



**Vitor Hespanhol Côrtes**

**Aplicação de redes convolucionais  
autoencoder integradas ao método ES-MDA  
para ajuste de histórico assistido de modelos de  
reservatórios fraturados**

**MONOGRAFIA DE FINAL DE CURSO**

Monografia apresentada como requisito parcial para obtenção do título de Especialista em Ciência de Dados pelo programa de Pós-graduação "Business Intelligence Master" do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio.

Orientador: PhD Smith Arauco Canchumuni

Rio de Janeiro  
Dezembro de 2018

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

### **Vitor Hespanhol Côrtes**

Graduado em engenharia de telecomunicações pela Universidade Federal Fluminense (UFF) e especialista em engenharia de petróleo e gás natural pela Universidade Petrobras.

#### Ficha Catalográfica

Hespanhol Côrtes, Vitor

Aplicação de redes convolucionais autoencoder integradas ao método ES-MDA para ajuste de histórico assistido de modelos de reservatórios fraturados / Vitor Hespanhol Côrtes; orientador: Smith Arauco Canchumuni. – Rio de Janeiro: PUC-Rio, Departamento de Engenharia Elétrica, 2018.

v., 40 f: il. color. ; 30 cm

Monografia (especialização) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica .

Inclui bibliografia

1. Engenharia de Reservatórios;. 2. Redes Convolucionais;. 3. Autoencoder;. 4. Ajuste de Histórico;. 5. Filtro de Kalman;. 6. ES-MDA;. 7. Reservatórios Fraturados.. I. Canchumuni, Smith Arauco. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica . III. Título.

CDD: 620.11

## **Agradecimentos**

Agradeço a minha família pelo incentivo, apoio e compreensão ao longo deste período, sem os quais não seria possível a realização deste trabalho.

Aos colegas de turma e professores, gostaria também de agradecer por toda amizade, troca de experiências e orientação ao longo do curso.

## Resumo

Hespanhol Côrtes, Vitor; Canchumuni, Smith Arauco (Orientador). **Aplicação de redes convolucionais autoencoder integradas ao método ES-MDA para ajuste de histórico assistido de modelos de reservatórios fraturados**. Rio de Janeiro, 2018. 40p. Monografia – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

O desenvolvimento e gerenciamento da produção de um campo de petróleo depende largamente de ferramentas de simulação de reservatórios. Contudo, para se obter uma previsão de produção confiável de um modelo de simulação, é necessário que ele esteja ajustado aos dados históricos de produção do campo. Dentre os avanços recentes na área de ajuste de histórico, o método *ensemble smoother with multiple data assimilation* (ES-MDA) vem sendo utilizado com sucesso no ajuste de modelos de reservatórios. Contudo, este método assume que as variáveis do modelo possuam uma função de distribuição Gaussiana, tendo sua performance reduzida ao lidar com atributos discretos, como por exemplo a propriedade de ocorrência/ausência de corredores de fratura. Portanto, o presente trabalho tem por objetivo integrar uma rede convolucional *autoencoder* ao processo, sendo capaz de transformar as propriedades representativas dos corredores de fratura em uma propriedade com distribuição Gaussiana, que sirva de parâmetro de entrada para o ES-MDA, além de ser capaz de decodificar sua saída nas propriedades originais.

## Palavras-chave

Engenharia de Reservatórios; Redes Convolucionais; Autoencoder; Ajuste de Histórico; Filtro de Kalman; ES-MDA; Reservatórios Fraturados.

## Abstract

Hespanhol Côrtes, Vitor; Canchumuni, Smith Arauco (Advisor). **Aplication of convolutional autoencoders integrated to ES-MDA method for assisted history matching of fractured reservoir models**. Rio de Janeiro, 2018. 40p. Monograph – Departamento de Engenharia Elétrica , Pontifícia Universidade Católica do Rio de Janeiro.

The development and management of oil fields relies on reservoir simulation tools. However, in order to obtain a reliable forecast, the reservoir model response must be fitted to measured production data. Among recent advances in history matching area, the ensemble smoother with multiple data assimilation method (ES-MDA) has been successfully applied to adjust the reservoir model parameters. However, the method assumes that these parameters have a Gaussian distribution function, and its performance is reduced when dealing with discrete attributes, such as fracture occurrence/absence property. Therefore, the present work aims to integrate a convolutional autoencoder to the process, being able to codify the representative property of the fracture corridors into a property with Gaussian distribution, which will be used as input parameter for the ES-MDA method, and also being able to decode its output into the original property.

## Keywords

Reservoir Engineering; Convolutional Neural Network; Autoencoder; History Matching; Kalman Filter; ES-MDA; Fractured Reservoir.

## Sumário

1	Introdução	10
1.1	Descrição do Problema	11
1.2	Organização da Monografia	12
2	Ajuste de Histórico	13
2.1	Métodos Baseados em Filtro de Kalman	15
2.1.1	<i>Ensemble Smoother with Multiple Data Assimilation</i>	17
2.1.2	<i>Localization</i>	18
3	Descrição do modelo de reservatórios	20
3.1	Construção das realizações de fraturas	21
4	<i>Autoencoder</i>	23
4.1	Arquitetura da Rede	24
4.2	Treinamento e Validação	25
5	Resultados	27
5.1	Dados de produção	28
5.2	Mapas de ocorrência de fraturas	31
5.3	Mapas de saturação de água	32
6	Conclusão	34
	Referências bibliográficas	36
A	Algoritmo ES-MDA	39

## Lista de figuras

Figura 1.1	Método de ajuste de histórico proposto.	12
Figura 2.1	Processo de assimilação de dados pelo método EnKF (3).	16
Figura 2.2	Processo de assimilação de dados pelo método ES (3).	16
Figura 2.3	Processo de assimilação de dados pelo método ES-MDA (3).	16
Figura 3.1	Modelo de reservatórios SPE 9.	20
Figura 3.2	Mapa de permeabilidade de matriz (a) e posição dos poços no modelo (b).	21
Figura 3.3	Imagem de treinamento representando a propriedade de ocorrência (em vermelho) e ausência (em preto) de fraturas (26).	22
Figura 3.4	Exemplos de realizações de fraturas criadas.	22
Figura 4.1	Estrutura geral de uma rede autoencoder (27).	23
Figura 4.2	Representação da rede <i>autoencoder</i> proposta.	23
Figura 4.3	Treinamento da rede.	25
Figura 4.4	Atributos de entrada (a), codificado (c) e decodificado (b) pela rede <i>autoencoder</i> de três realizações de fraturas.	26
Figura 4.5	Histogramas do atributo original e codificado para três realizações de fraturas.	26
Figura 5.1	Comparação dos valores do erro quadrático médio (MSE) dos modelos <i>a priori</i> e <i>a posteriori</i> .	27
Figura 5.2	<i>Boxplot</i> do erro quadrático médio dos modelos a cada iteração.	28
Figura 5.3	Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 1.	29
Figura 5.4	Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 2.	29
Figura 5.5	Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 3.	30
Figura 5.6	Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 4.	30
Figura 5.7	Ajuste de histórico da pressão de fundo de 4 poços injetores.	31
Figura 5.8	Comparação entre 3 modelos de fraturas antes e após o ajuste em relação ao reservatório real ( $\mathbf{m}_{true}$ ).	32
Figura 5.9	Comparação entre mapas saturação de água do reservatório real ( $\mathbf{m}_{true}$ ) e dos modelos originais ( <i>a priori</i> ) após 3000 dias de produção.	33
Figura 5.10	Comparação entre mapas saturação de água do reservatório real ( $\mathbf{m}_{true}$ ) e dos modelos ajustados ( <i>a posteriori</i> ) após 3000 dias de produção.	33

