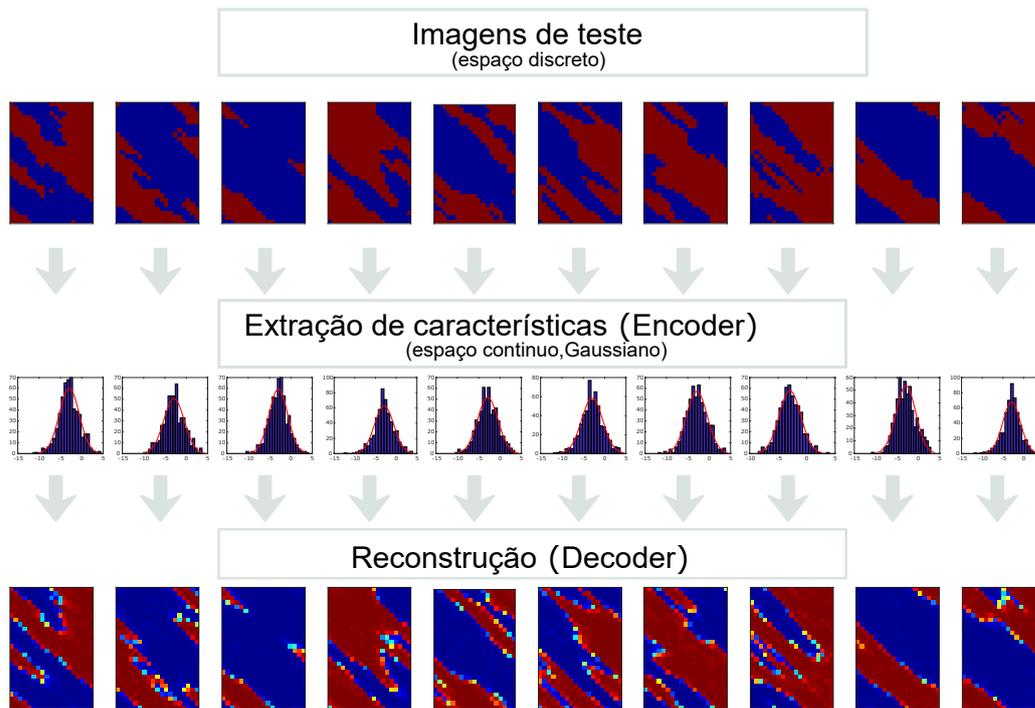


Figura 49: Imagens de fácies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede AE.

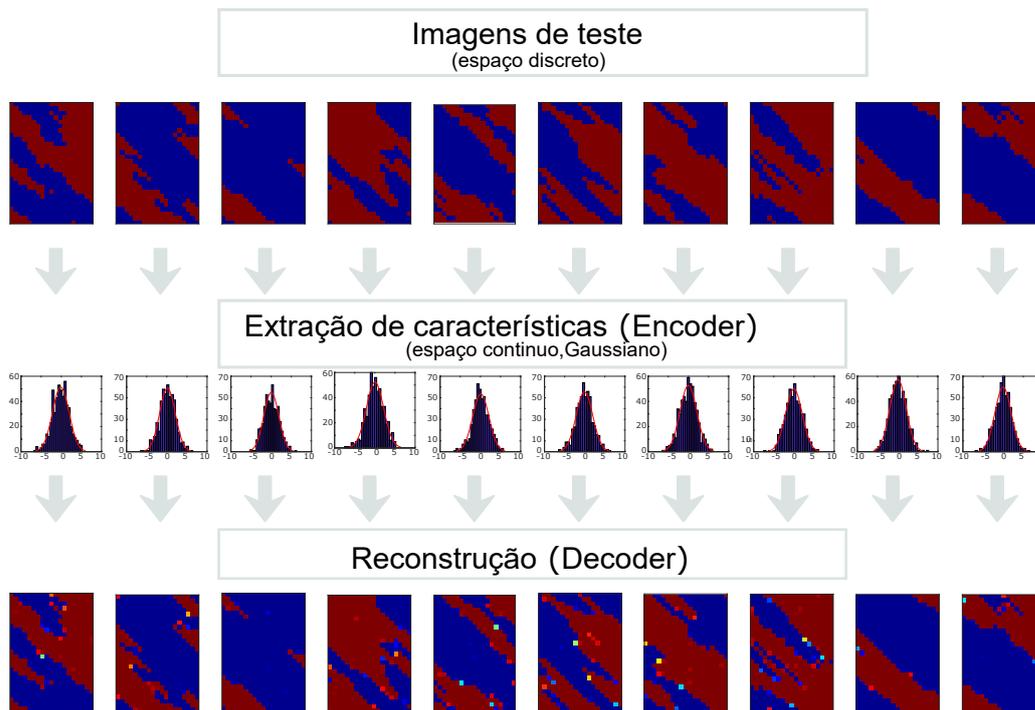


Figura 50: Imagens de fácies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede RBM.

arquitetura AE(532), mostram pixels mais definidos para cada fácies, mas alguns pixels não correspondem à fácies da imagem de entrada.

Para lidar com as principais deficiências das duas primeiras arquitetu-

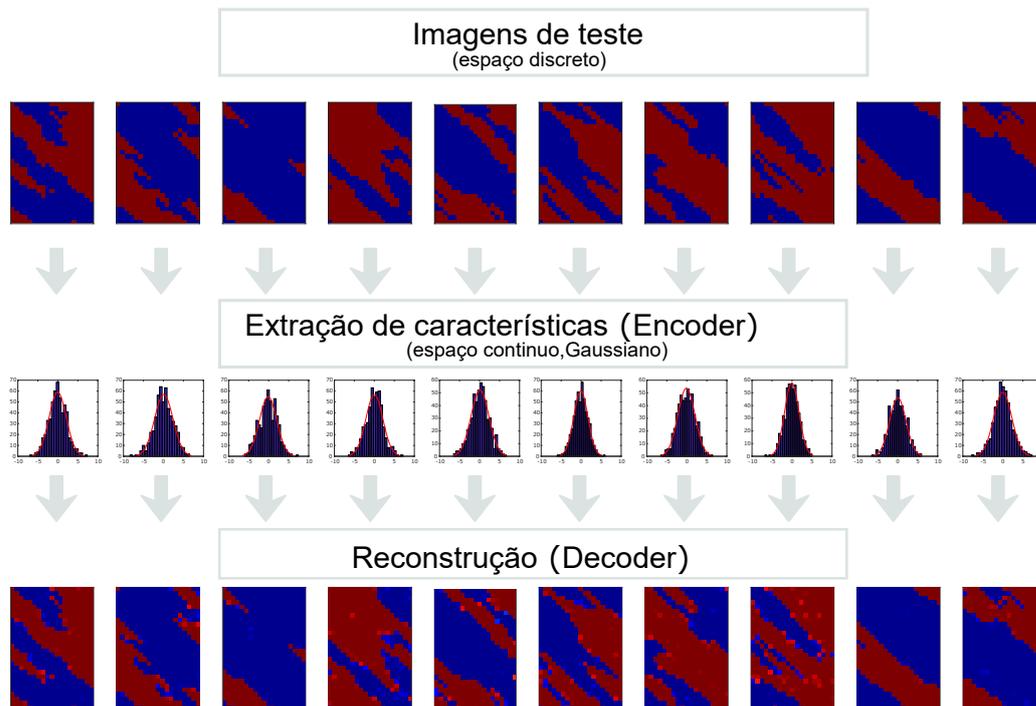


Figura 51: Imagens de fâcies da base de teste do caso de estudo PUNQ-S3, utilizando uma rede DBN como AE.

ras, foi testada a rede DBN(532-256-532). Ela consegue capturar e reconstruir a imagem de entrada, com pixels bem definidos, além de obter uma parametrização contínua e gaussiana. Embora os resultados de reconstrução pela arquitetura DBN(532-256-532) sejam melhores que as outras, o tempo de treinamento é maior, o que leva a utilizar recursos como processamento paralelo e placas de vídeo GPU, sobretudo quando a dimensão do modelo de fâcies é incrementado. A Tabela 11 mostra o tempo de treinamento das 3 arquiteturas, utilizando uma placa de vídeo “GeForce GTX 1080” e uma “CPU”, Intel(R) Core(TM) i7-4770 com clock de 3.4Ghz de 32GB RAM.

Tabela 11: Tempo de Treinamento do Caso de Estudo PUNQ-S3 em CPU vs GPU

Rede	Base	t_{CPU} (Min)	t_{GPU} (Min)
AE(532)	PUNQ-S3	34,5	5,5
RBM(532)	PUNQ-S3	72	8,9
DBN(532-256-532)	PUNQ-S3	92,9	18,4

No teste da rede AE(532) foi utilizada a a função de treinamento de *autoencoder* do Matlab e para as outras arquiteturas foi utilizada a biblioteca DeeBNet [131], versão 3.2, que permite uso de GPU. Para o caso da rede

AE(532) foi implementada também utilizando a biblioteca *Keras/Tensorflow* em *Python*. Embora decidiu-se utilizar a implementação no Matlab devido ao fato de que a rede será utilizada posteriormente em integração com o método ES-MDA de ajuste de histórico a qual foi implementada em Matlab.

5.2.2.2 MOD-3F

O segundo caso de estudo pertence a imagens que contem 3 fácies, a qual foi criada de forma similar do caso de estudo PUNQ-S3, detalhado no capítulo 4. A base MOD-3F foi treinada com o fim de mostrar que a metodologia de parametrização baseada em DL pode ser estendida para imagens com mais de duas fácies.

A Figura 52 ilustra as 30 primeiras imagens de realizações de fácies geológicas da base de treinamento, e como pode ser visto essas imagens apresentam a mesma característica da base PUNQ-S3, com a única diferença que está contém 3 fácies.

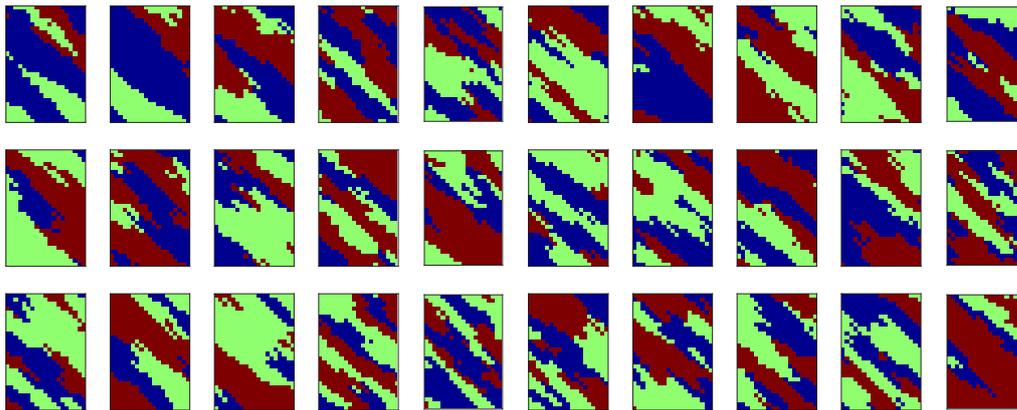


Figura 52: Imagens de fácies da base de treinamento do caso de estudo MOD-3F.

Neste experimento foi aplicada somente a arquitetura DBN(532-256-532), a qual apresentou melhores resultados no caso de estudo anterior.

Os resultados para a base de teste podem ser visualizados na Figura 53, onde similarmente aos resultados para o primeiro caso de estudo o vetor que representa cada imagem em um novo espaço, \mathbf{h} , apresenta continuidade e uma distribuição gaussiana, além de reconstruir a imagem de entrada, na qual pode-se claramente diferenciar as 3 fácies em cada realização.

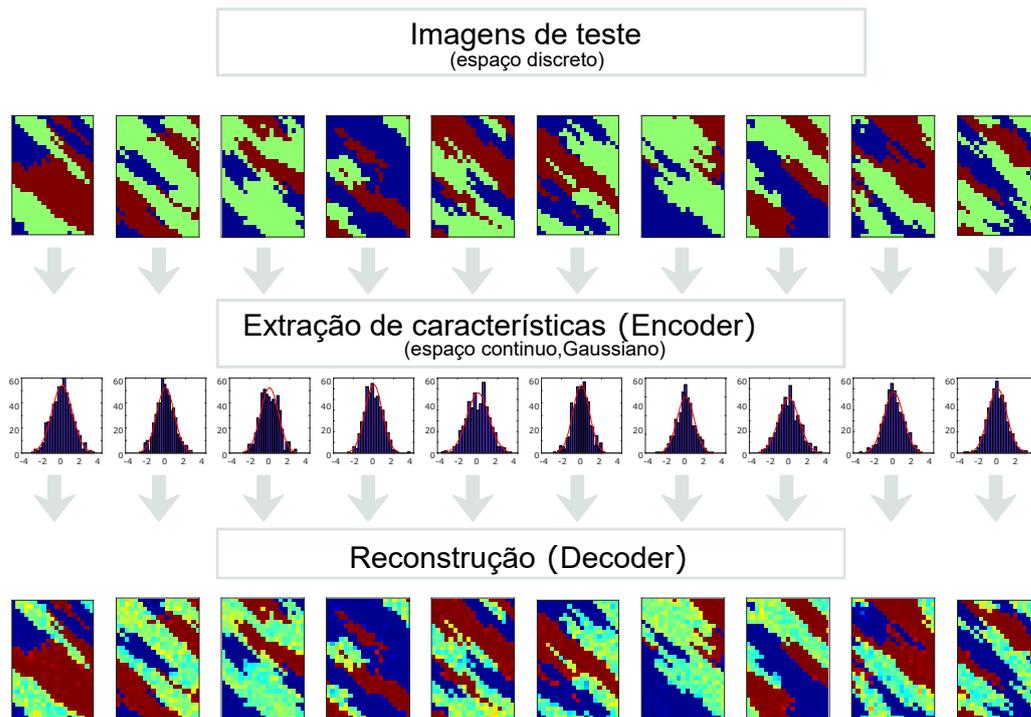


Figura 53: Imagens fâcies da base de teste do caso de estudo MOD-3F, utilizando uma rede DBN como AE.

5.2.2.3 MOD45

O terceiro caso de estudo corresponde a uma série de imagens de fâcies que foram construídas mediante algoritmos da MPS. O conjunto de treinamento e de teste foi criado usando a ferramenta SGeMs, baseado na pesquisa desenvolvida por Emerick [21].

As imagens correspondem a modelos canalizados de duas fâcies e a diferença para os casos de estudo prévios é que cada imagem apresenta canais mais complexos, ou seja, canais em diferentes direções, com algumas curvas na sua estrutura e cada imagem é de dimensão 45×45 .

É importante mencionar que foram aplicadas as mesmas arquiteturas apresentadas para os modelos anteriores, mas devido à sua complexidade as redes com arquitetura AE(2045) e RBM(2045) conseguiram reconstruir as imagens de fâcies com várias regiões onde os pixels não estavam bem-definidos.

Figura 54 apresenta os resultados do treinamento da arquitetura DBN-2045-1048-2045. Como pode ser visto as imagens de teste são representadas corretamente e são reproduzidas na saída da rede. Embora na saída alguns pixels não estão bem definidos, claramente as fâcies são distinguíveis. O experimento foi executado somente em GPU, tomando um tempo de treinamento de 3.2 horas.

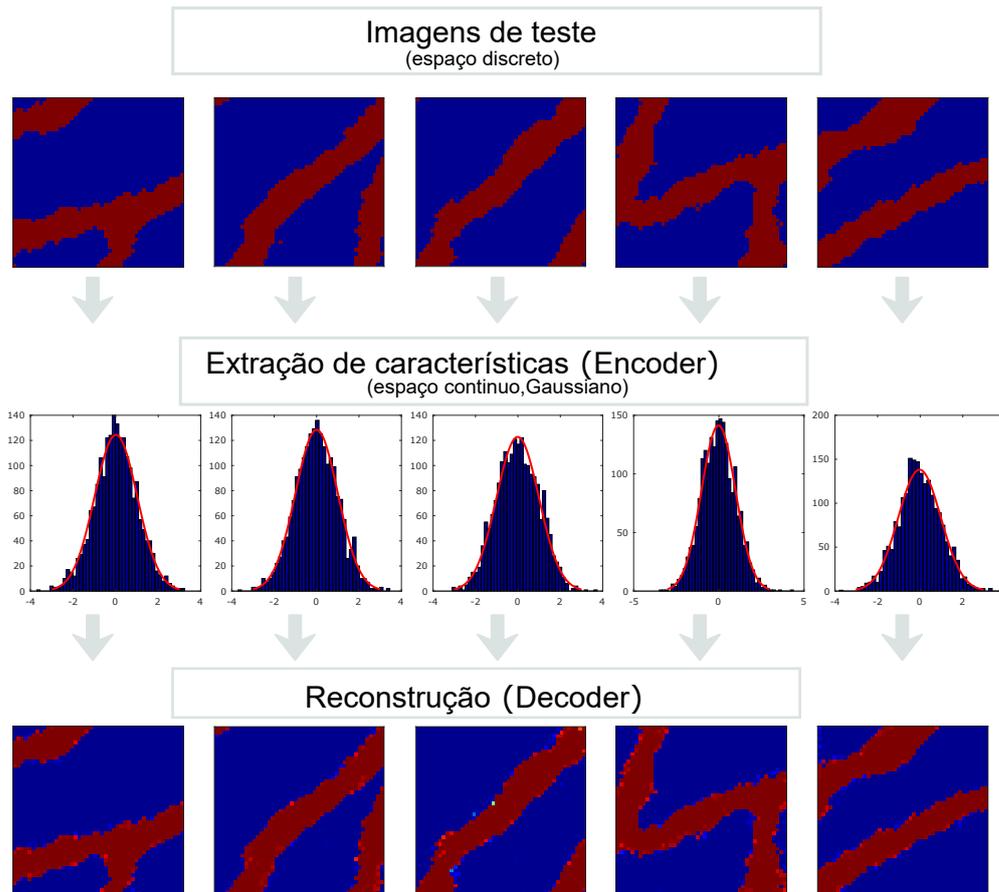


Figura 54: Imagens de fácies da base de teste do caso de estudo MOD45, utilizando uma rede DBN como AE.

5.2.2.4 MOD100

O caso de estudo MOD100 corresponde a um modelo maior de dimensão 100×100 . Igualmente ao caso de estudo MOD45, a base de dados é criada mediante algoritmos de MPS. A estrutura dos canais em cada imagem de fácies geológica apresenta canais em várias direções com diferentes curvas em cada canal. Essas características dificultam o problema.

A arquitetura utilizada para este caso de estudo corresponde a DBN(10000-4096-10000) e os resultados mostram claramente como o modelo treinado consegue reconstruir as imagens de teste, além de representá-lo no espaço contínuo e gaussiano. O experimento foi executado somente em GPU, tomando um tempo de treinamento de 20,7 horas.

Quando uma rede de arquitetura AE(100000) foi aplicada a esse caso de estudo, os resultados mostraram que a simples AE consegue reconstruir as imagens de entrada com muita ambiguidade, ou seja, os limites entre uma

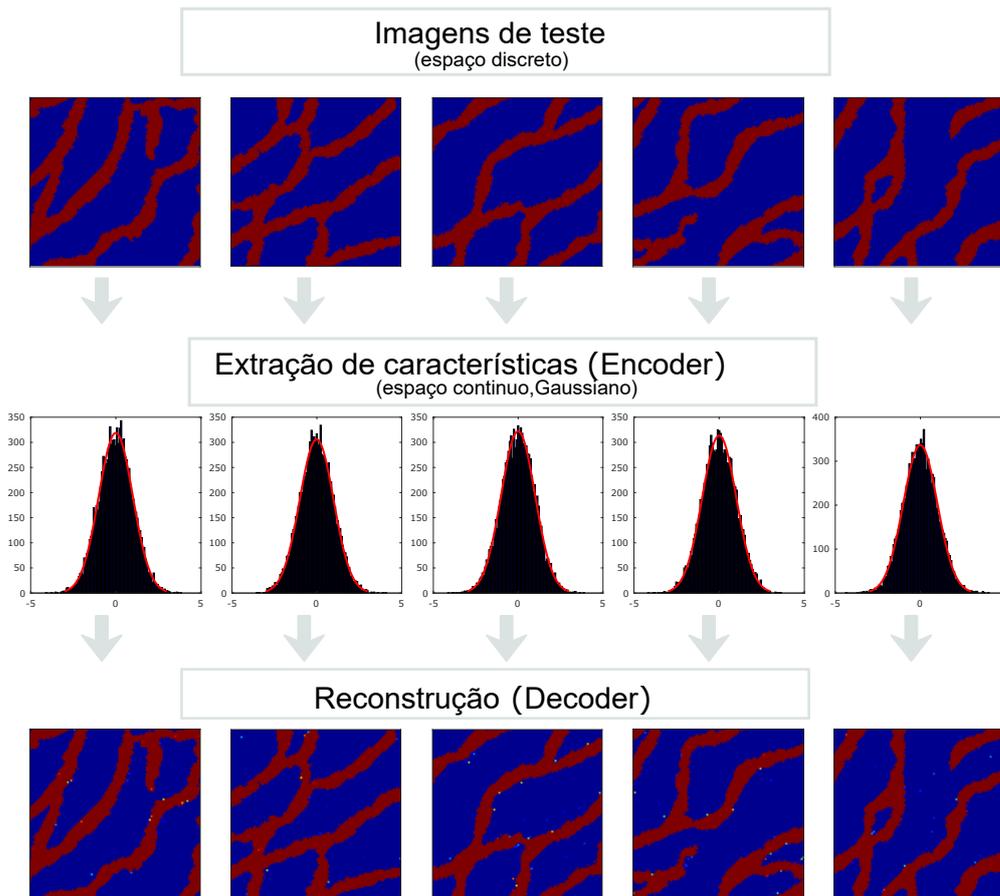


Figura 55: Imagens de fácies da base de teste do caso de estudo MOD100, utilizando uma rede DBN como AE.

fácies e outra apresentavam um efeito suavizado. Entretanto quando a rede com arquitetura RBM(10000) foi aplicada a esse caso de estudo, a rede não conseguiu convergir.

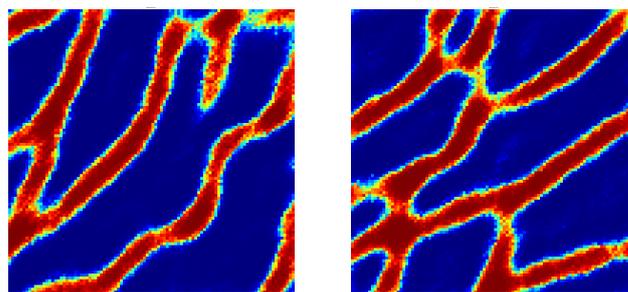


Figura 56: Reconstrução de duas imagens de fácies do caso de estudo MPS100, usando uma arquitetura AE.