## Lista de tabelas

Tabela 4.1	Descrição das camadas da rede <i>autoencoder</i> .	24
------------	--	----



## Lista de Abreviaturas

$\alpha$  – coeficiente de inflação

$\varepsilon$  – ruído aleatório

$\mathbf{C}_D$  – matriz de covariância do ruído nos dados medidos

$\mathbf{C}_{MS}$  – matriz de covariância cruzada entre os parâmetros do modelo e os dados simulados

$\mathbf{C}_S$  – matriz de covariância dos dados simulados

$\mathbf{d}$  – dados simulados a partir de um modelo de simulação  $\mathbf{m}$

$\mathbf{d}_{hist}$  – dados históricos representativo do reservatório real

E&P – Exploração e Produção de Petróleo

EnKF – *Ensemble Kalman Filter*

ES – *Ensemble Smother*

ES-MDA – *Ensemble Smoother with Multiple Data Assimilation*

GPU – *Graphics Processing Unit*

$\mathbf{K}$  – ganho de Kalman

$\mathbf{m}$  – modelo de reservatórios

$\mathbf{m}_{prior}$  – modelo *a priori*

$\mathbf{m}_{post}$  – modelo *a posteriori*

$\mathbf{m}_{true}$  – modelo representativo do reservatório real

MPS – *Multiple-Point Simulation*

MSE – erro quadrático médio

$N_a$  – número de iterações para assimilação de dados

$N_d$  – número total de dados históricos

$N_m$  – número total de modelos

$N_p$  – número total de parâmetros em cada modelo

PSO – *Particle Swarm Optimization*

$\mathbf{R}_{MS}$  – matriz de regularização do ganho de Kalman (*localization*)

SGeMS – *Stanford Geostatistical Modeling Software*

SNESIM – *Single Normal Equation Simulation*

SPE – *Society of Petroleum Engineers*

# 1

## Introdução

A simulação de reservatórios representa um papel fundamental para o desenvolvimento de campos de petróleo, permitindo avaliar a recuperação final da jazida ao se adotar diferentes estratégias de exploração e proporcionando uma maior precisão na previsão do comportamento de produção do reservatório.

A construção de um modelo de simulação é um processo que requer um amplo conhecimento das propriedades e dos processos geológicos e petrofísicos relacionados, sendo modelados de forma discretizada e representando suficientemente a complexidade do reservatório, servindo de arcabouço para a solução de equações fundamentais de fluxo que determinarão as repostas do modelo (pressões, vazões e saturações) aos dados de entrada e às restrições impostas (condições operacionais).

Inicialmente, a caracterização e a modelagem de um reservatório dependem da inferência e extrapolação de suas propriedades estáticas, como porosidade e permeabilidade, a partir de dados escassos de poços (comparados às dimensões do reservatório), dados sísmicos, além de todo conhecimento conceitual disponível sobre a geologia do reservatório (1), atribuindo portanto um alto grau de incerteza ao processo e resultando em uma grande quantidade de modelos representativos possíveis.

Contudo, ao longo da produção do campo, um conjunto de dados dinâmicos são observados nos poços, como pressões e vazões de fluidos (óleo, água e gás), que podem incorporar informações relevantes à modelagem do reservatório. Este processo de assimilação de dados é uma etapa de grande importância, uma vez que a preditividade de um modelo de simulação está relacionada a sua capacidade de reproduzir os dados históricos de produção.

Assim, o termo ajuste de histórico compreende o processo de atualizar um modelo de reservatório de modo que a simulação numérica da produção esteja consistente com os dados de produção medidos (1).

Portanto, a redução de incertezas na previsão da produção de um campo de petróleo através da adoção de estratégias eficientes de ajuste de histórico assistido são altamente desejáveis, de forma a permitir uma maior qualidade na tomada de decisão e na concepção de novos projetos, além de uma maior previsibilidade do desempenho de empresas de E&P.

## 1.1

### Descrição do Problema

Um grande número de aplicações para ajuste de histórico assistido tem utilizado métodos baseados no filtro de Kalman. Dentre elas, o método *ensemble smoother with multiple data assimilation* (ES-MDA), proposto por EMERICK e REYNOLDS (2), vem demonstrando excelentes resultados em termos da qualidade do ajuste como na quantificação de incertezas, além de se destacar por sua eficiência e simplicidade (3, 4).

Contudo, este método se apoia em hipóteses de linearidade e Gaussianidade dos dados de entrada (3), tendo sua performance reduzida à medida que estas variáveis se afastam de um comportamento Gaussiano. Assim, apesar dos bons resultados no ajuste de histórico ao considerar propriedades de matriz (porosidade, permeabilidade etc) como variáveis de entrada (2), muitas vezes sua utilização é inviabilizada quando é necessário ajustar propriedades que possuem valores categóricos, como por exemplo ao considerar o ajuste de fácies geológicas ou da propriedade de ocorrência/ausência de corredores de fratura, que apresentam uma função de distribuição de probabilidade discreta.

Apesar de na literatura serem reportadas algumas estratégias que se propõe a tratar propriedades discretas em conjunto a métodos baseados no filtro de Kalman (5, 6), na prática, quando há a necessidade de considerar este tipo de atributo, recorre-se à métodos estocásticos de ajuste de histórico assistido - como algoritmos genéticos (7, 8, 9) - considerando um número limitado de cenários ou mesmo à realização de um ajuste manual do modelo (10).