5.2.2.5 BRUGGE

O último caso de estudo corresponde uma base de dados de imagens de fácies geológicas do bem conhecido reservatório BRUGGE modificado. A for-

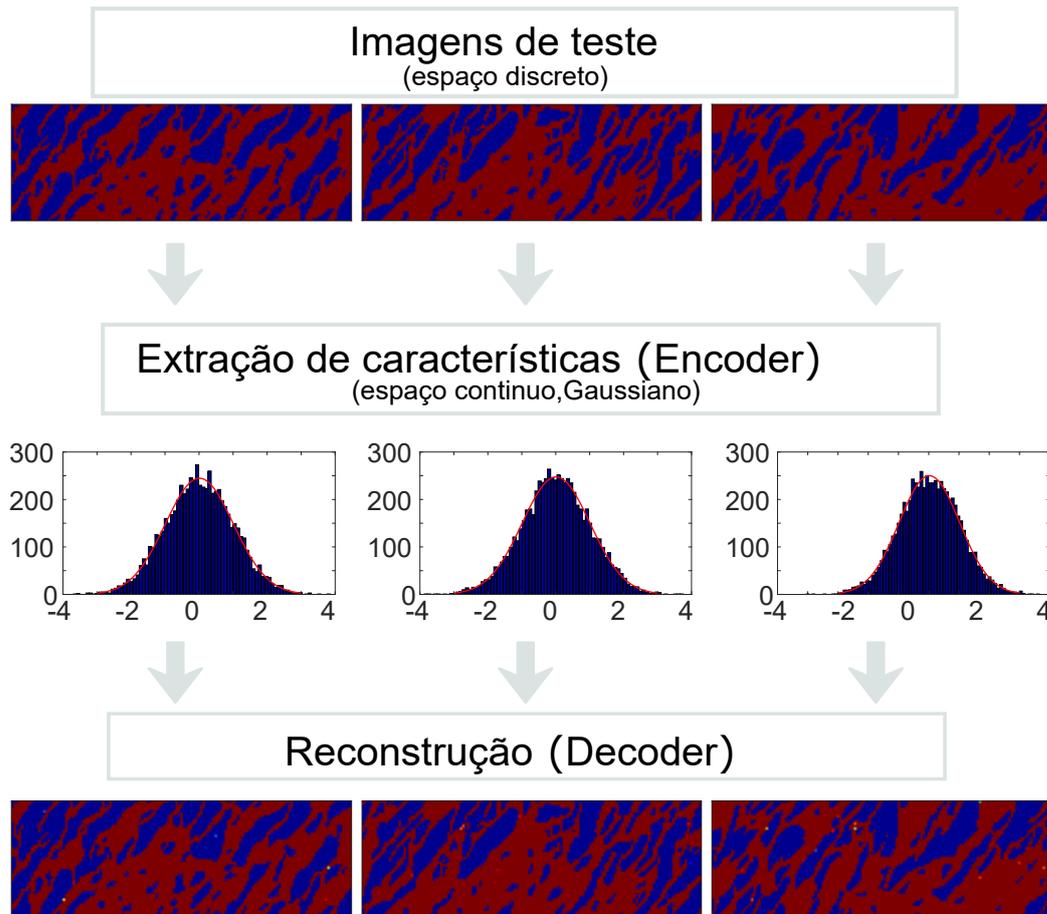


Figura 57: Imagens de fâcies de teste do caso de estudo BRUGGE, utilizando uma rede DBN como AE.

mação das imagens foi feita utilizando o algoritmo *snesim*, como uma imagem de treinamento apresentada por Mariethoz e Caers [132]. A rede com a qual foi feita o experimento é de arquitetura similar aos casos de estudo MPS45 e MPS100, somente variando o número de unidades na camada visível e oculta.

Os resultados obtidos na parametrização e reprodução mediante a rede DBN(6672-2048-6672) são ilustradas na Figura 57, onde similarmente aos resultados apresentados nos casos de estudo anteriores, a parametrização consegue representar uma imagem de fâcies em um espaço contínuo e com distribuição de probabilidade gaussiana. Assim, também é conseguido uma reconstrução das imagens de entrada.

5.2.2.6

Erro de Reconstrução

Para avaliação dos resultados da reconstrução obtida pelas redes DL, formulamos a seguinte equação do erro médio quadrático:

$$Er_t = \frac{100}{N_r} \sum_{j=1}^{N_r} \left(\frac{1}{N_g} \sum_{i=1}^{N_g} (x_{i,j} - \hat{x}_{i,j})^2 \right) \% \quad (5-40)$$

onde $x_{i,j}$ representa o valor da i -ésima posição do vetor \mathbf{x} (entrada à rede) da realização j e $\hat{x}_{i,j}$ representa o valor da i -ésima posição do vetor $\hat{\mathbf{x}}$ (saída da rede) da realização j , N_r e N_g correspondem ao número de realizações da base de teste e a dimensão do vetor \mathbf{x} respectivamente. Aplicando a Equação 5-40 para os diferentes experimentos feitos nesse capítulo, foram obtidos erros de reconstrução menores a 0.3 %.

Tabela 12: Erro de Reconstrução das Redes AEs para Modelos de Fácies

Rede	N_g	$Er_t(\%)$
DBN(532-256-532)-PUNQ-S3	532	0.140
DBN(532-256-532)-MOD-3F	532	0.221
DBN(2045-1024-2045)-MPS45	2025	0.251
DBN(10000-4096-10000)-MPS100	10000	0.266

A Tabela 12 apresenta os valores correspondentes para cada teste realizado utilizando para um número de 1000 realizações de teste ($N_r = 1000$).

Todos os resultados apresentados até aqui, serão reutilizados no próximo capítulo, onde conjuntamente com o método ES-MDA, é proposto uma nova metodologia para o problema de ajuste de histórico. A configuração feita nas redes DL foi sujeita de modo que o número de unidades na camada oculta seja $N_x = N_h$. Contudo, foram feitos outros experimentos, onde a camada oculta apresenta menor dimensão $N_x > N_h$, obtendo-se resultados de reconstrução ótimos. Estes não foram apresentados aqui, dado que o objetivo nesta tese é utilizar uma representação das imagens de fácies dentro do método ES-MDA com a abordagens de localização do ganho de Kalman que requer $N_x = N_h$.

Os resultados do treinamento dos diferentes casos de estudo mostram que as arquiteturas baseadas no DL conseguem representar corretamente as imagens correspondentes a modelos de fácies geológicas. Pode-se ver que quando as características das imagens de treinamento não apresentam muita complexidade (e.g. caso de estudo PUNQ-S3), uma rede AE e RBM pode parametrizar corretamente. Embora a medida que os modelos de fácies contenham características mais complexas, como por exemplo os três últimos casos de estudo, é preciso utilizar arquiteturas com mais camadas. Apesar dos resultados

apresentados neste capítulo correspondem somente a redes de 1 e 3 camadas ocultas, podem ser criadas redes com mais de 3 camadas. Optou-se por esta configuração porque para os casos de estudo apresentados nesta tese foram suficientes.

6

Metodologia de Atualização de Fácies Baseado em Deep Learning

Uma robusta parametrização pode auxiliar os métodos baseados no filtro de Kalman no ajuste de modelos de fácies geológicas sem perder as características de modelagem geológica inicial do reservatório. Sabe-se que esses métodos têm dificuldades quando os parâmetros do vetor de estado apresentam distribuições altamente não gaussianas, como é o caso de modelos de fácies que geralmente assumem valores discretos. Esse fato foi constatado em um modelo simples (PUNQ-S3) apresentado no Capítulo 4.

Este capítulo introduz uma nova metodologia que consiste na integração de uma parametrização de modelos de fácies geológicas baseada em arquiteturas de DL, com o método ES-MDA para a atualização simultânea de modelos de fácies geológicas e propriedades petrofísicas do reservatório no processo de ajuste de histórico. Este método, que será nomeado como ES-MDA-DL no restante desse texto, foi criado com o objetivo de fornecer um procedimento robusto, capaz de ajustar os dados dinâmicos de produção preservando o realismo geológico do modelo *a priori*.

6.1

Procedimento

A metodologia consiste em utilizar a representação de imagens de fácies, aprendida por técnicas de DL com o método ES-MDA. As técnicas baseadas em DL são utilizadas para aprender a representar imagem de fácies em parâmetros contínuos e gaussianos. O método ES-MDA é utilizado para atualizar esses parâmetros em um processo iterativo. Ao início do procedimento é usado uma função *encoder*, que transforma uma imagem de fácies geológica em um vetor de parâmetros contínuos. Esta representação é atualizada iterativamente com o método ES-MDA. Após cada passo iterativo os parâmetros são transformados em imagens de fácies mediante a função *decoder*. As funções *encoder* e *decoder*, são modelos que serão aprendidos durante o processo de treinamento das imagens de fácies por DL.

Na metodologia ES-MDA-DL foi utilizado o método ES-MDA, detalhado no Capítulo 3, e o treinamento de imagens de fácies geológicas apresentado no Capítulo 5. O procedimento completo de método ES-MDA-DL é ilustrado na Figura 58 e está composto por três etapas principais:

1. Etapa de treinamento de imagens de fácies geológicas mediante técnicas de DL com base em AE (**etapa de aprendizado**).
2. Etapa de acoplamento das propriedades petrofísicas do reservatório, com a representação dos modelos de fácies do reservatório (**etapa de representação**).
3. Etapa de atualização, que consiste no processo iterativo do método ES-MDA (**etapa de atualização**).

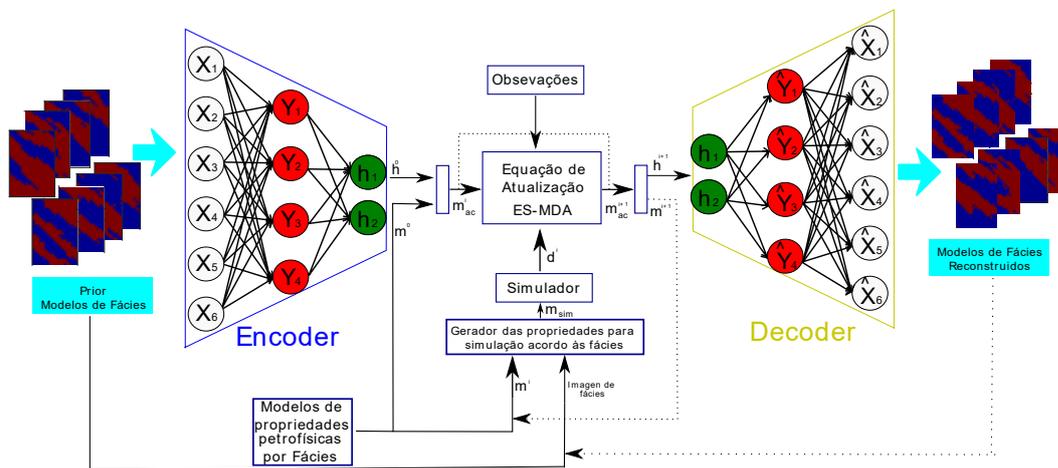


Figura 58: Procedimento do método ES-MDA-DL.

6.1.1 Etapa de Aprendizado

Esta etapa é anterior ao procedimento principal de ajuste de histórico. Nesta etapa foi utilizada uma arquitetura profunda para identificar as principais características das imagens de modelos de fácies. Nesse aprendizado dois modelos podem ser identificados, um de parametrização (função *encoder*) e um de reconstrução (função *decoder*).

$$Encoder : \text{Imagem} \longrightarrow \mathbf{h}$$

$$Decoder : \mathbf{h} \longrightarrow \text{Imagem Reconstruída}$$

A função *encoder* permite uma representação numérica contínua das imagens de fácies, onde esses novos parâmetros, \mathbf{h} , seguem a distribuição gaussiana. Além disso pode-se conseguir uma redução de dimensionalidade. A função *decoder* é usada para reconstruir as imagens de entrada a partir da nova representação delas.

A arquitetura profunda utilizada nesta etapa pode ser escolhida de acordo com a complexidade do problema. Algumas arquiteturas desenvolvidas neste texto foram apresentadas no Capítulo 5. Os resultados mostraram que conseguiram aprender as características da base de treinamento em cada caso de estudo. Entretanto, pode-se utilizar arquiteturas com maior número de camadas ocultas para o treinamento (dependendo da complexidade dos modelos de fácies). A Figura 59 ilustra o processo de treinamento de fácies utilizando uma arquitetura DL com três camadas ocultas.

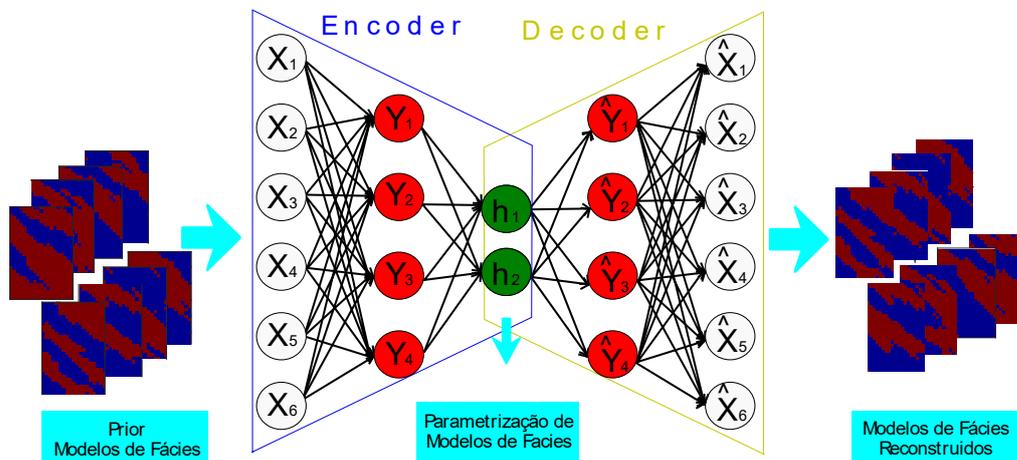


Figura 59: Arquitetura geral do aprendizado profundo de modelos de fácies geológicas.

Para o aprendizado é importante ter um conjunto amplo de imagens que representam fácies geológicas (base de treinamento). Ao final do aprendizado somente serão utilizadas as funções *encoder* e *decoder* nas próximas etapas.

6.1.2 Etapa de Representação

Esta etapa consiste em incrementar o vetor de parâmetros, \mathbf{m} , usualmente utilizado em métodos baseados no filtro de Kalman, com o vetor correspondente à representação das fácies geológicas, \mathbf{h} . Assim, o novo vetor de parâmetros, \mathbf{m}_{ac} , utilizado no procedimento de ajuste de histórico, consiste em

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \mathbf{m}^T]^T, \quad (6-1)$$

onde \mathbf{m} representa as propriedades petrofísicas (e.g., permeabilidade, porosidade para cada fácies) e \mathbf{h} corresponde à representação das imagens de fácies que é obtida da saída da função *encoder* do processo de aprendizado realizado na etapa prévia.

Para melhor entendimento considera-se um modelo de reservatório com duas fácies (F_1 e F_2), e com duas propriedades petrofísicas a atualizar (permeabilidade, κ , e porosidade, ϕ). Neste exemplo o vetor \mathbf{m} é formado por $\mathbf{m} = [\kappa^T, \phi^T]^T$ quando não é considerado as fácies geológicas. Entretanto, caso seja considerado o uso de modelos de fácies é conformado por $\mathbf{m} = [\kappa_{F_1}^T, \kappa_{F_2}^T, \phi_{F_1}^T, \phi_{F_2}^T]^T$, onde κ_{F_i} e ϕ_{F_i} representam a permeabilidade e porosidade na fácies i , respectivamente. Assim, o novo vetor de parâmetros é formado por

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \kappa_{F_1}^T, \kappa_{F_2}^T, \phi_{F_1}^T, \phi_{F_2}^T]^T, \quad (6-2)$$

onde \mathbf{h} é o vetor contínuo e com distribuição gaussiana que representa as imagens de fácies (camada oculta da arquitetura profunda).

6.1.3 Etapa de Atualização

A última etapa consiste em atualizar o novo vetor \mathbf{m}_{ac} mediante o método ES-MDA. Como este método é um procedimento iterativo, em cada passo são atualizados simultaneamente as propriedades petrofísicas de cada fácies e os modelos de fácies (Equação 6-2), com base nos dados históricos de produção observados. Ao final de cada iteração é utilizada a função *decoder* para reconstruir as imagens que representam as fácies atualizadas. Assim, dado a fácies atualizada e as propriedades petrofísicas para cada fácies, é gerado o modelo atualizado para poder simular o reservatório e obter as dados simulados, e continua-se com o processo iterativo.

Considerando o exemplo anterior, o modelo de simulação corresponde ao vetor $\mathbf{m} = [\kappa^T, \phi^T]^T$. Entretanto, quando é incluído os modelos de fácies, o vetor κ é formado pela mistura dos vetores κ_{F_1} e κ_{F_2}

$$\kappa_j = \begin{cases} \kappa_{F_1,j} & ; \text{ quando } j \in F_1 \\ \kappa_{F_2,j} & ; \text{ quando } j \in F_2 \end{cases} \quad (6-3)$$

onde o índice j representa a posição do vetor correspondente, F_1 e F_2 representam as fácies do modelo. Essas fácies são obtidas mediante a reconstrução da imagem a partir da representação do modelo de fácies na iteração atual. Assim para obter um modelo de permeabilidade geral precisa-se misturar as permeabilidades em cada fácies em base à imagem de fácies. Na Figura 60 apresenta-se um exemplo de reconstrução da propriedade de permeabilidade

para um modelo de duas fácies, onde a propriedade é reconstruída baseada no modelos de fácies.

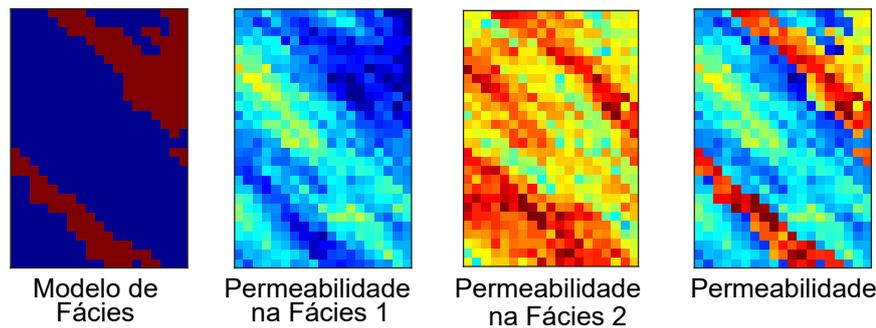


Figura 60: Acoplamento de permeabilidade em cada fácies a um modelo de permeabilidade geral de acordo com a imagem de fácies (esquerda).

Na Figura 58 mostra-se o procedimento esquematizado geral do método ES-MDA-DL proposto, para atualização conjunta de propriedades petrofísicas e modelos de fácies num processo de ajuste de histórico. Aqui pode-se perceber que a rede profunda previamente treinada (Figura 59), foi separado para incorporar a atualização dos modelos de fácies mediante o método ES-MDA. O processo iterativo é ilustrado na parte direita da Figura, onde a função *decoder* é utilizada dentro desse processo. Isto é feito para que em cada iteração as propriedades do reservatório sejam atualizadas com base no aprendizado, ou seja, conservando as características iniciais do modelo de reservatório.

6.2 Casos de Estudo

A metodologia desenvolvida foi testada em quatro modelos de reservatório sintéticos. O primeiro caso de estudo corresponde ao modelo de reservatório PUNQ-S3. Neste caso de estudo o resultado do método ES-MDA-DL são comparados com os resultados dos métodos ES-MDA padrão e ES-MDA-PCA (apresentados no Capítulo 4). Os outros, três casos de estudo foram comparados com o método ES-MDA-OPCA [21]. A razão para comparação com o método ES-MDA-OPCA é porque ele foi o método com melhor desempenho em um estudo anterior [21].

A Tabela 13 detalha algumas das características principais dos casos de estudo, como seu nome, a dimensão do modelo de reservatório, o nome da base de treinamento e o número de fácies do reservatório. No Capítulo 5 foram apresentados resultados do treinamento das imagens de fácies para cada caso de estudo. Entretanto, no primeiro caso de estudo, no Capítulo 5 somente foi apresentado os resultados de treinamento da primeira camada, mas o reservatório PUNQ-S3 apresenta cinco camadas, cada uma com duas fácies.

O treinamento das demais camadas foi feito de forma independentemente utilizando a mesma configuração da rede profunda.

Tabela 13: Casos de Estudo para os testes do Método ES-MDA-DL

Reservatório	Dimensão	Base de Treinamento	Fácies
PUNQ-S3	$19 \times 28 \times 5$	PUNQ-S3	2
MPS45	$45 \times 45 \times 1$	MOD45	2
MPS100	$100 \times 100 \times 1$	MOD100	2
BRUGGE	$139 \times 48 \times 1$	BRUGGE	2

A Tabela 14, resume os parâmetros usados no método ES-MDA-DL, incluindo o número de realizações do conjunto inicial, o número de modelos de fácies utilizadas para o treinamento e o número de épocas utilizadas no processo de treinamento. Estes parâmetros são gerais, para todos os casos de estudo.

Tabela 14: Parâmetros Gerais para os Testes do Método ES-MDA-DL

Parâmetro	PUNQ-S3	MPS45	MPS100	BRUGGE
N_e	200	200	200	200
N_t	5000	5000	5000	2000
N_{epocas}	2000	2000	2000	2000
N_a	4	4	10	6

Em continuação, serão apresentados os resultados da aplicação do ajuste de histórico nos casos de estudo apresentados na Tabela 13.

6.2.1 PUNQ-S3

O caso de estudo PUNQ-S3 já foi apresentado detalhadamente no Capítulo 4. É utilizado este experimento para comparar a abordagem utilizada para atualização de modelos de fácies mediante parametrização baseada no PCA (ES-MDA-PCA) e a metodologia ES-MDA-DL proposta neste capítulo.

As propriedades petrofísicas a atualizar neste caso de estudo são os modelos de porosidade e permeabilidade horizontal e vertical. Assim, o vetor de parâmetros (Equação 6-2) utilizado para este caso de estudo é representado por

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \phi_{F1}^T, \phi_{F2}^T, \log_{10}(\kappa_{h,F1})^T, \log_{10}(\kappa_{h,F2})^T, \log_{10}(\kappa_{V,F1})^T, \log_{10}(\kappa_{h,F2})^T]^T, \quad (6-4)$$

onde \mathbf{h} é a representação dos modelos de fácies (saída do *encoder* da rede profunda treinada), ϕ_{F1} e ϕ_{F2} correspondem aos modelos de porosidade na fácies 1 e 2, respectivamente, e $\kappa_{V,F1}$ e $\kappa_{H,F1}$ são os modelos de permeabilidade vertical e horizontal, respectivamente, na fácies 1.

Para referência foram apresentados os modelos escolhidos com o modelo verdadeiro na Figura 30. Assim também, as curvas de produção do modelo PUNQ-S3 foram apresentadas na Figura 31.

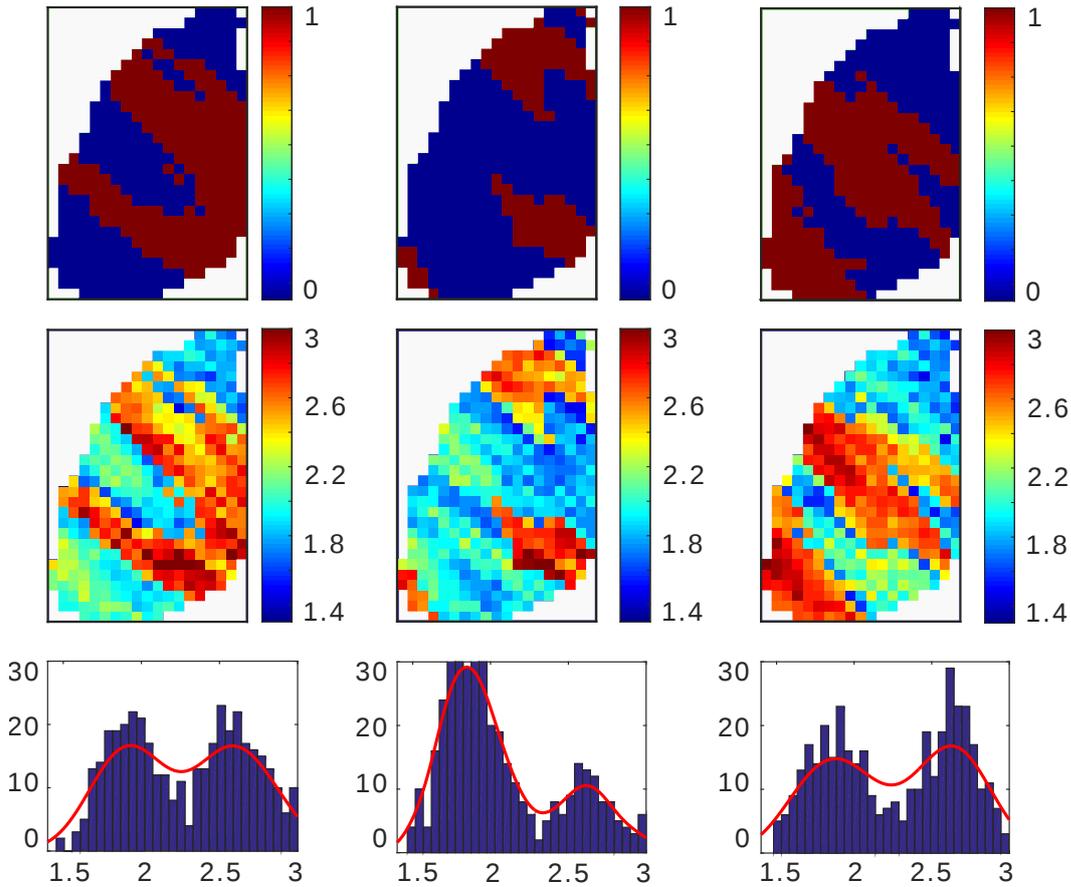


Figura 61: Primeiras três realizações do conjunto inicial (*a priori*) do modelo PUNQ-S3.

6.2.1.1 ES-MDA-DL

Os resultados aqui apresentados correspondem ao método ES-MDA-DL proposto, com quatro assimilações e valores constantes para o fator de inflação, ou seja, $\alpha_k = 4$ para $k = 1, \dots, 4$.

Figura 61 ilustra as três primeiras realizações *a priori* dos modelos de fácies (linha superior), no centro o logaritmo da permeabilidade horizontal, e seus histogramas correspondentes (linha inferior). Esses modelos pertencem à

primeira camada do caso de estudo PUNQ-S3. Nessa figura pode-se observar que os histogramas *a priori* apresentam estruturas de distribuição bi-modais para as realizações de permeabilidade horizontal, sendo similar o comportamento para todas as realizações do conjunto inicial.

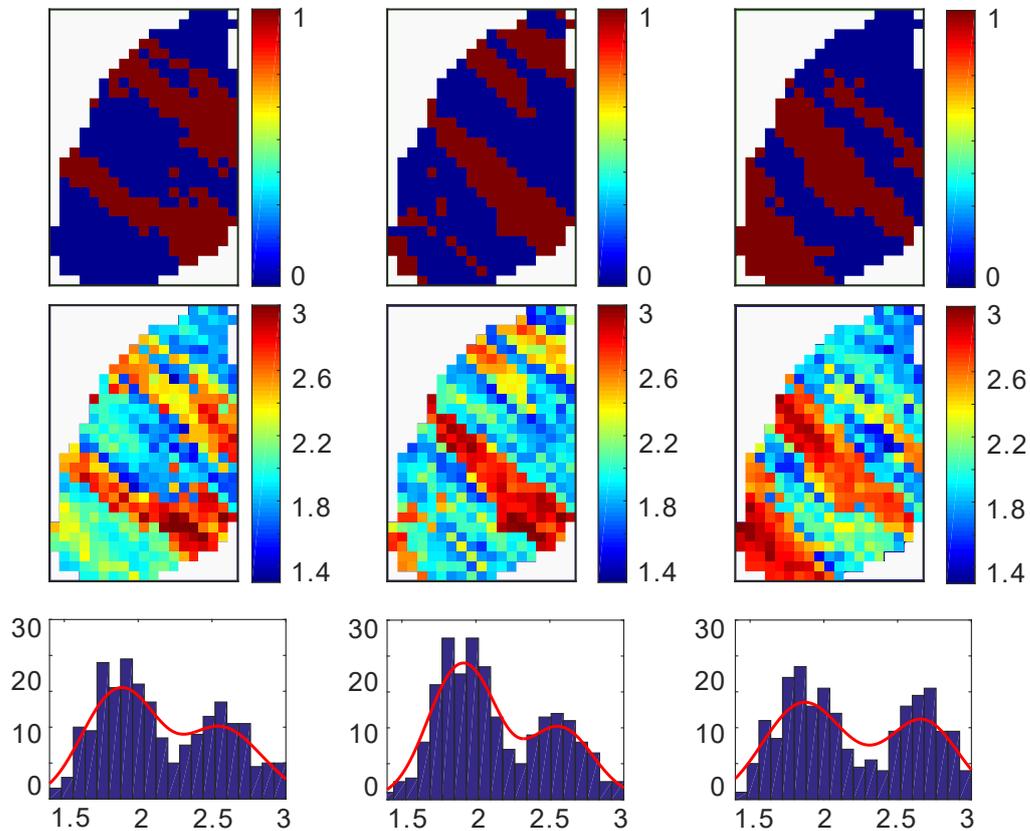


Figura 62: Primeiras três realizações do conjunto final (*a posteriori*) do modelo PUNQ-S3, utilizando ES-MDA-DL.

Ao aplicar o método ES-MDA-DL obtém-se resultados com estruturas similares de distribuição para os modelos atualizados. Isto pode ser visto na Figura 62, onde são ilustrados os modelos *a posteriori* das três primeiras realizações do conjunto para o caso de estudo PUNQ-S3. Tanto os mapas de permeabilidade como os histogramas mantêm as características encontradas nas realizações iniciais, e.g., a distribuição bi-modal do logaritmo da permeabilidade das realizações *a posteriori* (característica encontrada nas realizações *a priori*).

A Figura 63 ilustra os gráficos de produção, para o conjunto de realizações iniciais *a priori* e para os modelos ajustados ao final do procedimento de ajuste de histórico mediante o método ES-MDA-DL. A produção gerada pelo conjunto inicial corresponde às curvas com linhas cinzas e as azuis representam a produção do modelo ajustado. Também são mostrados os dados observados mediante os círculos vermelhos e a média dos modelos finais é representada pela

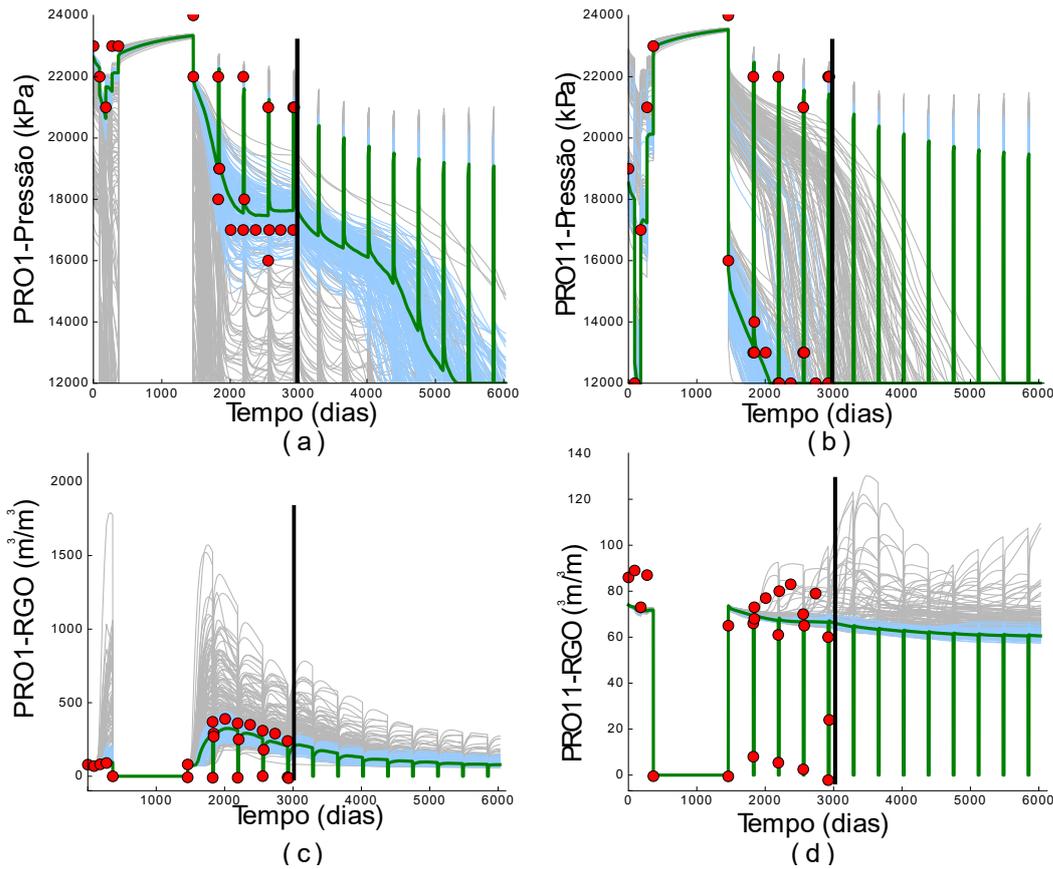


Figura 63: Curvas de produção resultantes do método ES-MDA-DL aplicado ao reservatório PUNQ-S3. (a) Pressão no poço PRO1. (b) Pressão no poço PRO11. (c) RGO no poço PRO1. (d) RGO no poço PRO11.

curva verde. Nesta figura apresenta-se a etapa de histórico (parte esquerda das subfiguras) e a etapa de previsão (parte direita das subfiguras), etapa onde já não existe observações.

Em todas as curvas de produção pôde-se perceber que os dados observados estão próximos das curvas finais (azuis), e a maioria dos modelos finais estão próximos dos dados de produção observados.

O procedimento aplicado conseguiu ajustar os modelos iniciais aos dados de produção observados, além de manter a geologia inicial dos modelos *a priori*. Para comparação dos resultados apresentados neste capítulo com os resultados obtidos no Capítulo 4, onde para o mesmo caso de estudo foi apresentado resultados utilizando os métodos ES-MDA e ES-MDA-PCA. Reescrevendo a função objetivo, (O_N) , utilizada no Capítulo 3 para comparação dos resultados baseado somente nos dados de produção observados. Assim, a função objetivo é calculada mediante

$$O_N = \frac{1}{2N_d} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}))^T \mathbf{C}_d^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m})) . \quad (6-5)$$

O método ES-MDA (sem parametrização) obteve um excelente ajuste de dados observados obtendo um valor de $O_N = 1.9$ (ver Figura 64). Entretanto os mapas do logaritmo da permeabilidade não conservarem características presentes nas realizações *a priori*, como a distribuição bi-modal (ver histogramas da Figura 33). Entretanto os resultados do método ES-MDA-PCA apresentam muita similaridade com os resultados do método ES-MDA-DL no ajuste dos dados de oridução observados. Isto pode ser visto nas curvas de produção das Figuras 36 e 63, e no gráfico de *boxplot* da função objetivo (Figura 64), onde durante o histórico as curvas *a posteriori* ficaram em torno dos dados de produção observados e o valor da função objetivo para esses dois métodos ficaram em $O_N = 2.1$ para o método de ES-MDA-PCA e $O_N = 2.08$ para o método ES-MDA-DL.

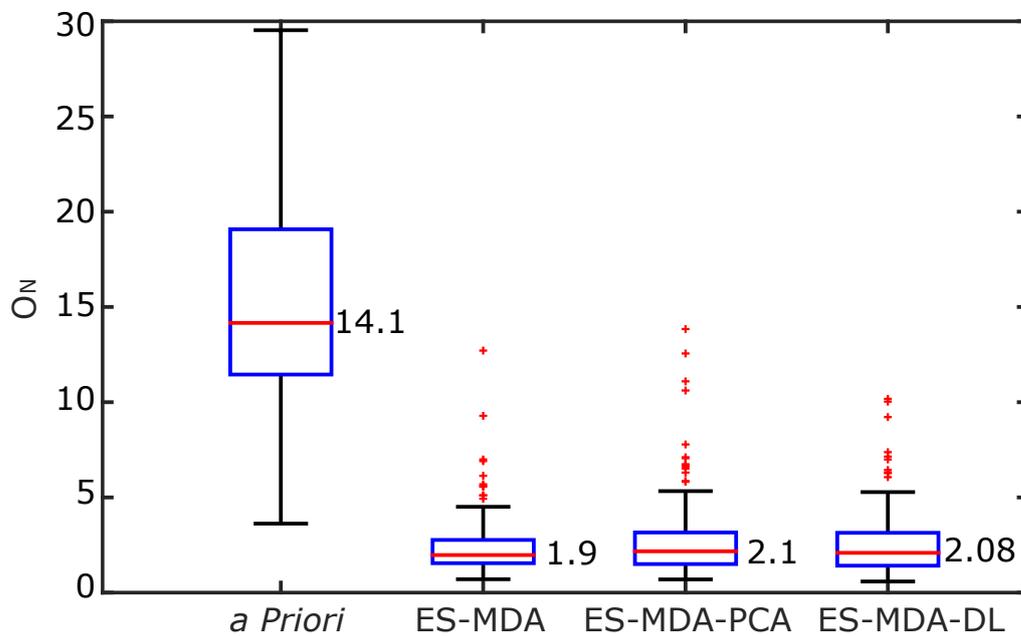


Figura 64: Função objetivo dos métodos estudados, para o caso de estudo PUNQ-S3.

Entretanto observando os mapas do logaritmo da permeabilidade (Figuras 33, 35 e 62), os métodos ES-MDA-PCA e ES-MDA-DL fornecem resultados que conservam a estrutura bi-modal da distribuição do logaritmo da permeabilidade e principais características de geologia do reservatório.

Os resultados desta seção da tese foi apresentando na *The Offshore Technology Conference- Brasil* (OTC-Brasil) [133], como um estudo comparativo dos métodos ES-MDA, ES-MDA-PCA e ES-MDA-DL aplicado ao caso de estudo PUNQ-S3. Entretanto, nesse trabalho o método ES-MDA-DL utilizou uma rede com arquitetura AE para seu aprendizado, e nesta tese o método ES-MDA-DL utilizou uma rede com arquitetura DBN.

6.2.2 MPS45

O caso de estudo MPS45 corresponde a um modelo sintético de reservatório, onde as imagens de fácies geológicas foram criadas pelo algoritmo de MPS. Esse modelo contém duas fácies, onde uma delas são canais curvilíneos, com maior complexidade em comparação ao caso de estudo PUNQ-S3. Para a criação desse modelo foram utilizadas informações da pesquisa realizada por Emerick [21], onde a partir de uma imagem de treinamento, a qual foi apresentada por Cears e Zhan [134], e utilizando o algoritmo *snesim* [93] da MPS, foram criadas realizações de modelos de fácies canalizadas.

Neste experimento foi atualizada conjuntamente, a parametrização dos modelos de fácies e uma propriedade petrofísica (permeabilidade). Assim, o vetor de parâmetros (Equação 6-2), utilizado para este caso de estudo é representado por

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \ln(\kappa_{F1})^T, \ln(\kappa_{F2})^T]^T, \quad (6-6)$$

onde \mathbf{h} é a representação dos modelos de fácies (saída do *encoder* da rede treinada), κ_{F1} e κ_{F2} são os modelos de permeabilidade para as fácies 1 e 2, respectivamente. Aqui é considerado que os valores de permeabilidade em cada fácies são constantes, e dados por

$$\begin{aligned} \kappa_{F1} &= 500 \text{ mD} , \\ \kappa_{F2} &= 5000 \text{ mD} . \end{aligned}$$

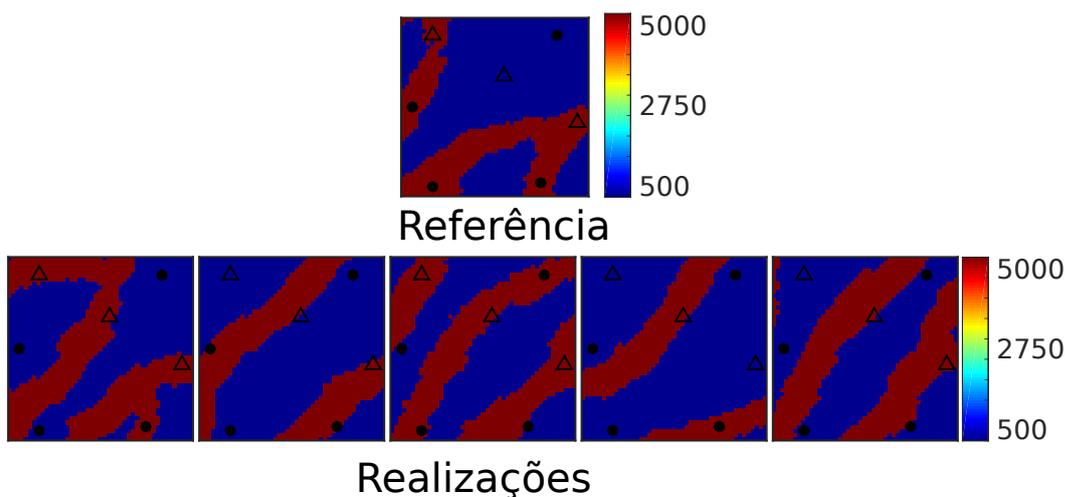


Figura 65: Modelo de permeabilidade de referência (parte de cima) e as primeiras cinco realizações do conjunto inicial (*a priori*) para o caso de estudo MPS45.

A Figura 65 ilustra o modelo escolhido como referência (linha superior) e as 5 primeiras realizações dos modelos de permeabilidade para o caso de estudo MPS45. Como pode ser visto no modelo de referência, existem canais curvilíneos na parte inferior. Nessa figura os círculos pretos indicam a localização dos poços produtores e os triângulos pretos dos poços injetores.

6.2.2.1 ES-MDA

Inicialmente são apresentados os resultados do método ES-MDA padrão. Neste experimento foram usadas $N_a = 4$ assimilações e um conjunto com $N_e = 200$ realizações. Os dados de produção observados correspondem à pressão em três poços, vazão de óleo e água em quatro poços, em um período aproximado de seis anos cada 30 dias, tendo um total de $N_d = 693$ dados de produção observados. Para esse caso de estudo não foi utilizada a abordagem de localização do ganho de Kalman.

A Figura 66 apresenta as cinco primeiras realizações após da aplicação do método ES-MDA. Claramente essas realizações não preservam as características geológicas dos modelos *a priori*. Entretanto a Figura 67, que apresenta as curvas de produção antes e depois do ajuste com ES-MDA, apesar dos modelos não preservarem as características geológicas, eles apresentam um ajuste excelente dos dados de produção observados (círculos vermelhos na Figura 67).

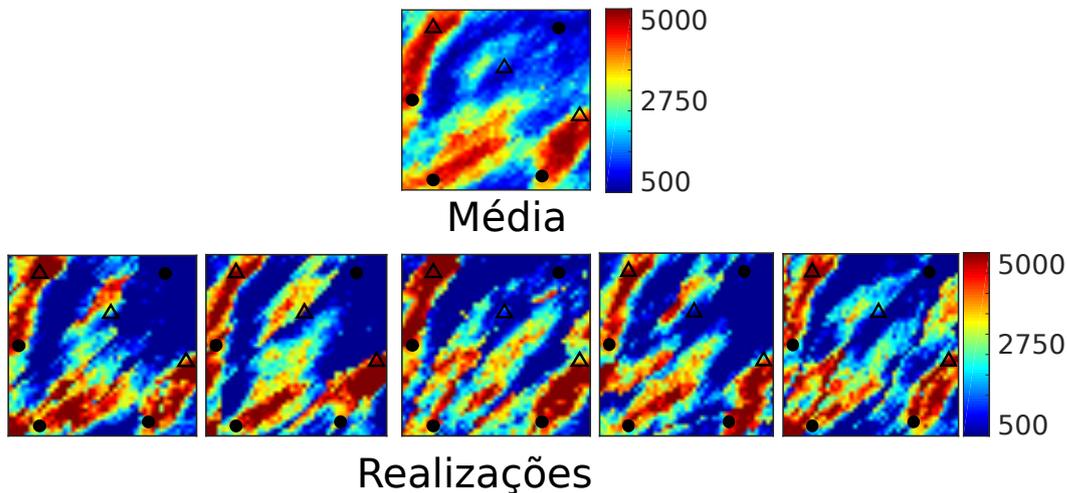


Figura 66: Primeiras cinco realizações de mapas de permeabilidade do conjunto final (*a posteriori*) do caso de estudo MPS45 aplicando o método ES-MDA.

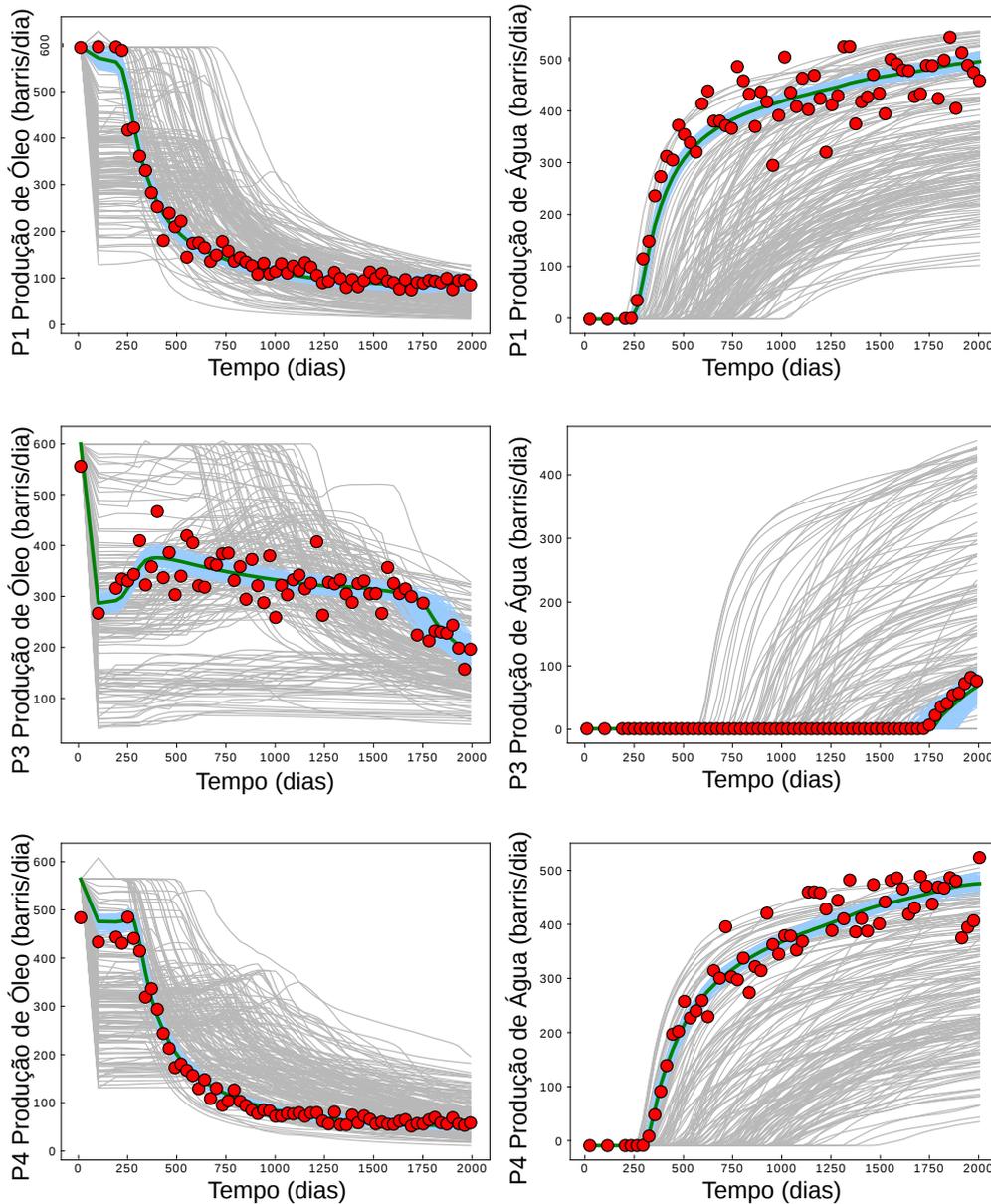


Figura 67: Curvas de produção resultantes do método ES-MDA aplicado ao caso de estudo MPS45. Vazão de produção de óleo e água, para os poços produtores 1, 3 e 4.

6.2.2.2 ES-MDA-DL

Introduzindo uma parametrização para atualizar as realizações de fácies conjuntamente com uma propriedade petrofísica. A Figura 68 apresenta as cinco primeiras realizações obtidas com o método ES-MDA-DL. Neste ajuste foram realizadas quatro iterações do método ES-MDA e utilizando $N_e = 200$ realizações no conjunto inicial. Além disso, a parametrização do método ES-MDA-DL corresponde à arquitetura DBN(2025-1024-2025), a mesma apresentada no Capítulo 5.