Portanto, o presente trabalho tem por objetivo propor uma aplicação baseada em *deep learning* que permita incorporar a propriedade de ocorrência/ausência de corredores de fratura ao processo de ajuste de histórico utilizando o método ES-MDA, análogo ao proposto por CANCHUMUNI *et al.* (4) para o tratamento de fácies. Para isso, será construída uma rede convolucional *autoencoder* que seja capaz de converter a propriedade original numa propriedade codificada com características Gaussianas, para que seja utilizada como variável de entrada para o método ES-MDA e, após cada ciclo de assimilação de dados, permita decodificar a saída para a propriedade que será utilizada na simulação do modelo de reservatórios (figura 1.1).

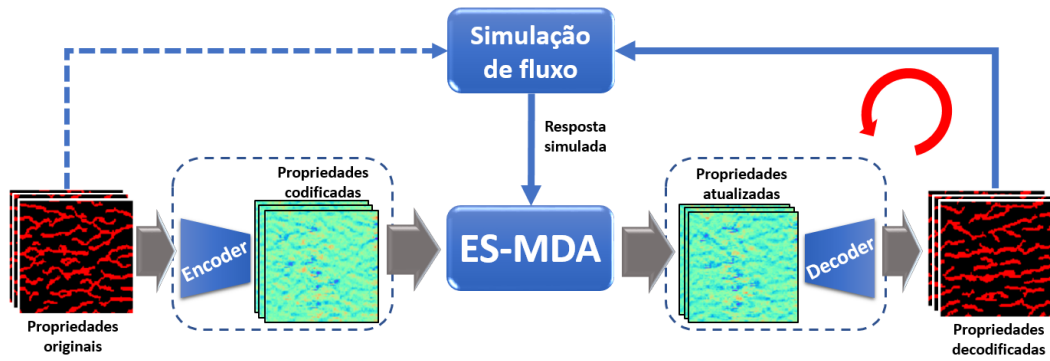


Figura 1.1: Método de ajuste de histórico proposto.

## 1.2

### Organização da Monografia

Para alcançar os objetivos propostos, a apresentação dos assuntos abordados nesta monografia está organizada da seguinte forma:

- ▷ uma breve discussão sobre metodologias de ajuste de histórico assistido e a descrição do método ES-MDA (Capítulo 2);
- ▷ apresentação do modelo de simulação utilizado e descrição da metodologia para construção das realizações de corredores de fraturas considerados no processo de ajuste de histórico (Capítulo 3);
- ▷ descrição da rede convolucional *autoencoder* proposta, além da apresentação dos resultados do treinamento da rede (Capítulo 4);
- ▷ demonstração e discussão dos resultados do processo de ajuste de histórico (Capítulo 5); e
- ▷ conclusões e propostas para futuros trabalhos (Capítulo 6).

## 2

### Ajuste de Histórico

O problema de ajuste de histórico pode ser classificado como um problema inverso uma vez que seu objetivo é o de determinar propriedades fisicamente plausíveis de um sistema (reservatório), ou obter informações sobre estas propriedades, dadas as respostas observadas do sistema (dados históricos de produção) a determinados estímulos (condições operacionais) (11).

Adicionalmente, pode ser considerado um problema de otimização, uma vez que ao construir-se uma representação para este sistema (modelo de reservatórios) deseja-se minimizar a diferença entre a resposta simulada e os dados históricos observados.

Portanto, a função objetivo ( $O$ ) a ser minimizada pode ser descrita pela seguinte equação:

$$O = (\mathbf{d}_{hist} - \mathbf{d})^T \mathbf{C}_D^{-1} (\mathbf{d}_{hist} - \mathbf{d}), \quad (2-1)$$

onde  $\mathbf{d}_{hist}$  é um vetor coluna com os dados históricos medidos,  $\mathbf{d}$  é um vetor coluna com os dados simulados pelo modelo e  $\mathbf{C}_D$  é a matriz covariância do ruído nos dados medidos, a qual define o nível de confiança dos dados e, portanto, o peso relativo que cada dado terá no ajuste (3).

Entretanto, os dados simulados ( $\mathbf{d}$ ) podem ser considerados como o resultado de uma transformação  $\mathbf{g}(\cdot)$  aplicada sobre os parâmetros do modelo de simulação. Assim, pode-se considerar que:

$$\mathbf{d} = \mathbf{g}(\mathbf{m}); \quad \text{e} \quad (2-2)$$

$$O(\mathbf{m}) = (\mathbf{d}_{hist} - \mathbf{g}(\mathbf{m}))^T \mathbf{C}_D^{-1} (\mathbf{d}_{hist} - \mathbf{g}(\mathbf{m})), \quad (2-3)$$

onde  $\mathbf{m}$  é um vetor coluna contendo os parâmetros do modelo e a transformação  $\mathbf{g}(\cdot)$  corresponde à simulação do modelo de reservatórios.

Ainda assim, é desejável que os parâmetros do modelo  $\mathbf{m}$  não se afastem significativamente da concepção geológica original. Assim, pode-se acrescentar um termo à função objetivo para que esta seja penalizada à medida que os parâmetros do modelo  $\mathbf{m}$  se afastem dos parâmetros do modelo geológico original  $\mathbf{m}_{prior}$ , ao longo do processo de ajuste de histórico:

$$O(\mathbf{m}) = (\mathbf{d}_{hist} - \mathbf{g}(\mathbf{m}))^T \mathbf{C}_D^{-1} (\mathbf{d}_{hist} - \mathbf{g}(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{prior}) \quad (2-4)$$

onde  $\mathbf{m}_{prior}$  é um vetor coluna contendo os parâmetros do modelo geológico original e  $\mathbf{C}_M$  é a matriz covariância dos parâmetros do modelo.

Portanto, o objetivo dos métodos de ajuste de histórico assistido tradicionais é obter modelos  $\mathbf{m}$  que minimizem a função  $O(\mathbf{m})$ , ou seja:

$$\mathbf{m}_{post} = \arg \min_{\mathbf{m}} \{O(\mathbf{m})\}, \quad (2-5)$$

onde  $\mathbf{m}_{post}$  é um vetor coluna contendo os parâmetros do modelo ajustado (ou modelo *a posteriori*). A literatura dispõe de vários métodos que se propõe a resolver este problema, sendo que as estratégias utilizadas para solucionar este problema de otimização podem ou não ser baseadas no cálculo do gradiente da função objetivo.

Os métodos baseados em gradiente utilizam as derivadas da função objetivo  $O(\mathbf{m})$  para determinar as atualizações a serem realizadas nos parâmetros do modelo (12). Uma estratégia simplificada seria utilizar o método de Newton (11) para atualizar iterativamente o modelo, conforme a seguinte equação:

$$\mathbf{m}^\ell = \mathbf{m}^{\ell-1} - \mathbf{H}_{\ell-1}^{-1} \nabla O(\mathbf{m}^{\ell-1}) \quad , \quad (2-6)$$

onde  $\ell$  é uma das iterações,  $\nabla O$  é o vetor das derivadas parciais de primeira ordem da função objetivo  $O(\mathbf{m})$  e  $\mathbf{H}$  é a matriz das derivadas parciais de segunda ordem de  $O(\mathbf{m})$  (denominada Hessiana). Assim, na primeira iteração ( $\ell = 1$ ) teríamos  $\mathbf{m}^0 = \mathbf{m}_{prior}$  e, ao se atingir um critério de parada predefinido, teríamos o modelo ajustado  $\mathbf{m}_{post} = \mathbf{m}^\ell$ .

Contudo, no contexto do ajuste de histórico de modelos de reservatórios, há poucas implementações baseadas em gradiente, como os métodos *Adjoint* (11, 13) e *Forward* (14), uma vez que, em geral, o cálculo das derivadas pode demandar um alto custo computacional, além de necessitar ser implementado intrinsecamente ao código do simulador de reservatórios.

Por outro lado, métodos estocásticos tem sido amplamente utilizados, dentre os quais podem ser citados os métodos baseados em algoritmos evolucionários (algoritmos genéticos) (7, 8, 9) e o método *Particle Swarm Optimization* (PSO) (15, 16, 17), devido à sua capacidade de encontrar múltiplas soluções para o problema de otimização, além de serem facilmente integráveis aos simuladores de reservatório comerciais.

Contudo, as principais desvantagens dos métodos estocásticos são a sua lenta convergência, devido à complexidade do problema e ao caráter explorató-

rio do método, o que leva à necessidade de uma severa redução dos parâmetros incertos (ajustáveis) a serem considerados, reduzindo significativamente o espaço de busca por soluções para ajuste de histórico (1). Estes métodos, em geral, buscam minimizar apenas o termo da função objetivo  $O(\mathbf{m})$  referente à diferença entre a resposta simulada e os dados históricos observados, representado pela equação 2-3.

Uma alternativa computacionalmente eficiente e de implementação relativamente simples são os métodos baseados em filtro de Kalman, que têm ganhado cada vez mais evidência na área de ajuste de histórico pela sua robustez e capacidade de se trabalhar com um grande número de parâmetros ajustáveis, além de serem facilmente acopláveis aos simuladores de reservatórios existentes (3).

## 2.1

### Métodos Baseados em Filtro de Kalman

O filtro de Kalman tem uma longa tradição de utilização na área de sistemas de controle, sendo uma técnica que se propõe a assimilar dados sequencialmente, de forma a se obter uma estimativa atualizada dos parâmetros ou estados de um modelo, através da correção de uma estimativa anteriormente obtida, podendo portanto ser classificado como um estimador recursivo (18).

Apesar de originalmente ter sido desenvolvido para tratar apenas de sistemas lineares, diversos métodos tem sido propostos para adaptar sua utilização a sistemas dinâmicos não lineares.

Dentre as abordagens que se demonstraram aplicáveis à área de ajuste de histórico de reservatórios, o primeiro foi o método *ensemble Kalman filter* (EnKF) que é uma implementação de Monte Carlo, que fornece uma alternativa para o filtro de Kalman tradicional, onde um conjunto de realizações (modelos) são utilizados para estimar a média e a matriz covariância dos parâmetros do modelo (11).

Contudo, mesmo apresentando resultados promissores, a assimilação sequencial dos dados (figura 2.1) representa um problema prático, pois é necessário parar e reiniciar a simulação de reservatórios a cada passo de tempo em que os dados são assimilados. Assim, há um prejuízo no tempo total do processo de ajuste, devido à necessidade de se lidar com arquivos de reinicialização das simulações, sendo ainda mais crítico caso sejam adotadas soluções de simulação distribuída (*clusters*) para executar várias simulações paralelamente.

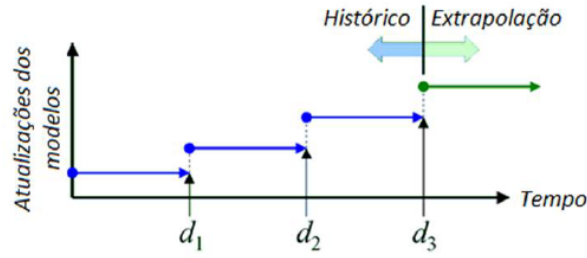


Figura 2.1: Processo de assimilação de dados pelo método EnKF (3).

Para contornar este problema, uma alternativa seria adotar o método *ensemble smother* (ES) (19), que apresenta uma formulação similar ao EnKF, mas com todos os dados sendo assimilados simultaneamente (figura 2.2). Entretanto, o ajuste de histórico obtido pelo método ES não tem demonstrado um desempenho satisfatório (2).

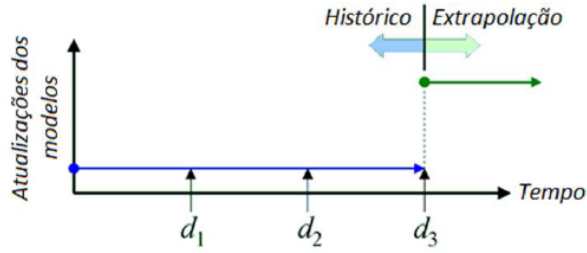


Figura 2.2: Processo de assimilação de dados pelo método ES (3).

Para melhorar a qualidade e a eficiência do ajuste de histórico obtido pelo ES, foi proposto o método *ensemble smoother with multiple data assimilation* (ES-MDA), podendo este ser considerado como um ES iterativo (figura 2.3), no qual o número de iterações é predefinido. Dentre os métodos baseados em filtro de Kalman, o ES-MDA é o que tem demonstrado os melhores resultados na literatura, tanto em termos de qualidade de ajuste como também na quantificação das incertezas na modelagem de reservatórios (3).

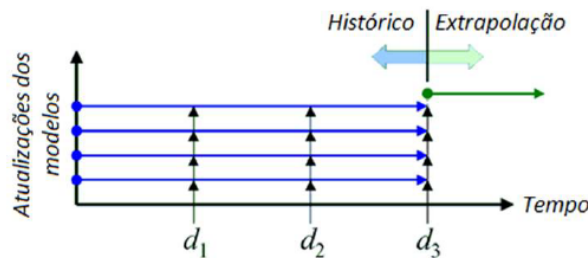


Figura 2.3: Processo de assimilação de dados pelo método ES-MDA (3).