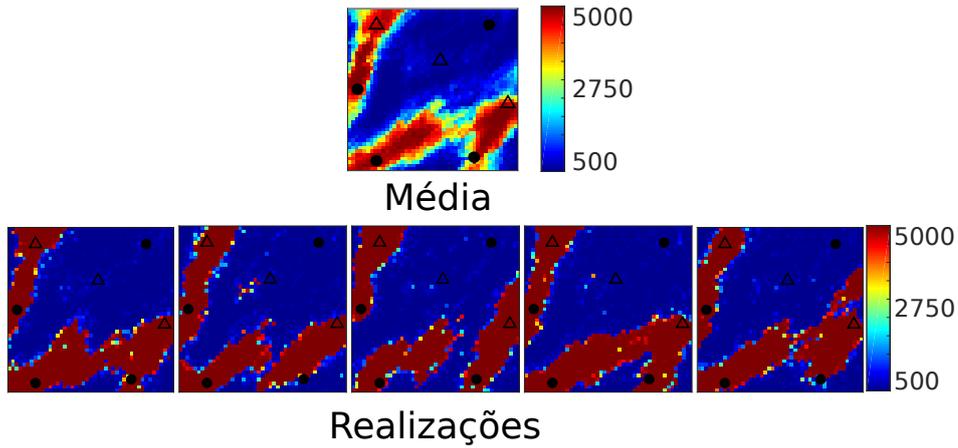


Figura 68: Primeiras cinco realizações de mapas de permeabilidade do conjunto final (*a posteriori*) do caso de estudo MPS45 aplicando o método ES-MDA-DL.

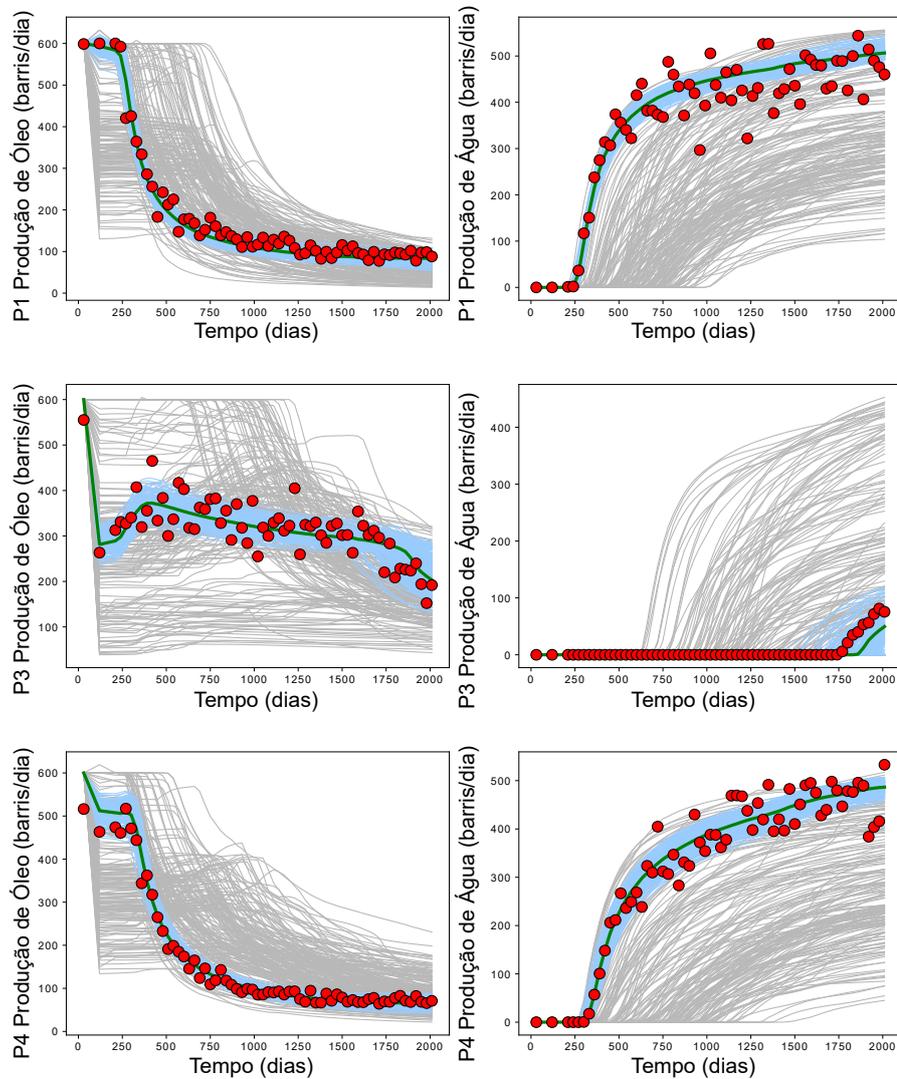


Figura 69: Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo MPS45. Taxa de produção de óleo e água para os poços produtores 1, 3 e 4.

Os resultados da Figura 68 mostram que ES-MDA-DL gerou realizações com fácies bem definidas e canais em posições similares ao modelo de referência (ver linha superior da Figura 65).

A Figura 69 mostra as curvas de produção para o caso de estudo MPS45 indicando que o método ES-MDA-DL resultou também em um ajuste satisfatório dos dados de produção observados (círculos vermelhos na Figura 69).

6.2.2.3 ES-MDA-OPCA

Com objetivo de comparar com os resultados do método ES-MDA-DL com outro tipo de parametrização, foi realizada um ajuste de histórico com o método ES-MDA-OPCA. Usando as mesmas configurações (número de iterações e o mesmo conjunto inicial).

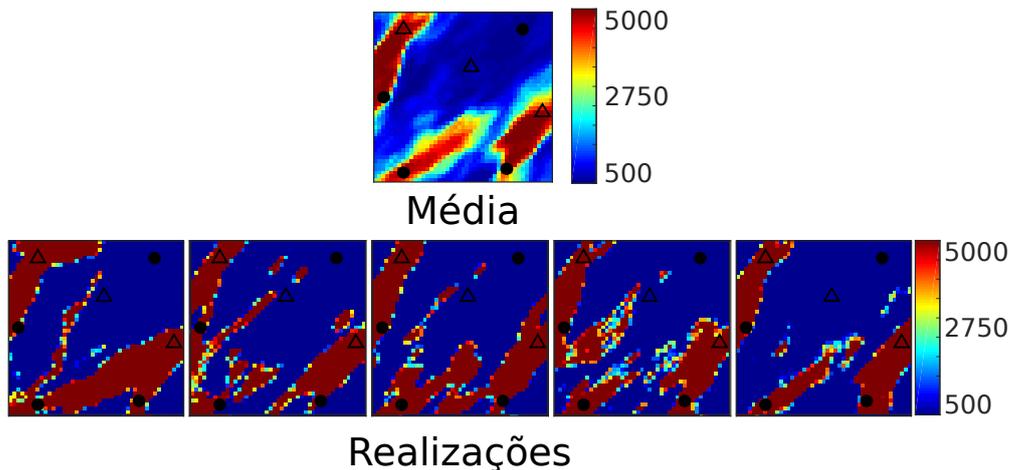


Figura 70: Média e as primeiras cinco realizações de mapas de permeabilidade do conjunto final (*a posteriori*) do caso de estudo MPS45 aplicando o método ES-MDA-OPCA.

As Figuras 70 e 71 apresentam os resultados obtidos para o método ES-MDA-OPCA no processo de ajuste de histórico para o caso de estudo MPS45. As realizações obtidas pelo método ES-MDA-OPCA apresentam fácies bem definidas. Entretanto, comparado ao ES-MDA-DL fica bastante evidente que a parametrização baseada em DL reproduz melhor os canais, gerando realizações mais próximas ao modelo de referência. Em termos de ajuste dos dados de produção observados, observa-se que ambos métodos resultaram em ajustes semelhantes. Note-se que os resultados do método ES-MDA-DL conseguiu modelos finais mais consistentes em comparação aos resultados do método ES-MDA-OPCA, que foi o método com o melhor desempenho em um estudo

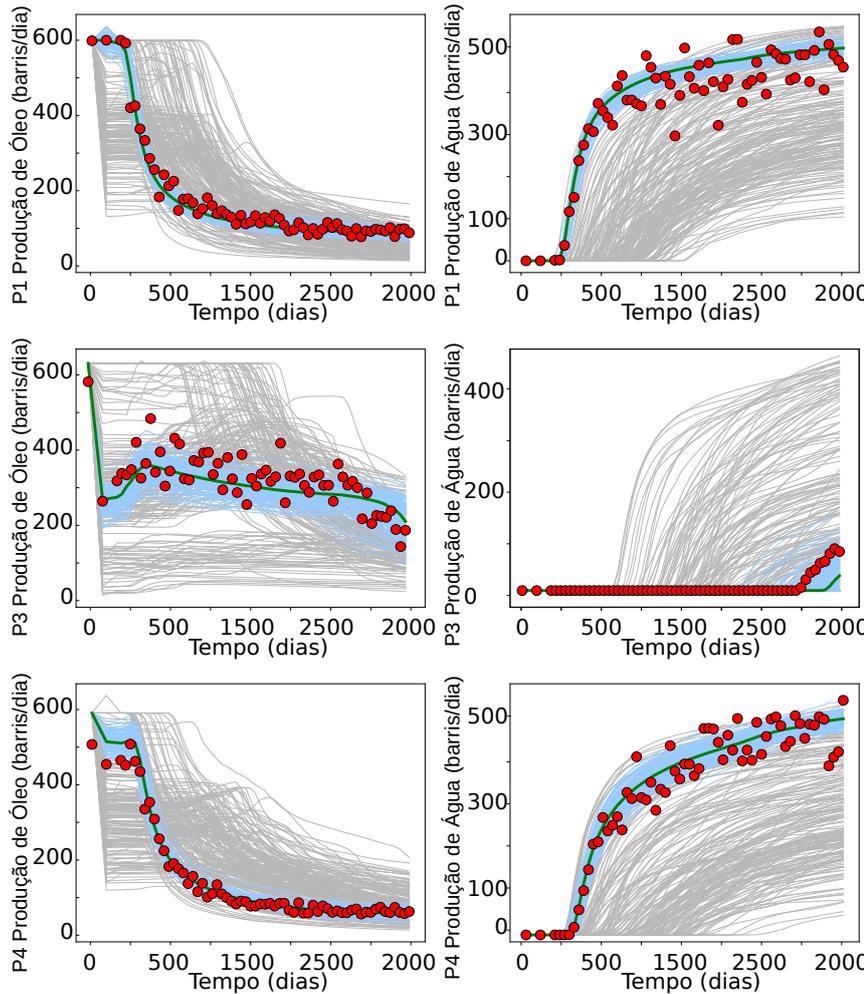


Figura 71: Curvas de produção resultantes do método ES-MDA-OPCA aplicado ao caso de estudo MPS45. Taxa de produção de óleo e água para os poços produtores 1, 3 e 4.

recente comparando estratégias baseadas em PCA [21].

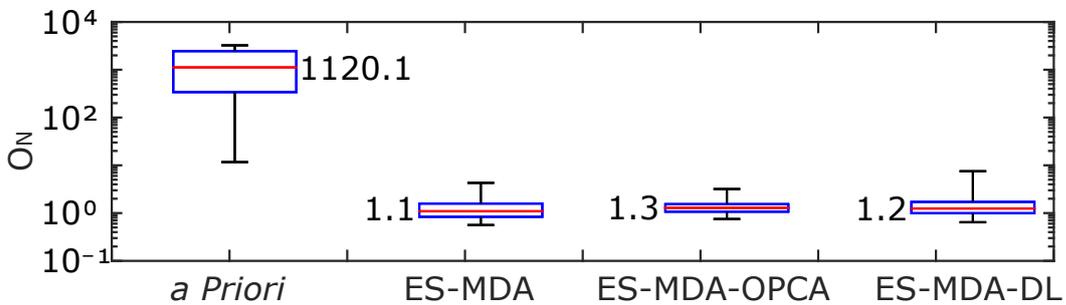


Figura 72: Função objetivo dos métodos estudados, para o caso de estudo MPS45.

Em resumo, o método ES-MDA padrão resultou com um melhor ajuste dos dados de produção observados, como pode ser visto no gráfico de *boxplot*

(Figura 72), que obteve um valor menor no cálculo da função objetivo (Equação 6-5). Contudo, gerou realizações inconsistentes com os modelos *a priori*.

Por outro lado, ES-MDA-DL e o ES-MDA-OPCA buscam preservar a geologia do reservatório ao mesmo tempo que buscam ajustar os dados de produção observados. Ao final, os resultados obtidos apresentam uma melhor representação das realizações de fácies (conservam a geologia inicial do modelo) e um ajuste de histórico muito próximos do método ES-MDA, $O_N = 1.3$ para o método ES-MDA-OPCA e $O_N = 1.2$ para o método ES-MDA-DL, considerando que quanto menor seja o valor da função objetivo, melhor é o ajuste realizado.

Na Figura 73 é apresentada o gráfico de *boxplot* da raiz quadrada do erro médio quadrático (RMSE) da permeabilidade do caso de estudo MPS45, obtida por cada método testado. Os valores da RMSE são calculados como

$$RMSE = \sqrt{\frac{1}{N_m} \|\mathbf{m} - \mathbf{m}_{real}\|_2^2}, \quad (6-7)$$

onde N_m representa a dimensão do vetor \mathbf{m} , sendo \mathbf{m} é o vetor de parâmetros do modelo (e.g., permeabilidade e porosidade) e \mathbf{m}_{real} é o valor de referência do vetor de parâmetros \mathbf{m} .

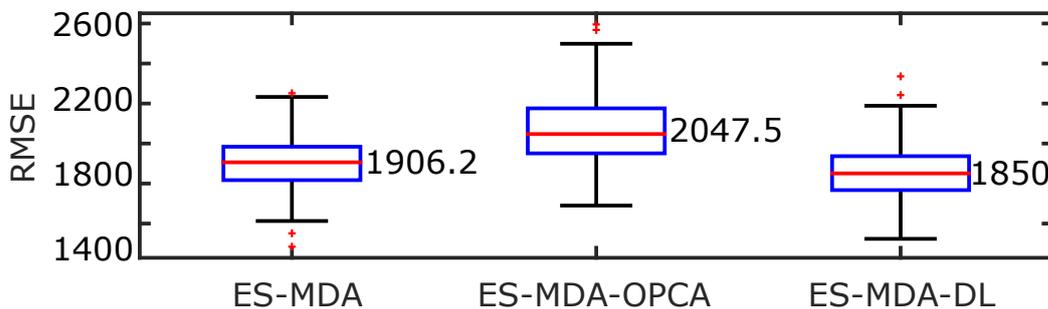


Figura 73: Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo MPS45.

A Equação 6-7 é uma forma de avaliar a diferença entre vetor de parâmetros estimado, \mathbf{m} , e o verdadeiro valor do vetor a ser estimado, \mathbf{m}_{real} . Assim pode-se observar que o método ES-MDA-DL obteve resultados com menor RSME, ou seja, as realizações de permeabilidade ajustadas são mais próximas do modelo de referência.

6.2.3 MPS100

O terceiro caso de estudo é um modelo mais complexo, com canais curvilíneos em diferentes direções. O caso de estudo MPS100 apresenta duas fácies e corresponde a um modelo sintético criado por algoritmos de MPS.

Similarmente ao caso de estudo anterior, o modelo de fácies geológicas foi criado utilizando o algoritmo *snesim*. Mas nesse caso foram criadas imagens de fácies com canais de maior complexidade, de maior dimensão (isto implica maior quantidade de parâmetros no reservatório a aprender no treinamento por DL e maior quantidade de parâmetros a atualizar no método ES-MDA)

Para o caso de estudo MPS100 cada fácies apresenta uma distribuição diferente de valores de permeabilidade (não é considerada constante em cada fácies). Portanto, este caso corresponde ao ajuste simultâneo de fácies e propriedades petrofísicas (permeabilidade). Assim, o vetor de parâmetros (Equação 6-2) utilizado para este caso é formado por

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \ln(\kappa_{F1})^T, \ln(\kappa_{F2})^T]^T, \quad (6-8)$$

onde \mathbf{h} é a representação dos modelos de fácies (saída do *encoder* da rede treinada), κ_{F1} e κ_{F2} são os modelos de permeabilidade para as fácies 1 e 2, respectivamente. Agora os valores de permeabilidade do conjunto inicial (*a priori*) para cada fácies foram obtidas com o algoritmo de simulação gaussiana sequencial. A permeabilidade da fácies *F2* foi gerada com um variograma esférico isotrópico, com uma média de 7 ln-mD e desvio padrão de 0.5 ln-MD. A permeabilidade na fácies *F1* foi gerada com um variograma esférico isotrópico, com uma média de 4 ln-mD e desvio padrão de 0.5 ln-mD.

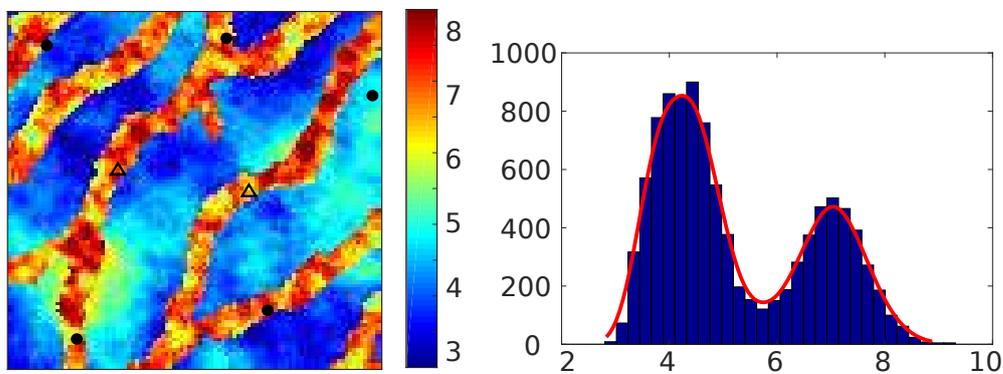


Figura 74: Mapa de logaritmo natural da permeabilidade de referência para o caso de estudo MPS100.

Figura 74 apresenta o logaritmo natural da permeabilidade do modelo escolhido como referência para o caso de estudo MPS100. Nessa figura pode-se observar que o mapa de permeabilidade apresenta uma distribuição bi-modal: um deles no valor aproximado de 4, que corresponde à fácies *F1*, e o segundo aproximadamente no valor de 7, referente à fácies *F2*. Assim, o conjunto inicial cumpre com essas mesmas características, como pode ser visto na Figura 75, onde são apresentadas as quatro primeiras realizações do logaritmo natural da permeabilidade (linha superior) com seus respectivos histogramas (linha

inferior). Similarmente ao caso de estudo anterior os círculos pretos indicam a localização dos produtores e os triângulos dos injetores.

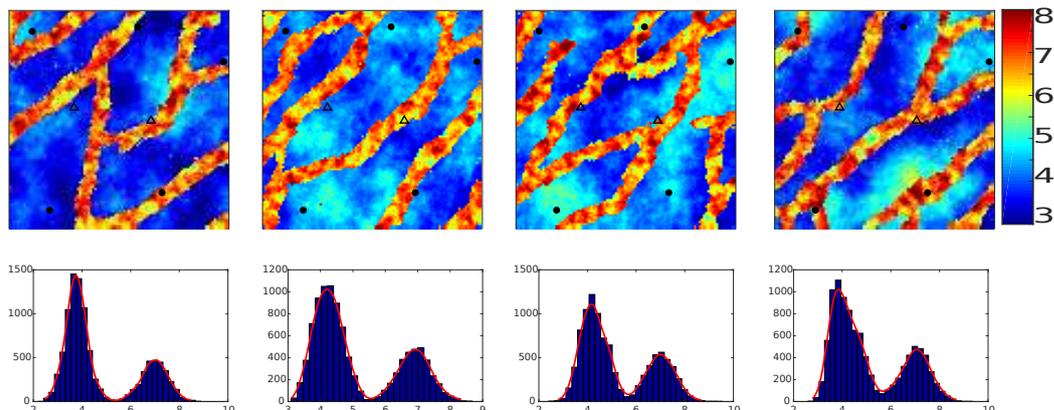


Figura 75: Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto inicial (*a priori*) para o caso de estudo MPS100.

6.2.3.1 ES-MDA

Primeiramente é aplicado o método ES-MDA padrão com um número de $N_a = 10$ assimilações e um conjunto inicial de $N_e = 200$ realizações no caso de estudo MPS100, somente atualizando a propriedade de permeabilidade (ajuste de histórico sem parametrização). Para posteriormente comparar com os resultados obtidos mediante os métodos ES-MDA-DL e ES-MDA-OPCA (ajuste de histórico com parametrização).

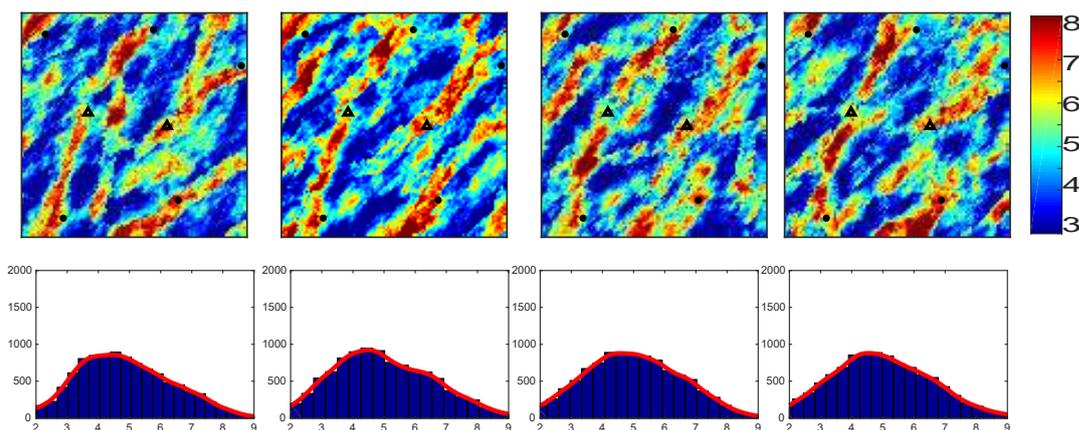


Figura 76: Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final do caso de estudo MPS1000 aplicando o método ES-MDA.

Neste caso de estudo sera utilizada a abordagem de localização do ganho de Kalman apresentada no Capítulo 2, seção 2.2.8. Nos resultados neste

experimento pode-se notar com maior facilidade como os modelos de permeabilidade ajustados são gravemente distorcidos, embora as curvas de produção finais apresentem um excelente ajuste dos dados de produção observados.

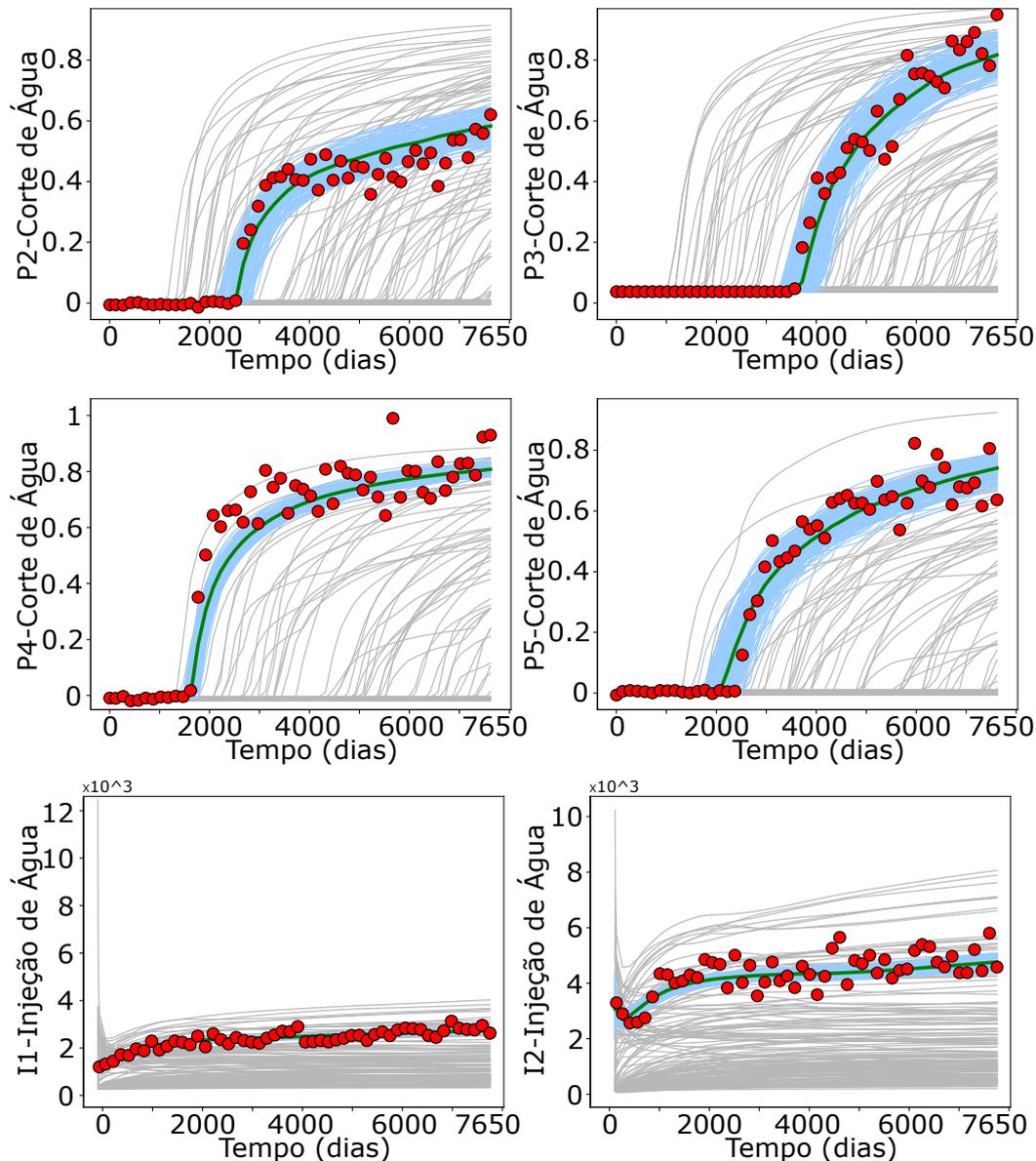


Figura 77: Curvas de produção do método ES-MDA aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.

A Figura 76 apresenta o logaritmo natural da permeabilidade das quatro primeiras realizações do conjunto e seus respectivos histogramas. Claramente percebe-se que não foram mantidas nem as características do modelo *a priori*, já que a distribuição dos modelos ajustados pelo método ES-MDA não apresentam distribuições bi-modais. Praticamente o resultado apresenta uma distribuição gaussiana, de fato, isto acontece devido a que o ES-MDA na sua formulação assume a hipóteses de gaussianidade.

Por outro lado, a Figura 77 ilustra as curvas de produção ajustas dos modelos finais. As curvas apresentam pouca dispersão, onde os dados de produção observados (círculos vermelhos) são ajustados a essas curvas. Isto é uma das características principais em todos os experimentos realizados para o método ES-MDA (realizam um bom ajuste dos dados de produção observados).

6.2.3.2 ES-MDA-DL

Utilizando a método proposto para o caso de estudo MPS100 com número de $N_a = 10$ assimilações no processo de atualização, em comparação aos casos de estudo anteriores. Isto é devido à alta complexidade que apresenta o caso de estudo MPS100. Assim como no experimento com o método ES-MDA padrão é utilizada a abordagem de localização do ganho de Kalman.

O procedimento de aprendizado também foi detalhado no Capítulo 5, mostrando como o treinamento conseguiu aprender as principais características das imagens de fácies e levá-las a um espaço contínuo onde apresenta uma distribuição gaussiana. Na etapa de treinamento foi utilizada a arquitetura profunda DBN(10000-4096-10000).

Os resultados obtidos com a aplicação do método ES-MDA-DL são apresentados nas Figuras 78 e 79. Note-se que o método ES-MDA-DL fornece resultados onde a geoestatística é mantida, ou seja, os histogramas apresentam distribuições bi-modais como nos modelos *a priori* e também conectividades que estão presentes no modelo de referência (Figura 74), aparecem nos modelos ajustados.

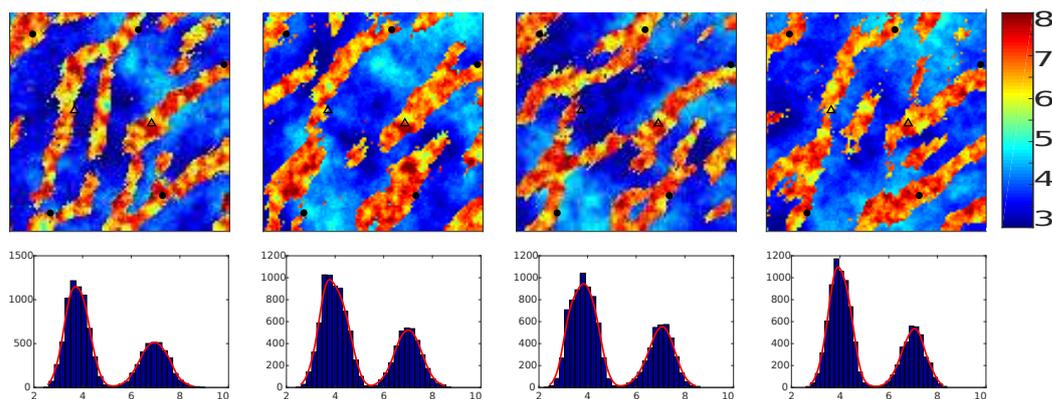


Figura 78: Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) do caso de estudo MPS1000 aplicando o método ES-MDA-DL.

Por outro lado, a Figura 79 mostra as curvas de produção dos modelos iniciais e finais. Embora os modelos ajustados apresentem uma dispersão maior em comparação com os resultados obtidos pelo método ES-MDA, eles

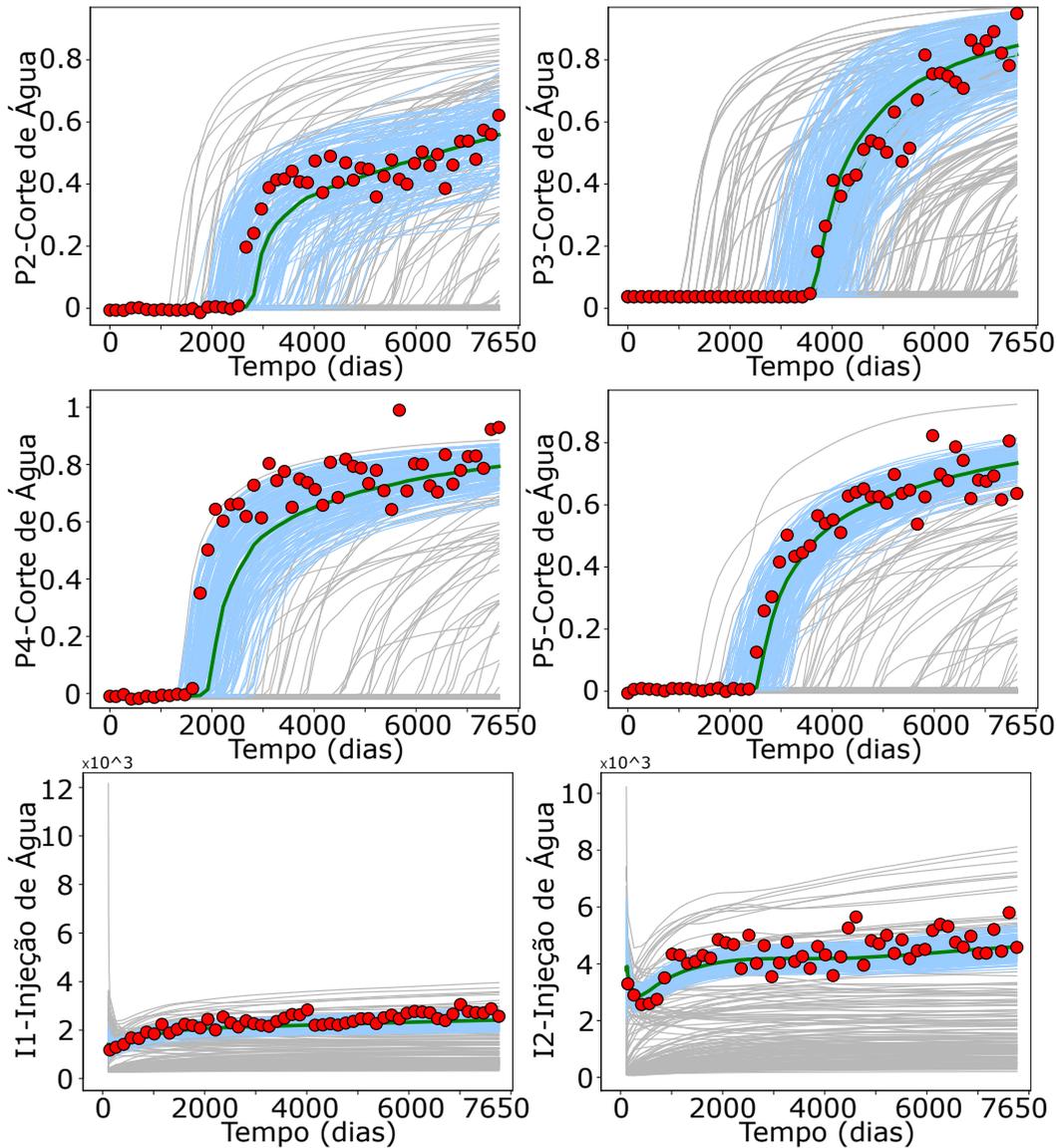


Figura 79: Curvas de produção do método ES-MDA-DL aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.

estão perto dos dados de produção observados (círculos vermelhos). Esta dispersão é atribuída ao fato do caso de estudo MPS100 apresenta uma maior complexidade e a metodologia ES-MDA-DL tenta conservar mais a geologia. Contudo, este ajuste pode ser considerado satisfatório embora seja possível uma melhoria (redução da dispersão) com o uso de um número maior de iterações.

6.2.3.3 ES-MDA-OPCA

Assim como no caso de estudo MPS45, foi usado o método ES-MDA-OPCA para comparação. Neste experimento, foram utilizadas $N_e = 200$

realizações iniciais de modelos de fácies da base de 5000 utilizadas para o treinamento feito pelo método ES-MDA-DL.

As Figuras 80 e 82 apresentam os resultados obtidos pelo método ES-MDA-OPCA. Na primeira figura observa-se que os mapas de permeabilidade são bastante similares, ou seja, as quatro realizações *a posteriori* apresentam estruturas semelhantes, como alguns pequenas alterações. Este comportamento não é percebido nos resultados obtidos mediante ES-MDA-DL, onde as realizações *a posteriori* apresentaram uma maior diversidade, e esses modelos contém estruturas semelhantes às realizações *a priori*, ilustradas na Figura 75.

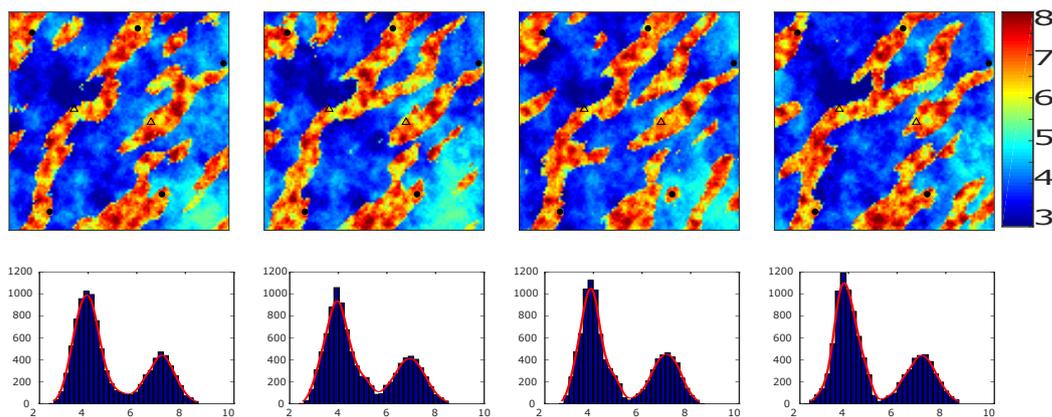


Figura 80: Primeiras quatro realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) do caso de estudo MPS1000 aplicando o método ES-MDA-OPCA.

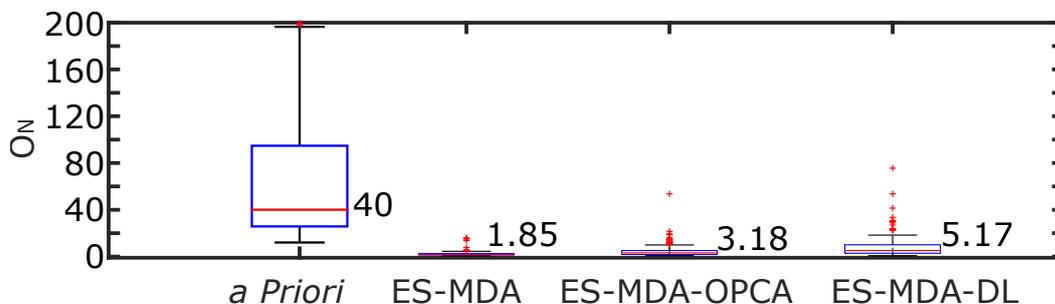


Figura 81: Função objetivo dos métodos estudados, para o caso de estudo MPS100.

A metodologia ES-MDA-OPCA fornece resultados que conservam em grande parte a modelagem geológica inicial das realizações de permeabilidade (histogramas com distribuição bi-modal), similar aos resultados obtidos pelo método ES-MDA-DL. Comparando as curvas de produção apresentadas na Figura 82, nota-se que as curvas ajustadas obtidas pelo método ES-MDA-OPCA apresentam uma menor dispersão que as obtidas mediante o método ES-MDA-DL. Isto é percebido no gráfico de *boxplot* da função objetivo (Figura 81).

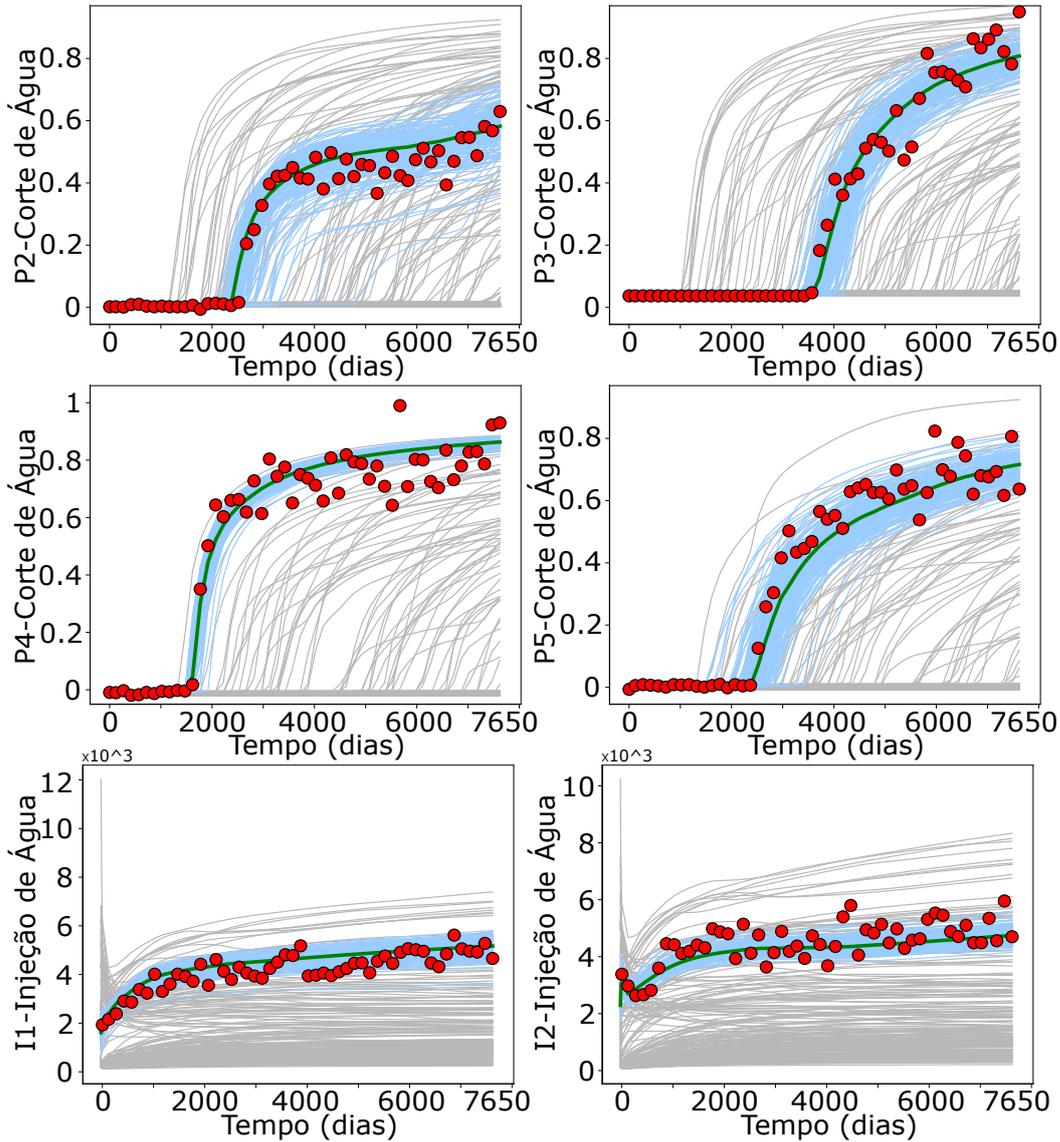


Figura 82: Curvas de produção do método ES-MDA-OPCA aplicado ao caso de estudo MPS100. Corte de água nos poços 2 a 5 e injeção de água nos injetores 1 e 2.

Neste caso de estudo o método ES-MDA padrão (sem parametrização) obteve o melhor resultado baseado na função objetivo (Equação 6-5). Entretanto, os métodos ES-MDA-OPCA e ES-MDA-DL conseguem conservar a geologia e distribuições bi-modais dos modelos *a priori* com um valor da função objetivo relativamente baixo (para o método ES-MDA-OPCA obteve um valor de $O_N = 3.18$ e o método ES-MDA-DL obteve um valor de $O_N = 5.1$). Onde incrementado o número de iterações pode-se obter valores de O_N menores.

Calculando a raiz quadrada do erro médio quadrático (Equação 6-7) da permeabilidade, baseado no modelo de permeabilidade de referência (Figura 74). Note-se que o menor valor do RMSE é obtido pelo método ES-MDA-DL, como é visto no gráfico de *boxplot* do RMSE para o caso de estudo MPS100

(Figura 83). Um menor valor do RMSE indica que os modelos finais ajustados estão mais próximos ao modelo de referência.

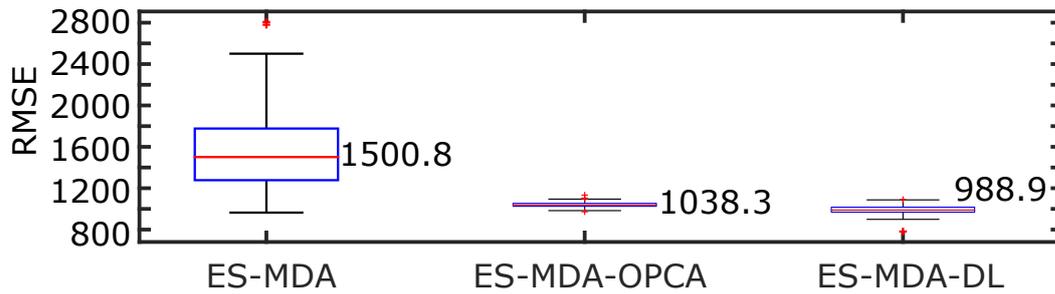


Figura 83: Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo MPS100.

6.2.4 BRUGGE

O último caso de estudo considerado é um modelo de reservatório sintético, que contém duas fácies com estruturas complexas. Esse caso de estudo corresponde a uma versão modificada do modelo BRUGGE [94, 95] com apenas uma camada. As realizações de fácies foram geradas com o algoritmo *snessim* de forma similar aos casos de estudo MPS45 e MPS100. Este experimento também inclui o ajuste simultâneo dos modelos de fácies e de modelos de permeabilidade. Assim, o vetor de parâmetros (Equação 6-2), utilizado é formado por

$$\mathbf{m}_{ac} = [\mathbf{h}^T, \ln(\kappa_{F1})^T, \ln(\kappa_{F2})^T]^T, \quad (6-9)$$

onde \mathbf{h} é a representação dos modelos de fácies (saída do *encoder* da rede treinada) e κ_{F1} e κ_{F2} são os modelos de permeabilidade para as fácies 1 e 2, respectivamente.

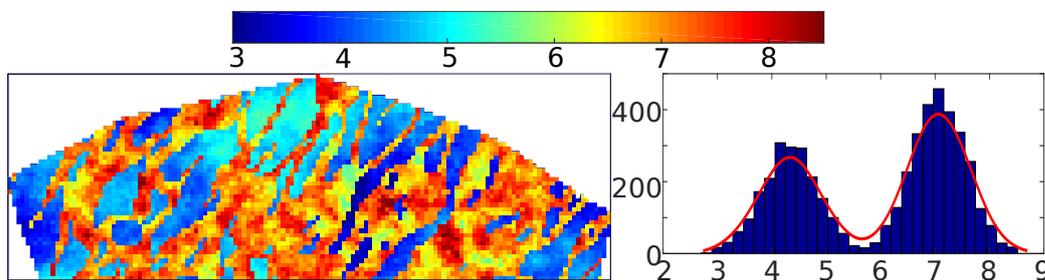


Figura 84: Mapa de logaritmo natural da permeabilidade de referência para o caso de estudo BRUGGE.

De forma similar ao caso de estudo MPS100 os valores de permeabilidade do conjunto inicial (*a priori*) dentro de cada fácies foram obtidas com o

algoritmo de simulação gaussiana sequencial. Na Figura 84 é apresentado o logaritmo natural da permeabilidade escolhido como referência para o caso de estudo BRUGGE.

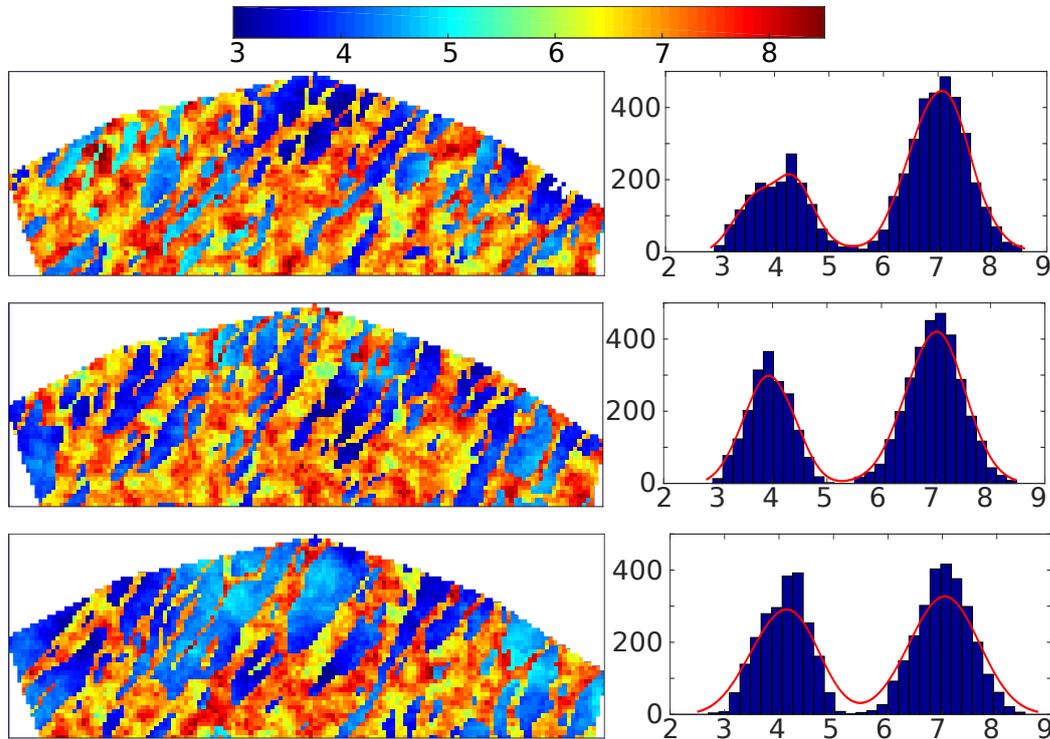


Figura 85: Primeiras três realizações do logaritmo natural da permeabilidade do conjunto inicial (*a priori*) para o caso de estudo BRUGGE.