## 2.1.1

**Ensemble Smoother with Multiple Data Assimilation**

O conceito básico do método ES-MDA, proposto por EMERICK e REYNOLDS (2), é realizar a assimilação de um mesmo conjunto de dados históricos ( $\mathbf{d}_{hist}$ ) múltiplas vezes de forma a se aproximar de um método iterativo, reduzindo assim os efeitos da não linearidade do problema, uma vez que a realização de um único passo de assimilação pode representar uma correção excessiva dos parâmetros do modelo (20).

O método se baseia na consideração de que, para casos lineares e Gaussianos, a assimilação dos mesmos dados múltiplas vezes com uma matrix de covariância do ruído dos dados medidos  $\mathbf{C}_D$  inflada (multiplicada por um fator  $\alpha > 1$ ) é equivalente a assimilar o dado uma única vez com a matriz de covariância original, desde que o coeficiente de inflação  $\alpha$  satisfaça a seguinte condição:

$$\sum_{\ell=1}^{N_a} \frac{1}{\alpha_\ell} = 1, \quad (2-7)$$

onde  $N_a$  é o número de ciclos de assimilação de dados (iterações) realizados pelo do método.

Assim, fica assegurado que a divisão do processo de assimilação de dados em passos menores seja realizada de forma estatisticamente consistente (20), o que, em conjunto à utilização de uma matriz de covariância inflada ( $\alpha \mathbf{C}_D$ ), reduz a magnitude da correção linear aplicada aos modelos a cada iteração.

Considerando uma realização  $\mathbf{m}_j$  de um conjunto de  $N_m$  modelos, o método ES-MDA aplicado para atualização do modelo durante uma iteração  $\ell$  pode ser descrito da seguinte forma:

$$\mathbf{m}_j^\ell = \mathbf{m}_j^{\ell-1} + \mathbf{C}_{MS}^\ell (\mathbf{C}_S^\ell + \alpha_\ell \mathbf{C}_D)^{-1} (\mathbf{d}_{hist} + \boldsymbol{\varepsilon}_j^\ell - \mathbf{g}(\mathbf{m}_j^{\ell-1})), \quad (2-8)$$

podendo ser reescrito como:

$$\mathbf{m}_j^\ell = \mathbf{m}_j^{\ell-1} + \mathbf{K}^\ell (\mathbf{d}_{hist} + \boldsymbol{\varepsilon}_j^\ell - \mathbf{g}(\mathbf{m}_j^{\ell-1})), \quad \text{sendo} \quad (2-9)$$

$$\mathbf{K}^\ell = \mathbf{C}_{MS}^\ell (\mathbf{C}_S^\ell + \alpha_\ell \mathbf{C}_D)^{-1}, \quad (2-10)$$

onde  $\mathbf{K}^\ell$  é denominado “ganho de Kalman”,  $\mathbf{C}_S$  é a matriz de covariância dos dados simulados,  $\mathbf{C}_{MS}$  é a matriz de covariância cruzada entre os parâmetros do modelo e os dados simulados,  $\boldsymbol{\varepsilon}$  é um ruído aleatório com uma distribuição de probabilidade  $\mathcal{N}(0, \alpha_\ell \mathbf{C}_D)$  e  $j = 1, 2, \dots, N_m$ .

A cada iteração  $\ell$  para assimilação de dados, as matrizes  $\mathbf{C}_S^\ell$  e  $\mathbf{C}_{MS}^\ell$  são estimadas considerando o resultado das simulações de todo o conjunto de modelos  $\{\mathbf{m}_j^{\ell-1}\}_{j=1}^{N_m}$ , atualizados na iteração anterior ( $\ell - 1$ ).

Portanto, partindo de um conjunto de realizações *a priori*  $\{\mathbf{m}_{prior,j}\}_{j=1}^{N_m}$  construído de forma a representar o modelo geológico conceitual do reservatório e incorporar as incertezas mapeadas, é possível aplicar o método para o ajuste de histórico assistido dos modelos. O conjunto de modelos iniciais corresponde às realizações *a priori* ( $\mathbf{m}_j^0 = \mathbf{m}_{prior,j}$ ). Ao final dos  $N_a$  ciclos de assimilação de dados é esperado que a resposta simulada dos modelos esteja mais aderente aos dados históricos de produção e que os modelos estejam mais próximos da representação real do reservatório.

O detalhamento dos passos para a aplicação do método ES-MDA será descrito no apêndice A.

### 2.1.2

#### **Localization**

Uma vez que os métodos baseados em filtro de Kalman se baseiam em hipóteses de Gaussianidade e linearidade, além de dispor (nas aplicações práticas de ajuste de histórico) de um conjunto relativamente pequeno de realizações para amostrar o espaço de soluções, o desempenho desses métodos pode ser prejudicado por correlações espúrias entre os dados medidos e os parâmetros do modelo ou pela redução excessiva das incertezas ao final do processo (3).

A técnica de *localization* se propõe a reduzir esses efeitos, fazendo com que a influência de um dado no ajuste dos parâmetros de um dos blocos do *grid* de simulação seja reduzida quanto maior for a distância que este dado tenha sido medido, em relação ao bloco. Em outras palavras, os dados medidos em um determinado poço serão responsáveis pelo ajuste dos parâmetros do modelo dentro de uma região de influência ao redor do poço.

Uma das formas de se aplicar o *localization* é considerar que o ganho de Kalman  $\mathbf{K}^\ell$  será regularizado por uma matriz de correlação  $\mathbf{R}_{MS}$ , cujos elementos são calculados, em geral, pela função de Gaspari-Cohn, com base no tamanho das regiões de influência de cada poço (21, 22). Essa regularização é realizada através da aplicação do produto elemento a elemento das duas matrizes (produto Schur), ou seja:

$$\mathbf{R}_{MS} \circ \mathbf{K}^\ell, \quad (2-11)$$

sendo possível portanto reescrever a equação 2-9, incorporando o *localization* ao processo ES-MDA:

$$\mathbf{m}_j^\ell = \mathbf{m}_j^{\ell-1} + (\mathbf{R}_{MS} \circ \mathbf{K}^\ell)(\mathbf{d}_{hist} + \boldsymbol{\epsilon}_j^\ell - \mathbf{g}(\mathbf{m}_j^{\ell-1})). \quad (2-12)$$

Portanto, a aplicação de ajuste de histórico descrita neste trabalho irá

considerar o método ES-MDA em conjunto à técnica de *localization*.

### 3

## Descrição do modelo de reservatórios

Para a realização do trabalho, foi construído um modelo de reservatórios baseado no modelo de referência SPE 9 (23). Este é um modelo de simulação *black-oil* com 24x25x15 células, contendo 25 poços produtores e um único injetor (figura 3.1).