No modelo de referência pode-se observar que o mapa do logaritmo natural da permeabilidade apresenta uma distribuição bi-modal. Assim, também o conjunto inicial cumpre com essas mesmas características, como pode ser visto na Figura 85, onde são apresentadas as três primeiras realizações do logaritmo natural da permeabilidade com seus respectivos histogramas.

6.2.4.1 ES-MDA

O primeiro experimento realizado para o caso de estudo BRUGGE, é a atualização dos modelos de permeabilidade (sem parametrização) utilizando o método ES-MDA padrão com $N_a = 6$ assimilações e um total de $N_e = 200$ realizações no conjunto inicial. Para esse caso de estudo foi utilizada a abordagem de localização do ganho de Kalman similar à utilizada no caso de estudo MPS100.

As Figuras 86 e 87 apresentam os resultados obtidos. Onde similar aos casos de estudo anteriores: o método ES-MDA faz um ajuste excelente dos

dados de produção observados. Contudo, os modelos resultantes não preservam as características da modelagem *a priori*.

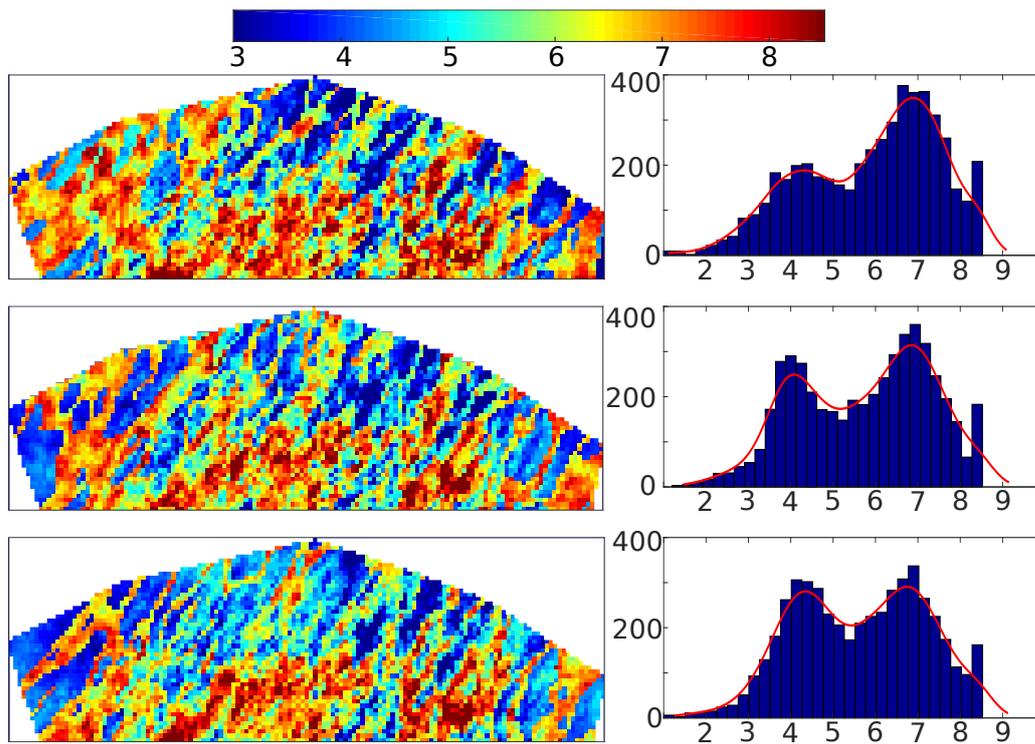


Figura 86: Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) para caso de estudo BRUGGE aplicando o método ES-MDA.

6.2.4.2 ES-MDA-DL

Para a aplicação do método ES-MDA-DL ao caso de estudo BRUGGE foi utilizada o treinamento de modelos de fácies apresentada na parte final do Capítulo 5. Esse procedimento de aprendizado utilizou a arquitetura profunda DBN(6672-2048-6672). Assim, com $N_a = 6$ assimilações e um total de $N_e = 200$ realizações no conjunto inicial foi obtida os resultados apresentados nas Figuras 88 e 89.

Na Figura 88 são apresentados os mapas do logaritmo natural da permeabilidade das três primeiras realizações *a posteriori* com seus respectivos histogramas. Note-se que esses histogramas conservam a distribuição bi-modal dos modelos iniciais (Figura 85). Além disso, os mapas de permeabilidade finais preservam a estrutura das realizações iniciais, ou seja, as variações entre os modelos ajustados e os modelos *a priori* são mínimas.

A Figura 89 ilustra as curvas de produção dos modelos finais obtidos pelo método ES-MDA-DL. Embora essas curvas ajustados apresentem uma maior

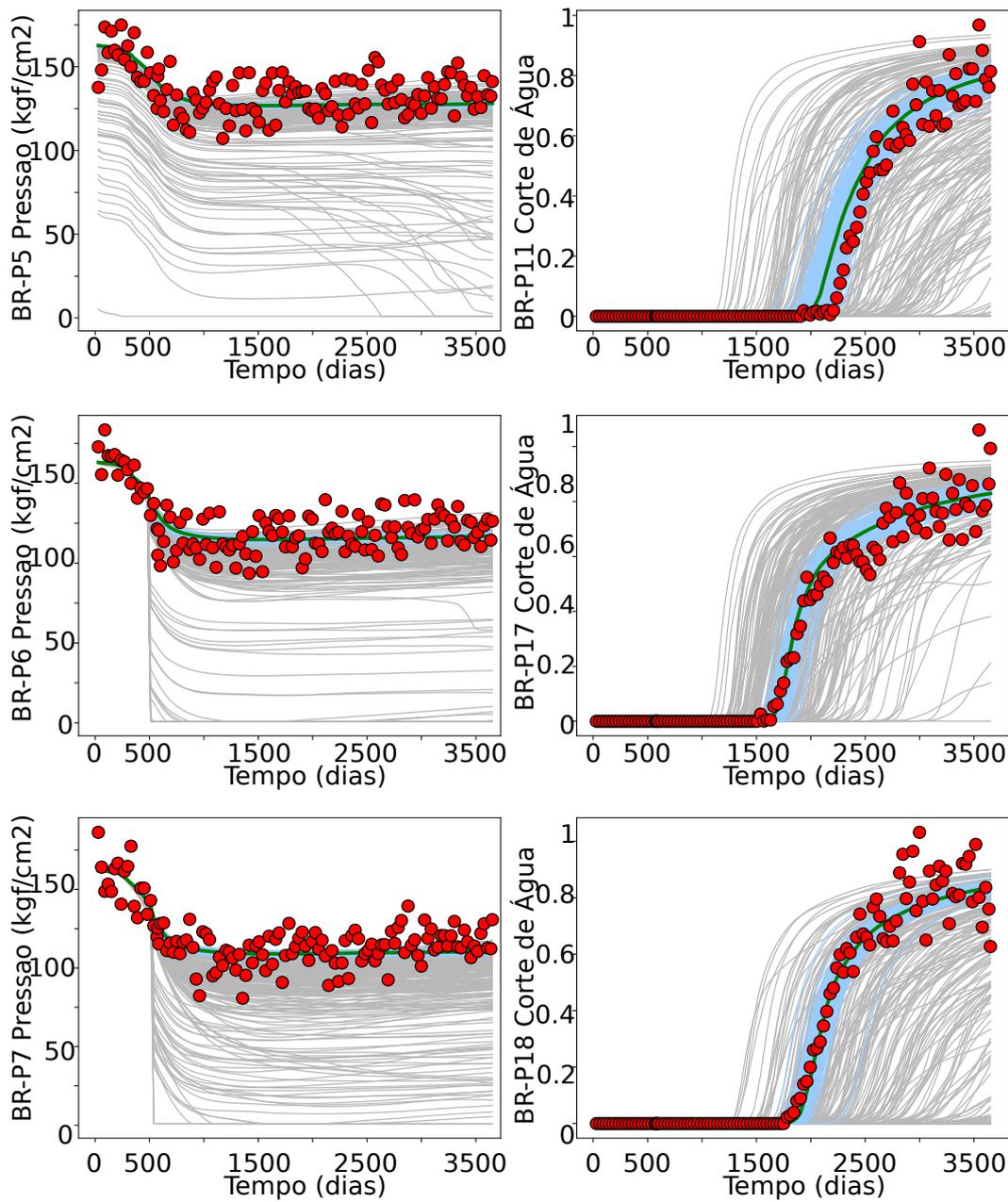


Figura 87: Curvas de produção resultantes do método ES-MDA aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.

dispersão em comparação com o resultado obtido pelo método ES-MDA, eles estão perto dos dados de produção observados, similar aos casos de estudo apresentados previamente o ES-MDA fornece um melhor ajuste nas curvas de produção. Embora, o método ES-MDA-DL conserva a modelagem *a priori* do reservatório.

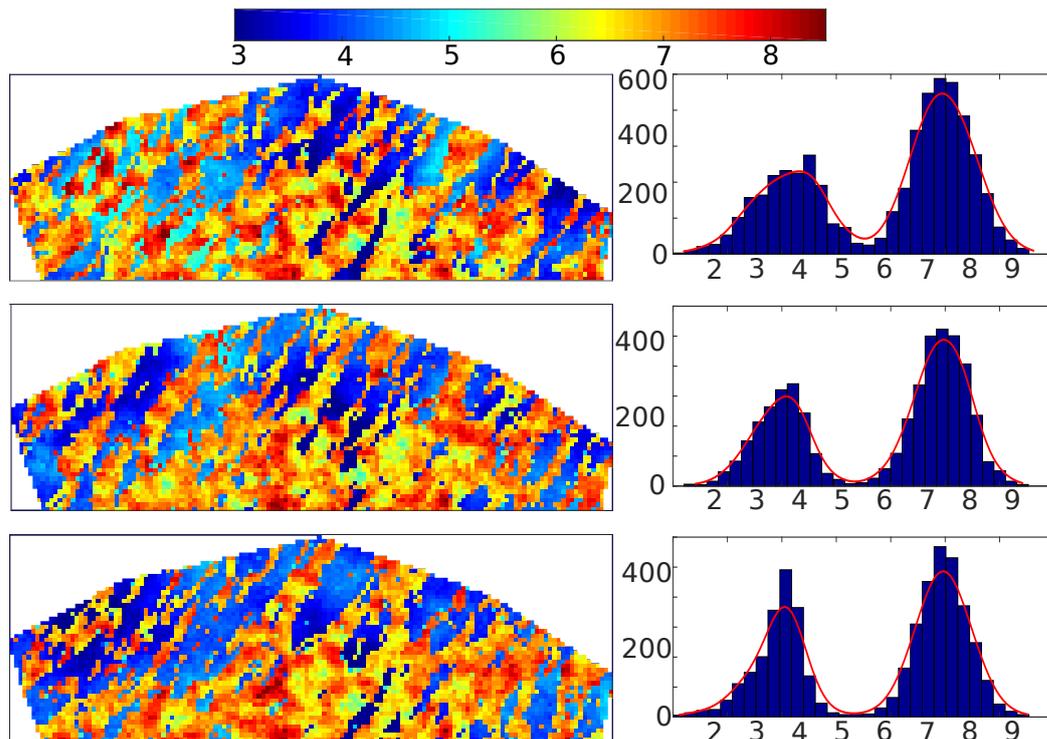


Figura 88: Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) para caso de estudo BRUGGE aplicando o método ES-MDA-DL.

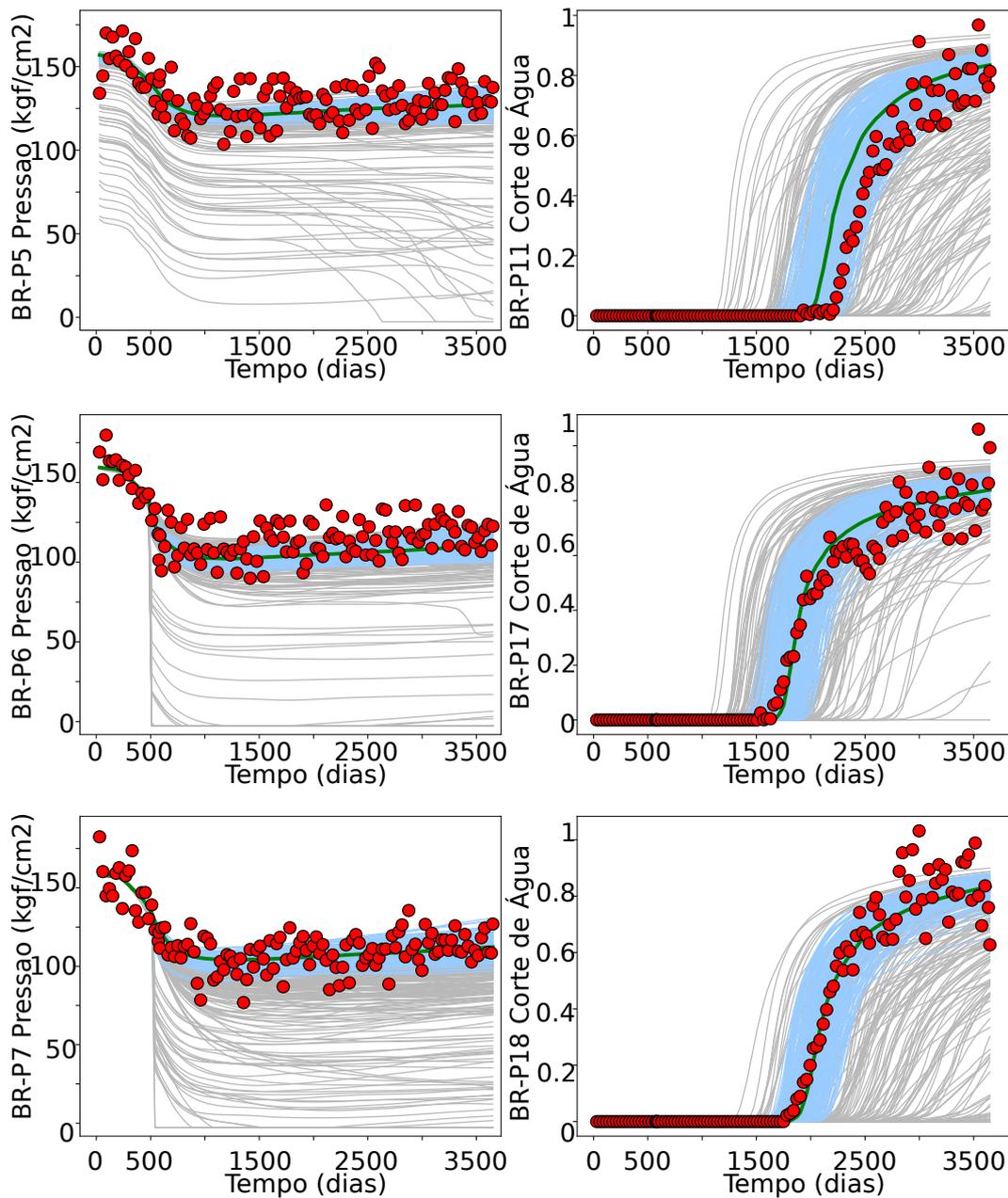
6.2.4.3 ES-MDA-OPCA

Similarmente aos casos de estudo MPS45 e MPS100 foi feito um experimento para o caso de estudo BRUGGE utilizando a metodologia ES-MDA-OPCA com o mesmo número de assimilações e número de realizações iniciais.

Neste experimento, foram utilizadas $N_r = 2000$ realizações de fácies geológicas para a construção das bases de componentes principais. Essas realizações são as mesmas usadas no treinamento do método ES-MDA-DL.

A Figura 90 mostra as três primeiras realizações depois do ajuste de histórico, mostrando resultados conservando a distribuição bi-modal do logaritmo natural da permeabilidade, similarmente aos resultados obtidos no ES-MDA-DL. Apesar que os mapas do logaritmo natural da permeabilidade apresentam estruturas similares às realizações iniciais, elas perdem os canais finos encontrados nos modelos *a priori* (Figura 85). Os resultados mostrados pelo ES-MDA-DL apresenta mapas do logaritmo natural da permeabilidade muito semelhantes aos modelos *a priori*.

Na figura 91 apresenta as curvas de produção resultantes depois do processo de ajuste de histórico, esses resultados ficaram muito parecidos aos obtidos pelo método ES-MDA-DL nas curvas de produção da pressão (parte



PUC-Rio - Certificação Digital N° 1313464/CA

Figura 89: Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.

esquerda) e com menor dispersão nas curvas de corte de água (parte direita).

Neste experimento, pode ser visto o mesmo comportamento como em todos os casos de estudo analisados. O ajuste de histórico dos dados de produção observados obtido pelas metodologias ES-MDA-OPCA e ES-MDA-DL são aceitáveis. Embora o ajuste de histórico, sem considerar modelos de fácies tenha menor dispersão. Isto pode ser detalhado analisando a função objetivo (Equação 6-5) para os experimentos realizados no caso de estudo

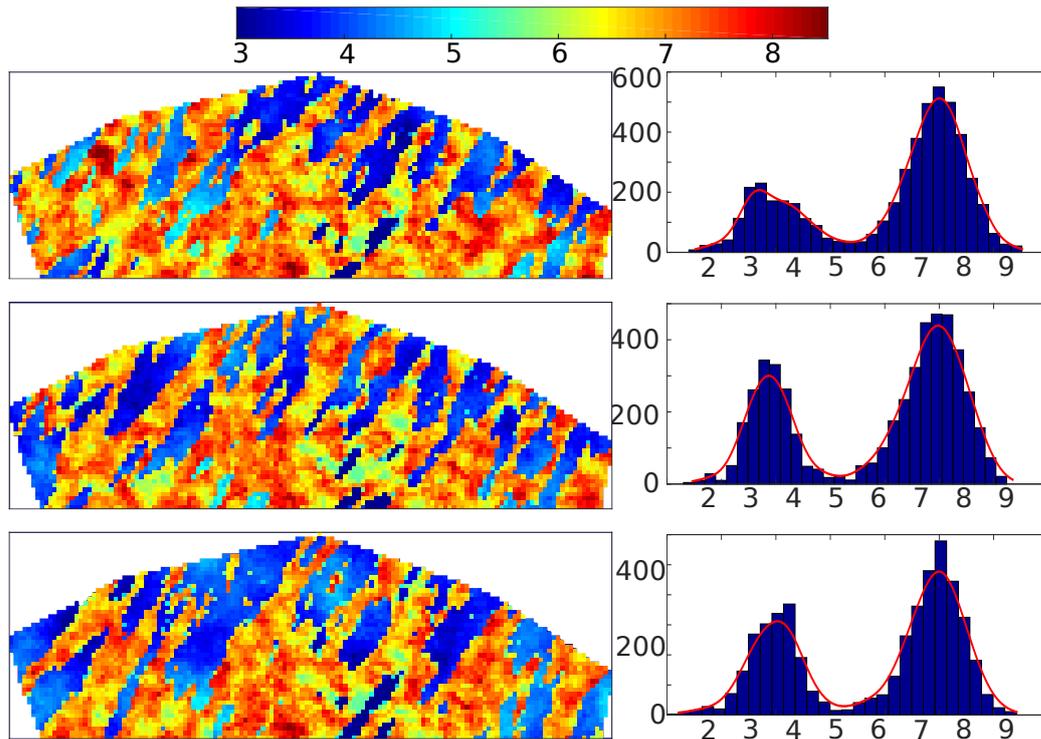


Figura 90: Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) para caso de estudo BRUGGE aplicando o método ES-MDA-OPCA.

BRUGGE.

Assim, a Figura 92 apresenta o gráfico *boxplot* da função objetivo para o caso de estudo BRUGGE. Note-se que o melhor ajuste baseado nessa métrica é para o método ES-MDA padrão ($O_N = 2.5$). Entretanto, o método ES-MDA-DL obteve um valor maior de $O_N = 10.2$. Este comportamento ocorre, porque atualizar modelos de fácies e tentar conservar a geologia *a priori* incrementa a complexidade do modelo. Esse problema pode ser resolvido incrementando o número de iterações (número de assimilações do método ES-MDA), como é apresentado no gráfico *boxplot* para ES-MDA-DL(20). Utilizando $N_a = 20$ iterações no método ES-MDA-DL pode-se obter um valor da função objetivo ($O_N = 4.52$) similar ao método ES-MDA-OPCA ($O_N = 4.47$).

Comparando os resultados utilizando a métrica da raiz quadrada do erro médio quadrático (Equação 6-7), e utilizando o modelo de permeabilidade de referência (Figura 84). A Figura 93 apresenta o gráfico de *boxplot* do RMSE para o caso de estudo BRUGGE. Note-se que o menor valor do RMSE é obtido pelo método ES-MDA-OPCA ($RMSE = 997.2$). Entretanto o método ES-MDA-DL obteve um valor $RMSE = 1037.8$ muito próximo ao obtido pelo método ES-MDA-OPCA, além disso, o incremento do número de iterações também diminui o valor do RMSE (ES-MDA-DL(20) obteve um

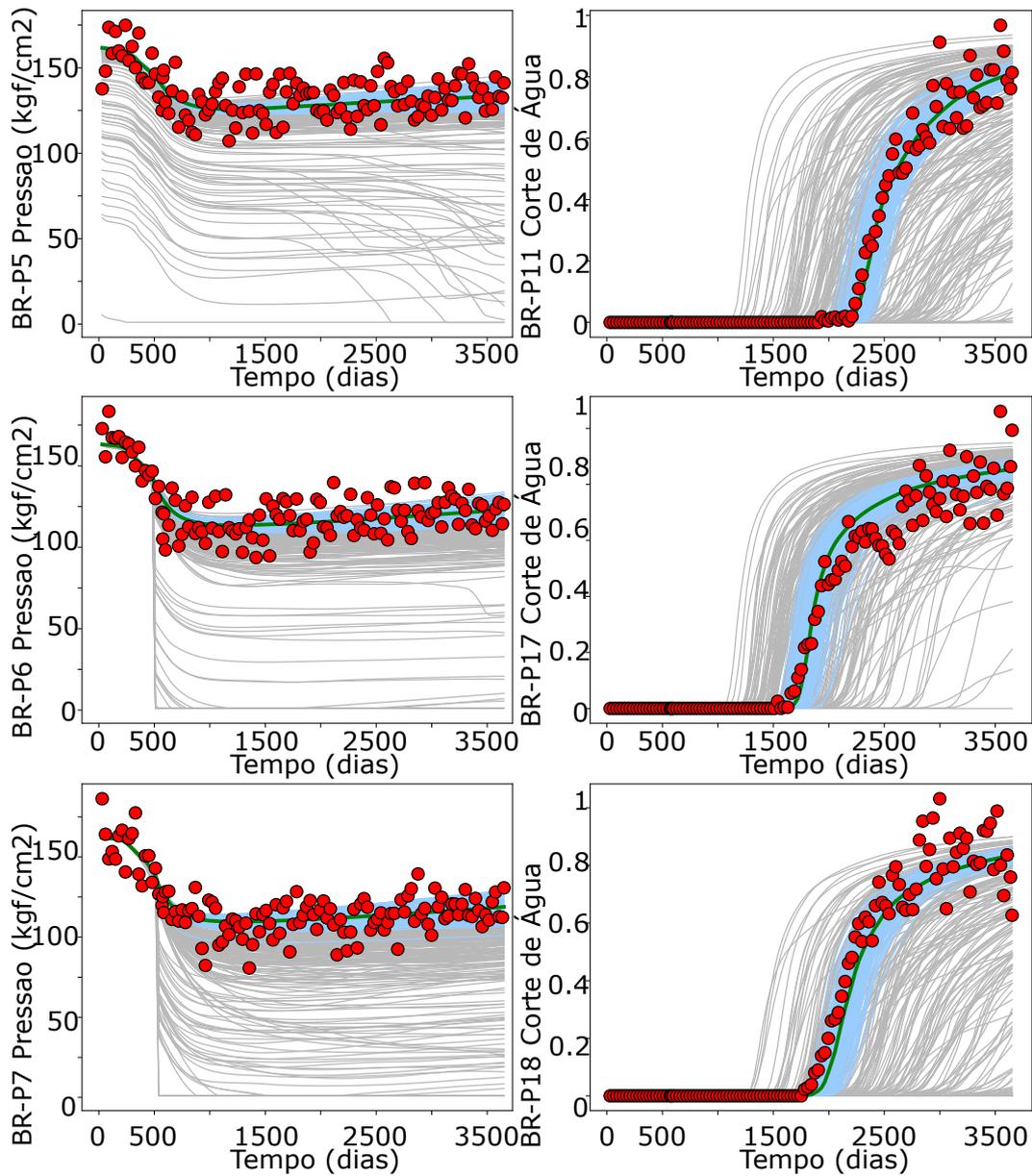


Figura 91: Curvas de produção resultantes do método ES-MDA-OPCA aplicado ao caso de estudo BRUGGE. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.

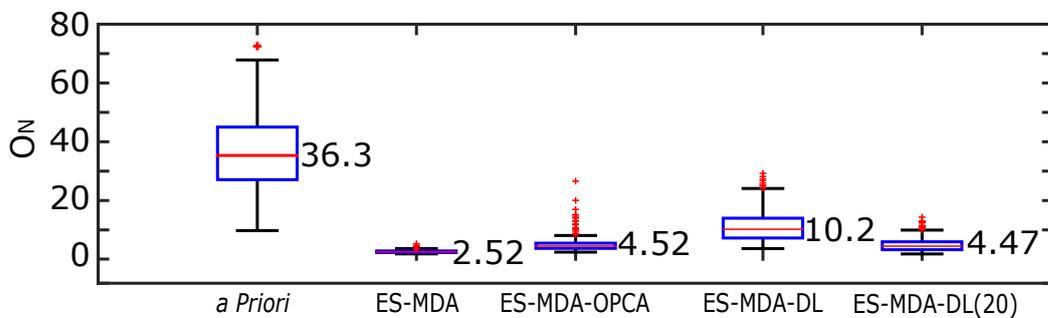


Figura 92: Função objetivo dos métodos estudados, para o caso de estudo BRUGGE.

$RMSE = 1003.9$). Um menor valor do RMSE indica que os modelos finais ajustados estão mais próximos ao modelo de referência.

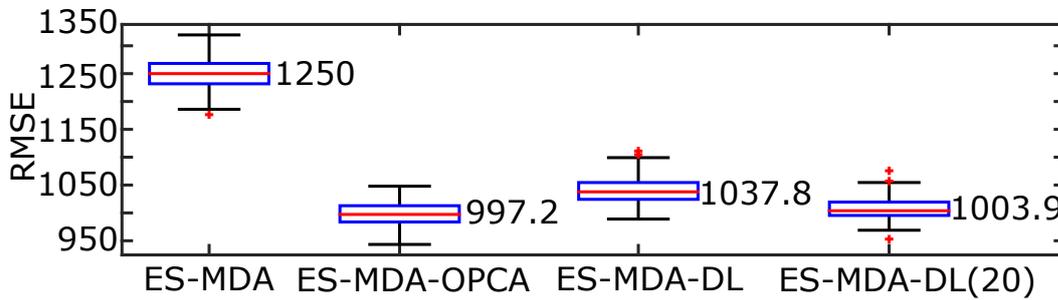


Figura 93: Raiz quadrada do erro médio quadrático da permeabilidade dos métodos estudados, para o caso de estudo BRUGGE.

Todos os resultados obtidos pela metodologia ES-MDA-DL nos casos de estudo apresentados, mostraram realizações *a posteriori* que preservam a geologia inicial do conjunto *a priori*. Embora foi percebido que o ajuste dos dados de produção observados em comparação com o método ES-MDA padrão e em alguns casos do método ES-MDA-OPCA, ficaram um uma distribuição com maior dispersão. Isto pode ser percebido melhor na Figura 89 onde na subfigura na linha superior direita, ilustra que os resultados de corte de água para o poço 11, na qual a curva de produção média (linha verde) ficou acima dos dados de produção observados (círculos vermelhos).

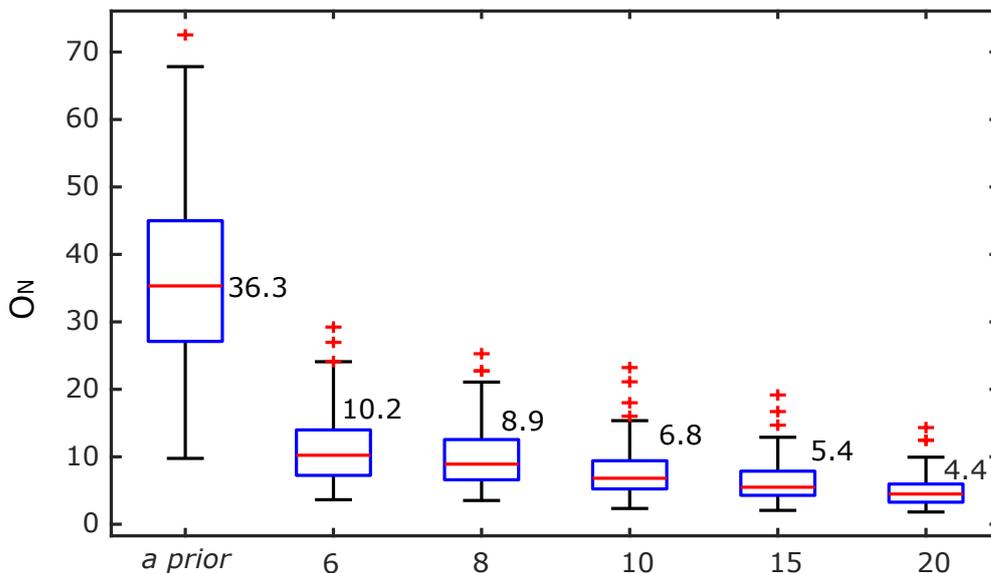


Figura 94: Influência do número de assimilações para o caso de estudo BRUGGE com o método ES-MDA-DL.

Para olhar o comportamento do método ES-MDA-DL quando é incrementado o número de iterações, foi realizado um último experimento variando

o número de assimilações (N_a) para o caso de estudo BRUGGE. Os resultados obtidos mostram que incrementando esse parâmetro obtém-se um melhor ajuste dos dados observados. Baseado na função objetivo (Equação 6-5) apresentada no Capítulo 3. Na Figura 94 é apresentado o gráfico de *boxplot* para diferentes valores de N_a . Note-se que à medida que é incrementado o valor de N_a , a dispersão e mediana dos valores da função objetivo é reduzida.

Figura 94 indicou que, baseado na função objetivo (Equação 6-5), o melhor resultado foi obtido para um número de 20 assimilações, este valor é próximo ao valor obtido pelo método ES-MDA-OPCA, que foi de $O_N = 4.5$.

Assim, nas Figuras 95 e 96 são apresentadas os resultados do método ES-MDA-DL para um número de 20 assimilações. Note-se que os mapas de logaritmo natural da permeabilidade obtidos não apresentam muita diferença com os resultados obtidos para um número de $N_a = 6$ assimilações. Embora as curvas de produção finais (curvas azuis) obtidas pelo método ES-MDA-DL com 20 iterações apresentam uma menor dispersão.

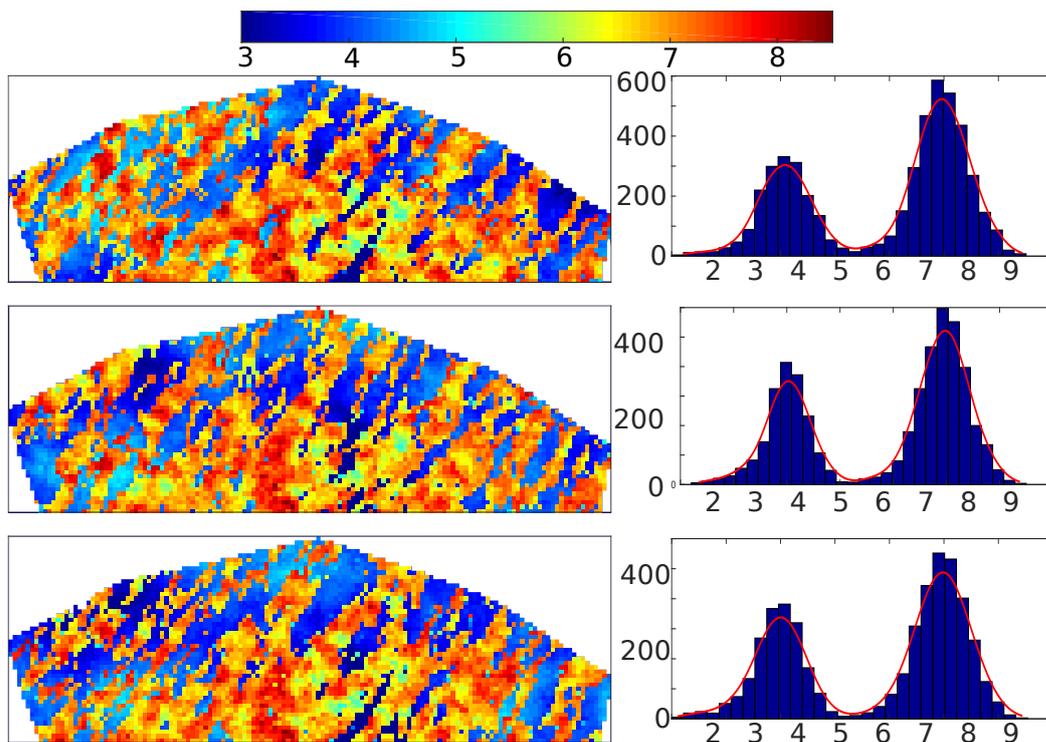
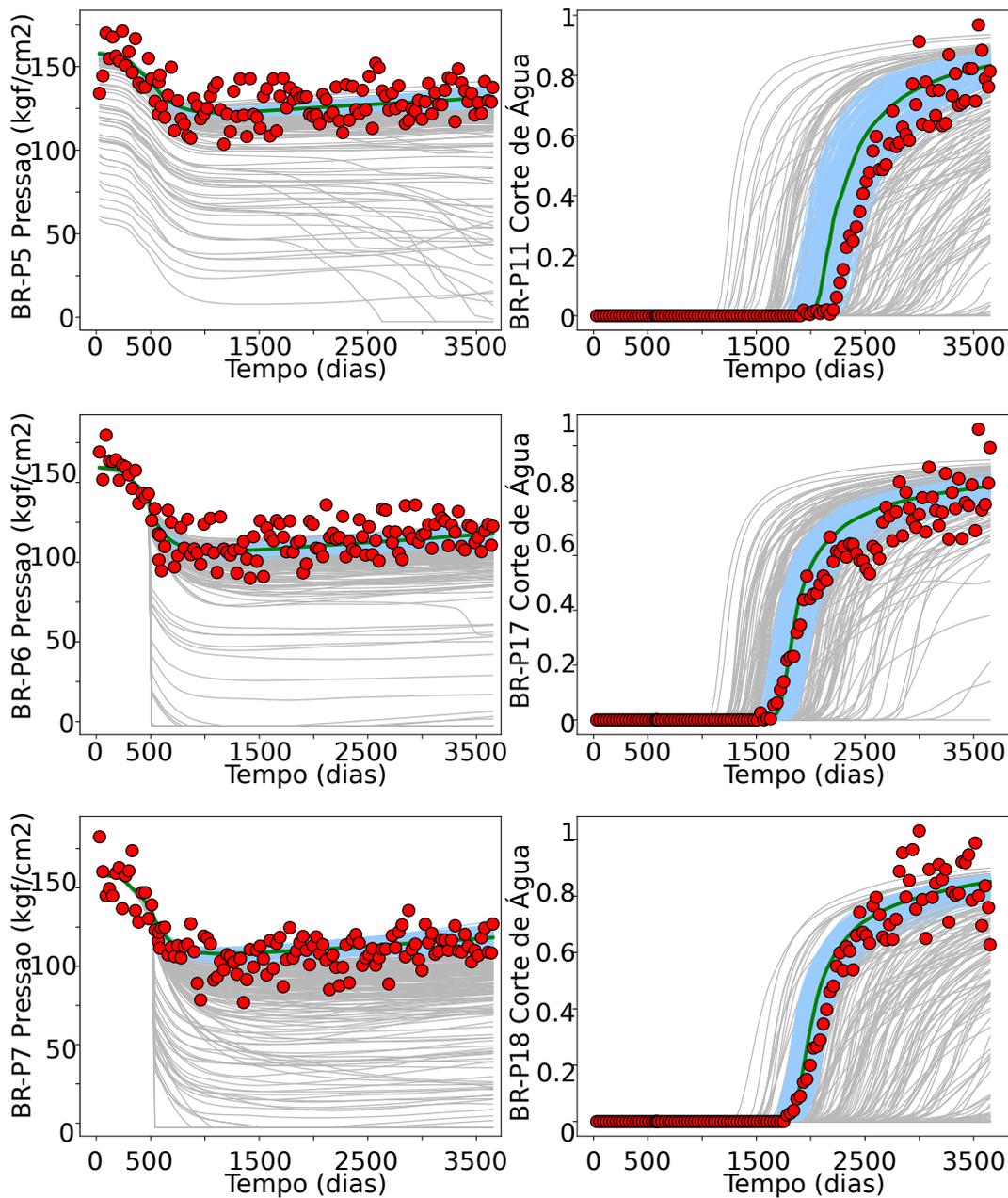


Figura 95: Primeiras três realizações do logaritmo natural da permeabilidade do conjunto final (*a posteriori*) do caso de estudo BRUGGE aplicando o método ES-MDA-DL, com um número de 20 assimilações.

Comparando ES-MDA-OPCA (Figura 91), nota-se que as curvas de produção finais em ambos ficaram muito semelhantes, com pouca dispersão. Observe-se que os dados de produção observados (círculos vermelhos nas figuras) estão próximos ao valor médio dos resultados obtidos (curva verde na figura).



PUC-Rio - Certificação Digital N° 1313464/CA

Figura 96: Curvas de produção resultantes do método ES-MDA-DL aplicado ao caso de estudo BRUGGE, com um número de 20 assimilações. Pressão nos poços 5, 6 e 7. Corte de água nos poços 11, 17 e 18.

Na linha superior direita na Figura 96, observa-se que a curva média está um pouco acima dos dados de produção observados, a diferença com a mesma gráfica em ES-MDA-OPCA, onde a curva média ficou dentro dos dados de produção observados. Aqui pode-se perceber porque o valor da função objetivo de ES-MDA-OPCA é menor do que ES-MDA-DL com 20 assimilações.

Os resultados apresentados nesse capítulo mostraram o bom desempenho

que fornece a metodologia proposta (ES-MDA-DL). Em todos os casos de estudo onde foi testada o método ES-MDA-DL consegui manter características importantes do reservatório, como por exemplo, as distribuições dos modelos de permeabilidade. Embora, os ajustes dos dados de produção observados tenham uma maior dispersão em comparação com o ajuste do método ES-MDA padrão (sem parametrização), isto pode ser contornado utilizando um número maior de assimilações (iterações) como foi apresentado no último experimento para o caso de estudo BRUGGE.

7

Conclusões e Recomendações

7.1

Conclusões

O problema de ajuste de histórico é bastante complexo, sobretudo quando é levada em consideração a importância de se preservar as características geológicas do reservatório em estudo. Assim, é de grande impacto tecnológico e industrial o desenvolvimento de métodos que permitam honrar os dados de produção e preservar a modelagem geoestatística do reservatório. O presente trabalho investigou o uso de métodos baseados no filtro de Kalman para a incorporação de dados dinâmicos (ajuste de histórico) em modelos de fácies geológicas. O trabalho introduziu ainda duas parametrizações para esse problema. A primeira baseada em análise de componentes principais e a segunda baseada em técnicas de *Deep Learning*. Essas parametrizações foram integradas ao método ES-MDA.

O método ES-MDA foi escolhido pelo fato de que a pesquisa realizada demonstrou a robustez dele em relação aos outros métodos avaliados baseados no filtro de Kalman. Esse método tem uma implementação prática, consegue obter resultados satisfatórios, possui capacidade de lidar com sistemas de alta dimensão e permite usar o simulador de reservatórios como uma “caixa preta” já que não há necessidade do cálculo de gradientes. Além de estimar e prever a produção, permite interpretar a incerteza na descrição do reservatório, assim conhecendo o desempenho das previsões. Já que o método parte de uma abordagem Bayesiana usando uma definição formal de probabilidade, pode-se interpretar a incerteza na descrição do reservatório. Entretanto, os resultados com a aplicação direta do método ES-MDA em modelos de reservatórios com distribuições complexas de fácies, mostraram como os modelos ajustados perderam suas características iniciais.

DL é uma área promissora, que cresce diariamente com diversas aplicações práticas. A parametrização baseada em DL desenvolvida nesta tese é inédita na literatura. Os resultados obtidos mostram que essa técnica tem grande potencial de uso em problemas complexos de ajuste de histórico. Os resultados obtidos em todos os casos de estudo mostram que as arquiteturas baseadas no

DL conseguem representar corretamente as imagens correspondentes a modelos de fácies geológicas.

Uma das conclusões do trabalho é que quando as imagens de fácies apresentam características simples (e.g. caso de estudo PUNQ-S3), o treinamento mediante uma rede AE ou RBM são suficientes para parametrizar corretamente. Entretanto, à medida que os modelos de fácies contenham características mais complexas é preciso utilizar arquiteturas com mais camadas. A arquitetura DBN com três camadas ocultas forneceram resultados bem sucedidos para os casos investigados nesta tese.

Os métodos desenvolvidos foram testados para diferentes casos de estudo com o objetivo de testar e validar diferentes métodos e procedimentos propostos. As principais conclusões são resumidas a seguir:

- O uso de técnicas baseadas em *importance sampling* melhoram o desempenho da assimilação de dados apenas para o caso de estudo “1D”. Isso se deve ao fato de que esse caso tem uma dimensão pequena (apenas 31 células). Contudo, para modelos de reservatório com maiores dimensões, o efeito dos pesos é bastante reduzido ao ponto de que esses métodos não tiveram uma melhoria nos resultados quando comparados no ES-MDA.
- As parametrizações desenvolvidas, baseadas em DL e na análise de componentes principais, podem lidar com hipóteses de não gaussianidade, que são assumidas pelo método ES-MDA. Assim, o método torna-se mais robusto no processo de ajuste de histórico.
- O método de parametrização desenvolvido com base na análise de componentes principais obteve bons resultados para casos mais simples, mas ainda mostra algumas deficiências. Dentre elas está sua tediosa implementação, já que há necessidade da integração de várias ferramentas. Isto pode se tornar mais complicado em reservatórios de alta dimensão e com maior número de fácies.
- A parametrização baseada em DL apresentou resultados mais robustos, em comparação com a baseada na análise de componentes principais. A implementação prática e robusta dos algoritmos de DL fornece uma grande facilidade para ser integrada ao método ES-MDA.
- As arquiteturas com três camadas ocultas de DL conseguiram resultados bem sucedidos. No caso de estudo mais simples (PUNQ-S3) uma arquitetura de só uma camada fornece bons resultados. A medida que a complexidade das imagens de fácies aumenta é preciso incrementar o número de camadas ocultas nas arquiteturas DL para seu treinamento.

- A metodologia ES-MDA-DL proposta apresentou resultados satisfatórios tanto no ajuste dos modelos aos dados observados quanto na conservação da geologia inicial dos reservatórios em estudo. Os mapas de permeabilidade apresentados ao final do ajuste de histórico para cada caso de estudo apresentam características geológicas e geoestatísticas semelhantes do conjunto inicial.
- Os resultados indicam que é necessário um maior número de iterações de ES-MDA quando é usado em conjunto com DL para se obter um ajuste satisfatório dos dados de produção. Possivelmente esse fato é explicado pelo aumento da não-linearidade do problema.

7.2

Recomendações para Trabalhos Futuros

O objetivo final da tese foi apresentar uma metodologia robusta que seja capaz de realizar um ajuste dos dados de produção, mas sem perder as características estabelecidas pelo geólogo.

Entre possíveis trabalhos futuros cabe citar:

- Investigação do método ES-MDA-DL em modelos mais complexos, com um número maior de fácies.
- Investigação de desempenho do método em um caso real com grandes dimensões em particular no desempenho computacional do treinamento das redes de DL.
- Investigação sobre o uso de outras arquiteturas e algoritmos de DL.
- Investigação sobre o uso de técnicas de geoestatística para reprocessamento das imagens geradas pelo ES-MDA-DL. O objetivo é preservar ao máximo a modelagem feita pelo geólogo.

Referências bibliográficas

- 1 JACQUARD, P.; JAIN, C. **Permeability distribution from field pressure data.** SPE Journal, 5(4):281–294, 1965.
- 2 JAHNS, H. O. **A rapid method for obtaining a two-dimensional reservoir description from well pressure response data.** SPE Journal, 6(12):315–327, 1966.
- 3 KALMAN, R. E. **A new approach to linear filtering and prediction problems.** Transactions of the ASME, Journal of Basic Engineering, 82:35–45, 1960.
- 4 EVENSEN, G. **Data Assimilation: The Ensemble Kalman Filter.** Springer Science & Business Media, Berlin, 2009.
- 5 AANONSEN, S. I.; NÆVDAL, G.; OLIVER, D. S.; REYNOLDS, A. C. ; VALLÈS, B. **Review of ensemble Kalman filter in petroleum engineering.** SPE Journal, 14(3):393–412, 2009.
- 6 OLIVER, D. S.; CHEN, Y. **Recent progress on reservoir history matching: A review.** Computational Geosciences, 15(1):185–221, 2011.
- 7 EVENSEN, G. **Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo Methods to forecast error statistics.** Journal of Geophysical Research, 99(C5):10143–10162, 1994.
- 8 HOUTEKAMER, P. L.; MITCHELL, H. L.; PELLERIN, G.; BUEHNER, M.; CHARRON, M.; SPACEK, L. ; HANSEN, B. **Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations.** Monthly Weather Review, 133(3):604–620, 2005.
- 9 REICHLE, R. H.; MCLAUGHLIN, D. B. ; ENTEKHABI, D. **Hydrologic data assimilation with the ensemble Kalman filter.** Monthly Weather Review, 130(1):103–114, 2002.
- 10 BERTINO, L.; EVENSEN, G. ; WACKERNAGEL, H. **Sequential data assimilation techniques in oceanography.** International Statistical Review, 71(2):223–241, 2003.
- 11 VAN LEEUWEN, P. J.; EVENSEN, G. **Data assimilation and inverse methods in terms of a probabilistic formulation.** Monthly Weather Review, 124:2898–2913, 1996.

- 12 SKJERVHEIM, J.-A.; EVENSEN, G.; HOVE, J. ; VABØ, J. G. **An ensemble smoother for assisted history matching**. In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, THE WOODLANDS, TEXAS, USA, 21–23 FEBRUARY. SPE 141929, 2011.
- 13 THULIN, K.; LI, G.; AANONSEN, S. I. ; REYNOLDS, A. C. **Estimation of initial fluid contacts by assimilation of production data with EnKF**. In: PROCEEDINGS OF THE SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION, ANAHEIM, CALIFORNIA, 11–14 NOVEMBER. SPE 109975, 2007.
- 14 CHEN, Y.; OLIVER, D. S. **Ensemble randomized maximum likelihood method as an iterative ensemble smoother**. *Mathematical Geosciences*, 44(1):1–26, 2012.
- 15 CHEN, Y.; OLIVER, D. S. **Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification**. *Computational Geosciences*, 17:689–703, 2013.
- 16 EMERICK, A. A.; REYNOLDS, A. C. **Ensemble smoother with multiple data assimilation**. *Computers & Geosciences*, 55:3–15, 2013.
- 17 STORDAL, A. S.; ELSHEIKH, A. H. **Iterative ensemble smoothers in the annealed importance sampling framework**. *Advances in Water Resources*, 86:231–239, 2015.
- 18 STORDAL, A. S. **Iterative Bayesian inversion with Gaussian mixtures: finite sample implementation and large sample asymptotics**. *Computational Geosciences*, 19(1):1–15, 2014.
- 19 RESOPTIMA. **ResX**: <<http://resoptima.com/solutions/>>. Access, 2017.
- 20 JOLLIFFE, I. **Principal Component Analysis**. Wiley Online Library, 2002.
- 21 EMERICK, A. A. **Investigation on principal component analysis parameterizations for history matching channelized facies models with ensemble-based data assimilation**. *Mathematical Geosciences*, 49(1):85–120, 2017.
- 22 BENGIO, Y. **Learning Deep Architectures for AI**. Now Publishers Inc, 2009.
- 23 LECUN, Y.; BENGIO, Y. ; HINTON, G. **Deep learning**. *Nature*, 521:436–444, 2015.
- 24 DENG, L.; LI, J.; HUANG, J.-T.; YAO, K.; YU, D.; SEIDE, F.; SELTZER, M.; ZWEIG, G.; HE, X. ; WILLIAMS, J. **Recent advances in deep learning for speech research at Microsoft**. In: ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), p. 8604–8608. IEEE, 2013.