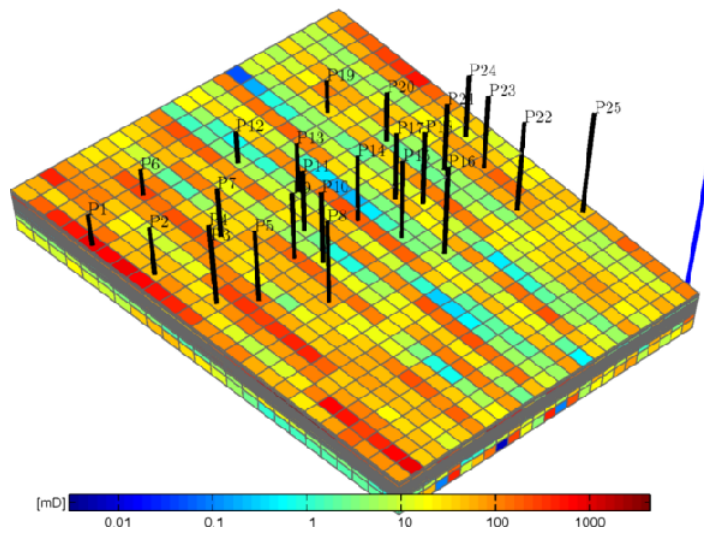


Figura 3.1: Modelo de reservatórios SPE 9.

Foram realizadas algumas adaptações no modelo original para melhor representar o problema, além de otimizar a utilização de recursos computacionais (tempo de simulação):

- ▷ incluídos 8 poços injetores adicionais;
- ▷ a resolução horizontal foi ampliada de 24x25 para 100x100; e
- ▷ considerada apenas a primeira camada do modelo.

A figura 3.2a representa a permeabilidade de matriz do modelo de simulação construído. Para fins práticos, a porosidade de matriz foi mantida constante em 17% em todo o modelo. A figura 3.2b é uma representação da posição dos poços no modelo, sendo que os poços destacados em vermelho foram arbitrariamente definidos como sendo poços que teriam cruzado corredores de fratura. Esta informação (*hard data*) será utilizada para condicionar as realizações de fraturas do modelo aos dados dos poços.

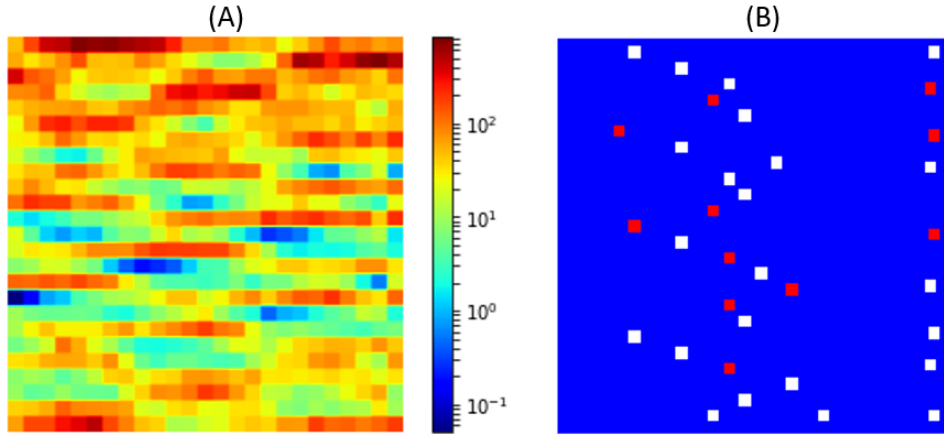


Figura 3.2: Mapa de permeabilidade de matriz (a) e posição dos poços no modelo (b).

### 3.1

#### Construção das realizações de fraturas

O modelo SPE 9 é um modelo de porosidade simples, portanto foi necessário construir um conjunto de realizações de ocorrência de fraturas que servirá para o treinamento/validação da rede convolucional *autoencoder*, além do conjunto de realizações que serão as variáveis *a priori* do processo de ajuste de histórico ES-MDA.

Para a construção das realizações foi utilizado o método *Single Normal Equation Simulation* (SNESIM). O SNESIM é um algoritmo de simulação multiponto (*Multiple-Point Simulation* – MPS) que simula variáveis categóricas e, a fim de reproduzir padrões, varre uma imagem de treinamento calculando a probabilidade de ocorrência de cada um dos padrões possíveis, que são guardados em árvores de busca (24), permitindo assim a construção de múltiplas realizações. Este método também permite condicionar as realizações aos dados de poços.

Assim, a partir de uma imagem de treinamento (figura 3.3), foram geradas 6200 realizações de ocorrência de fraturas. Essa propriedade irá representar as regiões do modelo onde há fraturas ativas (em vermelho) ou onde não há fraturas (em preto), sendo associadas aos valores de 1 e 0, respectivamente. Nas células onde as fraturas estão ativas, foram considerados valores de porosidade de fratura de 1% e permeabilidade de fratura de 1000 mD.

A construção das realizações de fratura foi realizada através da ferramenta *open-source* de geoestatística SGeMS (*Stanford Geostatistical Modeling Software*), que possui o método SNESIM implementado (25). A Figura 3.4 apresenta 9 amostras das 6200 realizações de fraturas criadas.

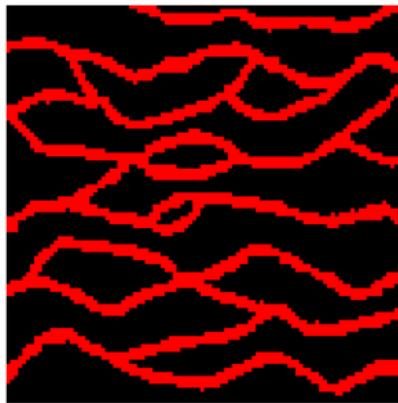


Figura 3.3: Imagem de treinamento representando a propriedade de ocorrência (em vermelho) e ausência (em preto) de fraturas (26).

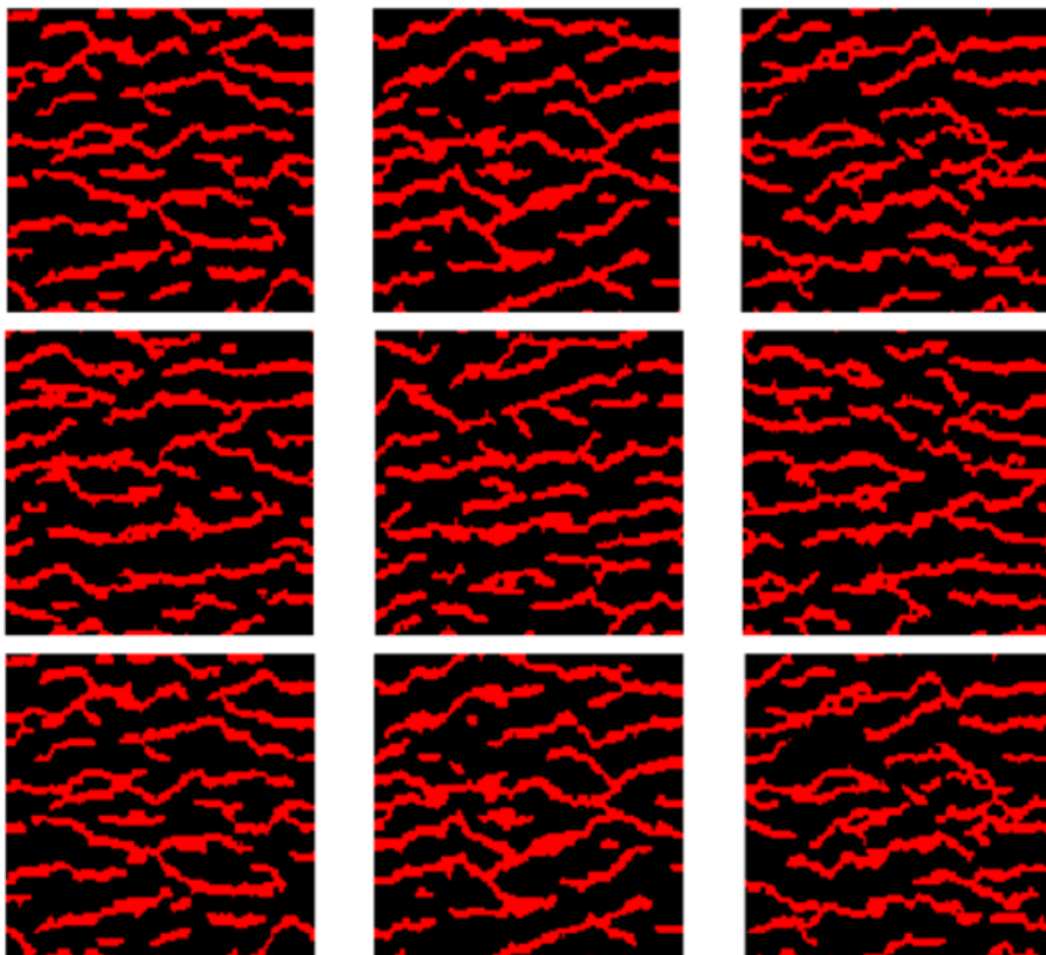


Figura 3.4: Exemplos de realizações de fraturas criadas.

## 4

### *Autoencoder*

Uma rede *autoencoder* é uma rede neural treinada para representar em sua saída uma cópia aproximada de sua entrada. Internamente, ela possui uma camada escondida cuja saída  $\mathbf{h}$  é uma propriedade codificada que representa o atributo de entrada. Portanto, este tipo de rede pode ser decomposto em duas partes (figura 4.1): uma função *encoder*  $\mathbf{h} = f(\mathbf{x})$  e uma função *decoder* que produz a reconstrução  $\mathbf{r} = g(\mathbf{h})$ , onde  $\mathbf{r}$  é uma representação aproximada da entrada  $\mathbf{x}$  (27).

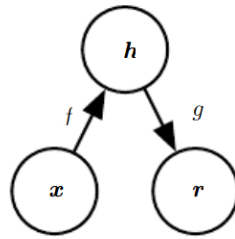


Figura 4.1: Estrutura geral de uma rede autoencoder (27).

Assim, será proposta uma rede convolucional *autoencoder* (figura 4.2) capaz de codificar um atributo categórico (ocorrência/ausência de fraturas) em uma propriedade com características Gaussianas que será utilizada como atributo de entrada para o processo de ajuste de histórico ES-MDA.

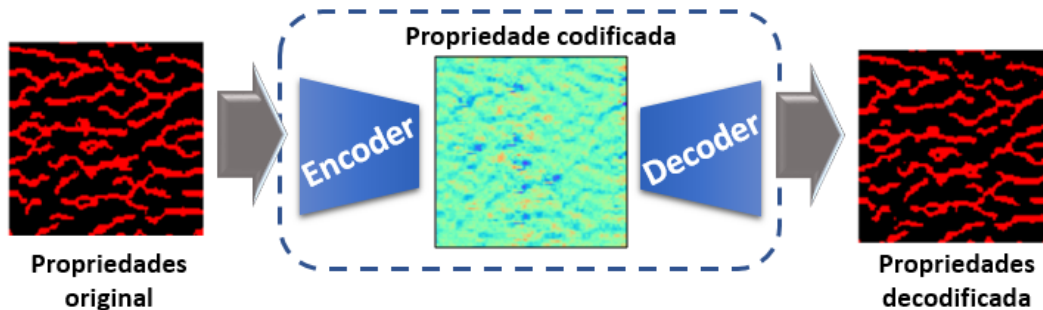


Figura 4.2: Representação da rede *autoencoder* proposta.

## 4.1

## Arquitetura da Rede

A rede *autoencoder* projetada para a solução do problema é descrita na Tabela 4.1, sendo composta de basicamente de um conjunto de camadas convolutivas (em duas dimensões), além de uma camada *fully connected* (*Dense*) ao final do processo de codificação (*encoder*), para generalização. A maior parte das camadas possuem funções de ativação lineares, exceto a camada de saída do processo de decodificação (*decoder*) que possui uma função de ativação sigmoial, pois o atributo de saída deve estar limitado entre 0 e 1.

Com o objetivo de aplicar a técnica de *localization* (seção 2.1.2) para melhorar o desempenho do método ES-MDA, a rede foi projetada de forma que a propriedade codificada mantivesse uma distribuição espacial similar ao atributo de entrada. Assim, a saída da rede *encoder* foi dimensionada com mesma resolução do atributo original (100x100) e todas as camadas convolutivas foram configuradas para preservar o formato da entrada (*padding*=“same”).

Tabela 4.1: Descrição das camadas da rede *autoencoder*.