- 25 KRIZHEVSKY, A.; SUTSKEVER, I. ; HINTON, G. **Imagenet classification with deep convolutional neural networks**. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, p. 1097–1105, 2012.
- 26 CHO, K.; VAN, B.; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H. ; BENGIO, Y. **Learning phrase representations using RNN encoder-decoder for statistical machine translation**. arXiv preprint arXiv:1406.1078, 2014.
- 27 HAU CHEN, C. **Handbook of pattern recognition and computer vision**. World Scientific, 2015.
- 28 HIRSCHBERG, J.; MANNING, C. **Advances in natural language processing**. Science, 349(6245):261–266, 2015.
- 29 DENG, L.; YU, D. **Deep learning: Methods and applications**. Foundations and Trends in Signal Processing, 7(3–4):197–387, 2014.
- 30 THIKSHAJA, U. K.; PAUL, A. **A brief review on deep learning and types of implementation for deep learning**. Deep Learning Innovations and Their Convergence With Big Data, p. 20, 2017.
- 31 ROSA, A. J.; DE SOUZA CARVALHO, R. ; XAVIER, J. A. D. **Engenharia de Reservatórios de Petróleo**. Interciência, Rio de Janeiro, 2006.
- 32 CHIERICI, G. L. **Principles of Petroleum Reservoir Engineering**. Springer Science & Business Media, Berlin, 2012.
- 33 MACHADO, A. F. **Análise quantitativa de mapas de pressão e saturação no processo de ajuste de histórico**. Master's thesis, Unicamp, 2010.
- 34 RWECHUNGURA, R. W.; DADASHPOUR, M. ; KLEPPE, J. **Advanced history matching techniques reviewed**. In: SPE MIDDLE EAST OIL AND GAS SHOW AND CONFERENCE, MANAMA, BAHRAIN, 25-28 SEPTEMBER. SPE 142497, 2011.
- 35 LI, R.; REYNOLDS, A. C. ; OLIVER, D. S. **History matching of three-phase flow production data**. SPE Journal, 8(4):328–340, 2003.
- 36 GAO, G.; REYNOLDS, A. C. **An improved implementation of the LBFGS algorithm for automatic history matching**. SPE Journal, 11(1):5–17, 2006.
- 37 TAVAKOLI, R.; REYNOLDS, A. C. **History matching with parameterization based on the SVD of a dimensionless sensitivity matrix**. SPE Journal, 15(12):495–508, 2010.
- 38 SHIRANGI, M. G.; EMERICK, A. A. **An improved TSVD-based Levenberg-Marquardt algorithm for history matching and comparison with Gauss-Newton**. Journal of Petroleum Science and Engineering, 143:259–271, 2016.

- 39 SCHULZE-RIEGERT, R.; AXMANN, J.; HAASE, O.; RIAN, D. ; YOU, Y.-L. **Evolutionary algorithms applied to history matching of complex reservoirs**. SPE Reservoir Evaluation & Engineering, 5(02):163–173, 2002.
- 40 MOHAMED, L.; CHRISTIE, M. A. ; DEMYANOV, V. **Reservoir model history matching with particle swarms: Variants study**. In: SPE OIL AND GAS INDIA CONFERENCE AND EXHIBITION. Society of Petroleum Engineers, 2010.
- 41 OUENES, A.; BRÉFORT, B.; MEUNIER, G. ; DUPÉRE. **A new algorithm for automatic history matching: Application of simulated annealing method (SAM) to reservoir inverse modeling**. SPE Unsolicited Manuscript, 1(SPE 26297), 1993.
- 42 KABIR, C.; YOUNG, N. **Handling production-data uncertainty in history matching: the meren reservoir case study**. SPE Reservoir Evaluation & Engineering, 7(02):123–131, 2004.
- 43 ALESSIO, L. D.; BOURDON, L. M.; COCA, S. ; OTHERS. **Experimental design as a framework for multiple realisation history matching: F6 further development studies**. In: SPE ASIA PACIFIC OIL AND GAS CONFERENCE AND EXHIBITION. Society of Petroleum Engineers, 2005.
- 44 PEAKE, W.; ABADAH, M.; SKANDER, L. ; OTHERS. **Uncertainty assessment using experimental design: Minagish oolite reservoir**. In: SPE RESERVOIR SIMULATION SYMPOSIUM. Society of Petroleum Engineers, 2005.
- 45 KING, G.; LEE, S.; ALEXANDRE, P.; MIGUEL, M.; BLEVENS, M.; PILLOW, M.; CHRISTIE, G. ; OTHERS. **Probabilistic forecasting for mature fields with significant production history: a nemba field case study**. In: SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION. Society of Petroleum Engineers, 2005.
- 46 JAZWINSKI, A. H. **Stochastic Processes and Filtering Theory**. Courier Corporation, New York, 2007.
- 47 PSIAKI, M. L. **The blind tricyclist problem and a comparative study of nonlinear filters: A challenging benchmark for evaluating nonlinear estimation methods**. IEEE Control Systems, 33(3):40–54, 2013.
- 48 JULIER, S.; UHLMANN, J. **Unscented filtering and nonlinear estimation**. Proceedings of the IEEE, 92(3):401–422, 2004.
- 49 ARULAMPALAM, S.; RISTIC, B. **Comparison of the particle filter with range-parameterized and modified polar EKFs for angle-only tracking**. In: SIGNAL AND DATA PROCESSING OF SMALL TARGETS, ORLANDO, FL, 24 APRIL. SPIE 12.391985, 2000.
- 50 OLIVER, D. S.; REYNOLDS, A. C. ; LIU, N. **Inverse Theory for Petroleum Reservoir Characterization and History Matching**. Cambridge University Press, Cambridge, UK, 2008.

- 51 EVENSEN, G. **The ensemble Kalman filter: Theoretical formulation and practical implementation.** *Ocean Dynamics*, 53(4):343–367, 2003.
- 52 LIU, N.; OLIVER, D. S. **Critical evaluation of the ensemble Kalman filter on history matching of geologic facies.** In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, THE WOODLANDS, TEXAS, 31 JANUARY-2 FEBURARY. SPE 92867, 2005.
- 53 EMERICK, A. A.; REYNOLDS, A. C. **Investigation of the sampling performance of Ensemble-based methods with a simple reservoir model.** *Computational Geosciences*, 17(2):325–350, 2013.
- 54 EMERICK, A. A. **Analysis of the performance of ensemble-based assimilation of production and seismic data.** *Journal of Petroleum Science and Engineering*, 139:219–239, 2016.
- 55 EMERICK, A. A.; REYNOLDS, A. C. **History matching of production and seismic data for a real field case using the ensemble smoother with multiple data assimilation.** In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, THE WOODLANDS, TEXAS, USA, 18–20 FEBRUARY, 2013.
- 56 LE, D. H.; EMERICK, A. A. ; REYNOLDS, A. C. **An adaptive ensemble smoother with multiple data assimilation for assisted history matching.** In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, HOUSTON, TEXAS, USA, 23-25 FEBRUARY. SPE 173214, 2015.
- 57 STORDAL, A. S.; KARLSEN, H. A.; NÆVDAL, G.; SKAUG, H. J. ; VALLÈS, B. **Bridging the ensemble Kalman filter and particle filters: The adaptive gaussian mixture filter.** *Computational Geosciences*, 15(2):293–305, 2011.
- 58 HOUTEKAMER, P. L.; MITCHELL, H. L. **Data assimilation using an ensemble Kalman filter technique.** *Monthly Weather Review*, 126(3):796–811, 1998.
- 59 CHEN, Y.; OLIVER, D. S. **Cross-covariance and localization for EnKF in multiphase flow data assimilation.** *Computational Geosciences*, 14(4):579–601, 2010.
- 60 EMERICK, A. A.; REYNOLDS, A. C. **Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications.** *Computational Geosciences*, 15(2):251–269, 2011.
- 61 GASPARI, G.; COHN, S. E. **Construction of correlation functions in two and three dimensions.** *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757, 1999.

- 62 LORENTZEN, R. J.; FJELDE, K. K.; FRØYEN, J.; LAGE, A. C.; NÆVDAL, G. ; VEFRING, E. H. **Underbalanced and Low-head drilling operations: Real time interpretation of measured data and operational support.** In: PROCEEDINGS OF THE SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION, NEW ORLEANS, LOUISIANA, 30 SEPTEMBER-3 OCTOBER. SPE 71384, 2001.
- 62 NÆVDAL, G.; MERETHE, J. L.; AANONSEN, S. I. ; VEFRING, E. H. **Reservoir monitoring and continuous model updating using ensemble Kalman filter.** In: SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION. Society of Petroleum Engineers, 2003.
- 64 GU, Y.; OLIVER, D. S. **History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter.** SPE Journal, 10(02):217–224, 2005.
- 65 PUNQ-S3. **Test case:** <<http://www.imperial.ac.uk/earth-science/research/research-groups/perm/standard-models/>>. Imperial College, 2017.
- 66 FLORIS, F.; BUSH, M.; CUYPERS, M.; ROGGERO, F. ; SYVERSVEEN, A. R. **Methods for quantifying the uncertainty of production forecasts: A comparative study.** Petroleum Geoscience, 7(S):S87–S96, 2001.
- 67 BIANCO, A.; COMINELLI, A.; DOVERA, L.; NÆVDAL, G. ; VALLES, B. **History matching and production forecast uncertainty by means of the ensemble Kalman filter: A real field application.** In: EUROPEC/EAGE CONFERENCE AND EXHIBITION. Society of Petroleum Engineers, 2007.
- 68 SKJERVHEIM, J.-A.; EVENSEN, G.; AANONSEN, S. I.; RUUD, B. O. ; JOHANSEN, T.-A. **Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter.** SPE Journal, 12(03):282–292, 2007.
- 69 HAVERL, M.; AGA, M. ; REISO, E. **Integrated workflow for quantitative use of time-lapse seismic data in history matching—a north sea field case.** In: 67TH EAGE CONFERENCE & EXHIBITION, 2005.
- 70 CHEN, Y.; OLIVER, D. S. **History matching of the norne full-field model with an iterative ensemble smoother.** SPE Reservoir Evaluation & Engineering, 17(02):244–256, 2014.
- 71 ARMSTRONG, M.; GALLI, A.; BEUCHER, H.; LOC'H, G.; RENARD, D.; DOLIGEZ, B.; ESCHARD, R. ; GEFFROY, F. **Plurigaussian Simulations in Geosciences.** Springer Science & Business Media, Berlin, 2011.
- 72 LOC'H, G.; GALLI, A. **Truncated plurigaussian method: Theoretical and practical points of view.** Geostatistics Wollongong, 96(1):211–222, 1997.

- 73 LÓPEZ, S. D. **Mapeamento da incerteza de realizações condicionais da simulação plurigaussiana truncada**. PhD thesis, Universidade de São Paulo, 2015.
- 74 ZHAO, Y.; REYNOLDS, A. C. ; LI, G. **Generating facies maps by assimilating production data and seismic data with the ensemble Kalman filter**. In: PROCEEDINGS OF THE SPE SYMPOSIUM ON IMPROVED OIL RECOVERY, TULSA, OKLAHOMA, USA, 20-23 APRIL. SPE 113990, 2008.
- 75 AGBALAKA, C. C.; OLIVER, D. S. **Application of the EnKF and localization to automatic history matching of facies distribution and production data**. *Mathematical Geosciences*, 40(4):353–374, 2008.
- 76 AGBALAKA, C. C.; OLIVER, D. S. **Automatic history matching of production and facies data with nonstationary proportions using EnKF**. In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, THE WOODLANDS, TEXAS, 2-4 FEBRUARY. SPE 118916, 2009.
- 77 SEBACHER, B.; HANEA, R. ; HEEMINK, A. **A probabilistic parametrization for geological uncertainty estimation using the ensemble Kalman filter (EnKF)**. *Computational Geosciences*, 17(5):813–832, 2013.
- 78 DEUTSCH, C. V.; WANG, L. **Hierarchical object-based stochastic modeling of fluvial reservoirs**. *Mathematical Geology*, 28(7):857–880, 1996.
- 79 GUARDIANO, F. B.; SRIVASTAVA, R. M. **Multivariate geostatistics: Beyond bivariate moments**. In: Soares, A., editor, *GEOSTATISTICS TROIA2 92*, volumen 5 de *Quantitative Geology and Geostatistics*, p. 133–144. Springer Netherlands, 1993.
- 80 STREBELLE, S. **Conditional simulation of complex geological structures using multiple-point statistics**. *Mathematical Geology*, 34(1):1–21, 2002.
- 81 MARIETHOZ, G.; CAERS, J. **Multiple-point Geostatistics – Stochastic Modeling with Training Images**. John Wiley & Sons, Ltd., 2014.
- 82 CHANG, H.; ZHANG, D. ; LU, Z. **History matching of facies distribution with the enkf and level set parameterization**. *Journal of Computational Physics*, 229(20):8011–8030, 2010.
- 83 MANNSETH, T. **Relation between level set and truncated pluri-Gaussian methodologies for facies representation**. *Mathematical Geosciences*, 46(6):711–731, 2014.

- 84 SARMA, P.; CHEN, W. H. **Generalization of the ensemble Kalman filter using Kernels for nongaussian random fields**. In: PROCEEDINGS OF THE SPE RESERVOIR SIMULATION SYMPOSIUM, THE WOODLANDS, TEXAS, 2-4 FEBRUARY. SPE 119177, 2009.
- 85 VO, H. X.; DURLOFSKY, L. J. **A new differentiable parameterization based on principal component analysis for the low-dimensional representation of complex geological models**. *Mathematical Geosciences*, 46(7), 2014.
- 86 VO, H. X.; DURLOFSKY, L. J. **Data assimilation and uncertainty assessment for complex geological models using a new PCA-based parameterization**. *Computational Geosciences*, 19(4):747–767, 2015.
- 87 CHEN, C.; GAO, G.; GELDERBLUM, P. ; JIMENEZ, E. **Integration of Cumulative-Distribution-Function mapping with Principal-Component Analysis for the history matching of channelized reservoirs**. *SPE Reservoir Evaluation & Engineering*, 19(2):278–293, 2016.
- 88 HE, J.; SARMA, P. ; DURLOFSKY, L. J. **Reduced-order flow modeling and geological parameterization for Ensemble-based data assimilation**. *Computers & Geosciences*, 55:54–69, 2013.
- 89 YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: Conceitos e Aplicações**. Oficina de Textos, São Paulo, Brasil, 2015.
- 90 LOC'H, G. L.; BEUCHER, H.; GALLI, A. ; DOLIGEZ, B. **Improvement in the truncated gaussian method: Combining several gaussian functions**. In: ECMOR IV-4TH EUROPEAN CONFERENCE ON THE MATHEMATICS OF OIL RECOVERY, 1994.
- 91 CHAMBERS, R. L.; YARUS, J. M. ; HIRD, K. B. **Petroleum geostatistics for nongeostatisticians: Part 1**. *The Leading Edge*, 19(5):474–479, 2000.
- 92 DOYEN, P. **Seismic reservoir characterization: An earth modeling perspective**, volumen 2. EAGE publications Houten, 2007.
- 93 STREBELLE, S. **Conditional simulation of complex geological structures using multiple-point statistics**. *Mathematical Geology*, 34(1):1–21, 2002.
- 94 PETERS, L.; ARTS, R.; BROUWER, G.; GEEL, C.; CULLICK, S.; LORENTZEN, R. J.; CHEN, Y.; DUNLOP, N.; VOSSEPOEL, F. C. ; XU, R. **Results of the Brugge benchmark study for flooding optimization and history matching**. *SPE Reservoir Evaluation & Engineering*, 13(03):391–405, 2010.
- 95 MOHAMED, L.; CHRISTIE, M. A.; DEMYANOV, V.; ROBERT, E. ; KACHUMA, D. **Application of particle swarms for history matching in the Brugge reservoir**. In: PROCEEDINGS OF THE SPE ANNUAL

- TECHNICAL CONFERENCE AND EXHIBITION, FLORENCE, ITALY, 19–22 SEPTEMBER. SPE 135264, 2010.
- 96 REMY, N.; BOUCHER, A. ; WU, J. **Applied Geostatistics with SGeMS: A User's Guide**. Cambridge University Press, Cambridge, UK, 2009.
- 97 HINTON, G. E.; OSINDERO, S. ; TEH, Y.-W. **A fast learning algorithm for deep belief nets**. *Neural Computation*, 18(7):1527–1554, 2006.
- 98 HINTON, G. E.; ZEMEL, R. S. **Autoencoders, minimum description length and Helmholtz free energy**. In: *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, p. 3–10, 1994.
- 99 HINTON, G. E.; SEJNOWSKI, T. J. **Learning and relearning in Boltzmann machines**. *Parallel Distributed Processing*, 1, 1986.
- 100 COLLOBERT, R.; BENGIO, S. **Links between perceptrons, MLPs and SVMs**. In: *PROCEEDINGS OF THE TWENTY-FIRST INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, p. 23. ACM, 2004.
- 101 COLLOBERT, R.; WESTON, J. **A unified architecture for natural language processing: Deep neural networks with multitask learning**. In: *PROCEEDINGS OF THE 25TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, p. 160–167. ACM, 2008.
- 102 HASTIE, T.; TIBSHIRANI, R. ; FRIEDMAN, J. **Overview of supervised learning**. In: *THE ELEMENTS OF STATISTICAL LEARNING*, p. 9–41. Springer, 2009.
- 103 COATES, A.; NG, A. ; LEE, H. **An analysis of single-layer networks in unsupervised feature learning**. In: *PROCEEDINGS OF THE FOURTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS*, p. 215–223, 2011.
- 104 REZENDE, D. J.; MOHAMED, S. ; WIERSTRA, D. **Stochastic backpropagation and approximate inference in deep generative models**. arXiv preprint arXiv:1401.4082, 2014.
- 105 GOODFELLOW, I.; BENGIO, Y. ; COURVILLE, A. **Deep Learning**. MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- 106 GOGNA, A.; MAJUMDAR, A. **Latent factor models for collaborative filtering**. PhD thesis, IIT-Delhi, 2017.
- 107 DENG, X.; TIAN, X.; CHEN, S. ; HARRIS, C. J. **Deep learning based nonlinear principal component analysis for industrial process fault detection**. In: *NEURAL NETWORKS (IJCNN), 2017 INTERNATIONAL JOINT CONFERENCE ON*, p. 1237–1243. IEEE, 2017.
- 108 VINCENT, P.; LAROCHELLE, H.; BENGIO, Y. ; MANZAGOL, P.-A. **Extracting and composing robust features with denoising autoencoders**. In: *PROCEEDINGS OF THE 25TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, p. 1096–1103. ACM, 2008.

- 109 ALAIN, G.; BENGIO, Y. **What regularized auto-encoders learn from the data-generating distribution.** The Journal of Machine Learning Research, 15(1):3563–3593, 2014.
- 110 BENGIO, Y. **Deep learning of representations: Looking forward.** In: INTERNATIONAL CONFERENCE ON STATISTICAL LANGUAGE AND SPEECH PROCESSING, p. 1–37. Springer, 2013.
- 111 LAN BOUREAU, Y.; LECUN, Y. **Sparse feature learning for deep belief networks.** In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, p. 1185–1192, 2008.
- 112 CONNOR, P.; HOLLENSEN, P.; KRIGOLSON, O. ; TRAPPENBERG, T. **A biological mechanism for bayesian feature selection: Weight decay and raising the LASSO.** Neural Networks, 67:121–130, 2015.
- 113 MAKHZANI, A.; FREY, B. **K-sparse autoencoders.** arXiv preprint arXiv:1312.5663, 2013.
- 114 RIFAI, S.; VINCENT, P.; MULLER, X.; GLOROT, X. ; BENGIO, Y. **Contractive auto-encoders: Explicit invariance during feature extraction.** In: PROCEEDINGS OF THE 28TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML-11), p. 833–840, 2011.
- 115 FAHLMAN, S. E.; HINTON, G. E. ; SEJNOWSKI, T. J. **Massively parallel architectures for AI: NETL, Thistle, and Boltzmann machines.** Proceedings of AAAI-83109, 113, 1983.
- 116 HINTON, G. E.; BROWN, A. D. **Spiking boltzmann machines.** In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, p. 122–128, 2000.
- 117 ACKLEY, D. H.; HINTON, G. E. ; SEJNOWSKI, T. J. **A learning algorithm for Boltzmann machines.** Cognitive Science, 9(1):147–169, 1985.
- 118 HINTON, G. E. **Training products of experts by minimizing contrastive divergence.** Neural Computation, 14(8):1771–1800, 2002.
- 119 BORIN, R. G. **Deteção de atividade vocal empregando máquinas de Boltzmann restritas.** PhD thesis, Universidade de São Paulo, 2016.
- 120 RABINER, L.; JUANG, B. **An introduction to hidden Markov models.** IEEE ASSP Magazine, 3(1):4–16, 1986.
- 121 DESJARDINS, G.; COURVILLE, A.; BENGIO, Y.; VINCENT, P. ; DELALLEAU, O. **Tempered markov chain Monte Carlo for training of restricted Boltzmann machines.** In: PROCEEDINGS OF THE THIRTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS, p. 145–152, 2010.
- 122 HOPFIELD, J. J. **Neural networks and physical systems with emergent collective computational abilities.** Proceedings of the National Academy of Sciences, 79(8):2554–2558, 1982.

- 123 HERNANDEZ, H. **Standard Maxwell-Boltzmann distribution: Definition and properties**. ForsChem Research Reports, 2, 2017.
- 124 BOTTOU, L. **Large-scale machine learning with stochastic gradient descent**. In: PROCEEDINGS OF COMPSTAT'2010, p. 177–186. Springer, 2010.
- 125 LAROCHELLE, H.; BENGIO, Y.; LOURADOUR, J. ; LAMBLIN, P. **Exploring strategies for training deep neural networks**. Journal of Machine Learning Research, 10(Jan):1–40, 2009.
- 126 BENGIO, Y.; COURVILLE, A. ; VINCENT, P. **Representation learning: A review and new perspectives**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8):1798–1828, 2013.
- 127 CHO, K.; ILIN, A. ; RAIKO, T. **Improved learning of Gaussian-Bernoulli restricted Boltzmann machines**. Artificial Neural Networks and Machine Learning–ICANN 2011, p. 10–17, 2011.
- 128 SUN, Q.; WU, C. ; LI LI, Y. **A new probabilistic neural network model based on backpropagation algorithm**. Journal of Intelligent & Fuzzy Systems, 32(1):215–227, 2017.
- 129 THIELE, J.; DIEHL, P. ; COOK, M. **A wake-sleep algorithm for recurrent, spiking neural networks**. arXiv preprint arXiv:1703.06290, 2017.
- 130 SALAKHUTDINOV, R. **Learning deep generative models**. University of Toronto, 2009.
- 131 KEYVANRAD, M. A.; HOMAYOUNPOUR, M. M. **A brief survey on deep belief networks and introducing a new object oriented toolbox (DeeBNet)**. arXiv preprint arXiv:1408.3264, 2014.
- 132 MARIETHOZ, G.; CAERS, J. **Multiple-point geostatistics: Stochastic modeling with training images**. John Wiley & Sons, 2014.
- 133 ARAUCO, S.; EMERICK, A. A. ; PACHECO, M. A. **Integration of ensemble data assimilation and deep learning for history matching facies models**. In: OTC BRASIL. Offshore Technology Conference, 2017.
- 134 CAERS, J.; ZHANG, T. **Multiple-point geostatistics: A quantitative vehicle for integrating geologic analogs into multiple reservoir models**. American Association of Petroleum Geologists, 2004.
- 135 EMERICK, A. A. **History Matching and Uncertainty Characterization Using Ensemble-based Methods**. PhD thesis, The University of Tulsa, 2012.