Função	Layer	Output Shape	Kernel Size	Activation Function
Encoder	<i>Input</i>	(100, 100, 1)	3x3	-
	Conv2D 01	(100, 100, 256)	3x3	linear
	Conv2D 02	(100, 100, 128)	3x3	linear
	Conv2D 03	(100, 100, 64)	3x3	linear
	Conv2D 04	(100, 100, 32)	3x3	linear
	Conv2D 05	(100, 100, 1)	3x3	linear
	<i>Flatten</i>	(10000, 1)	-	linear
	<i>Dense</i>	(10000, 1)	-	linear
	<i>Reshape (encoded)</i>	(100, 100, 1)	-	-
Decoder	Conv2D 06	(100, 100, 32)	3x3	linear
	Conv2D 07	(100, 100, 64)	3x3	linear
	Conv2D 08	(100, 100, 128)	3x3	linear
	Conv2D 09	(100, 100, 256)	3x3	linear
	Conv2D 10 ( <i>Output</i> )	(100, 100, 1)	3x3	sigmoid



## 4.2

### Treinamento e Validação

Uma vez definida a rede, o conjunto de 6200 imagens de ocorrência de fraturas disponíveis foi dividida da seguinte forma:

- ▷ 5000 imagens para treinamento;
- ▷ 1000 imagens para validação; e
- ▷ 200 imagens reservadas para a etapa de ajuste de histórico.

A rede foi treinada com um *batch size* de 40, sendo inicialmente considerada a função erro (*loss function*) *binary crossentropy*, conforme figura 4.3, sendo que a curva azul corresponde ao conjunto de treinamento e a curva vermelha ao conjunto de validação. O treinamento foi interrompido antes que as curvas de treinamento e validação começassem a divergir, contudo atingindo valores de *loss* inferiores a 0.1.

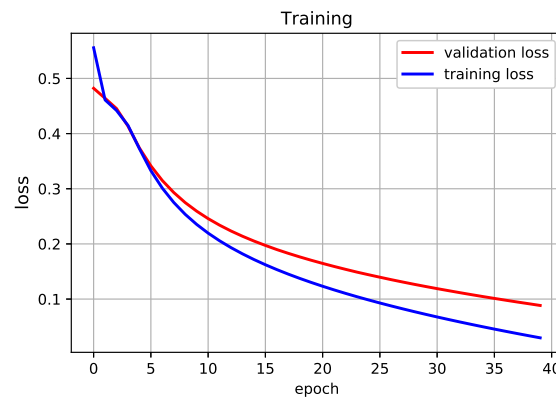


Figura 4.3: Treinamento da rede.

Para que a saída (após decodificação) possa ser utilizada como propriedade de ocorrência de fraturas na simulação de fluxo, os valores foram ainda truncados de forma que a imagem de saída apenas consistisse de valores discretos. A figura 4.4 mostra as saídas codificadas e decodificadas para 3 casos de ocorrência de fraturas do conjunto de validação, demonstrando uma excelente capacidade de recuperação da imagem original.

A figura 4.5 compara os histogramas do atributo original e codificado para três das realizações. Apesar de não ser possível garantir que a distribuição conjunta das realizações tenha características Gaussianas, ao se avaliar cada modelo percebe-se que o histograma do atributo codificado tem características de uma distribuição normal, enquanto originalmente poderia ser melhor representada por uma distribuição de Bernoulli (0 ou 1).

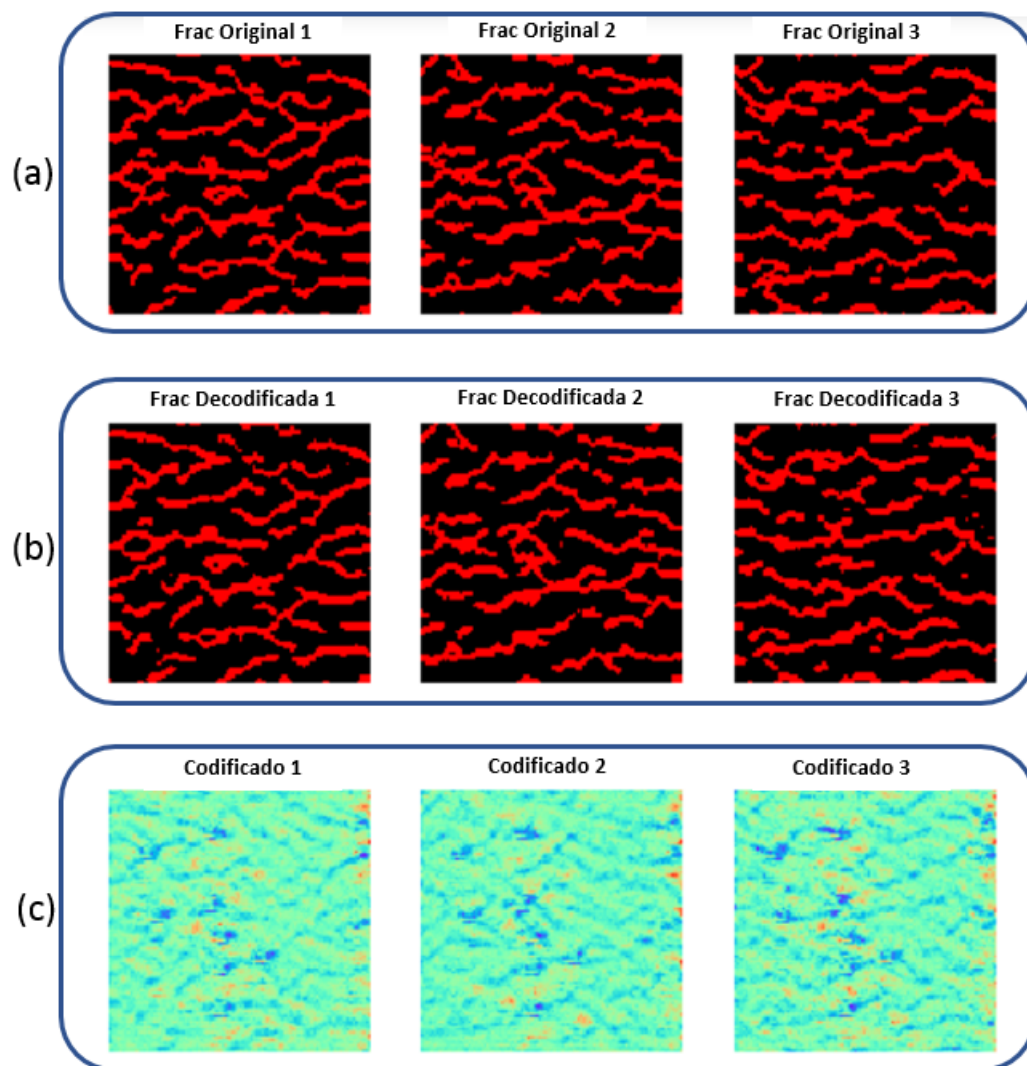


Figura 4.4: Atributos de entrada (a), codificado (c) e decodificado (b) pela rede *autoencoder* de três realizações de fraturas.

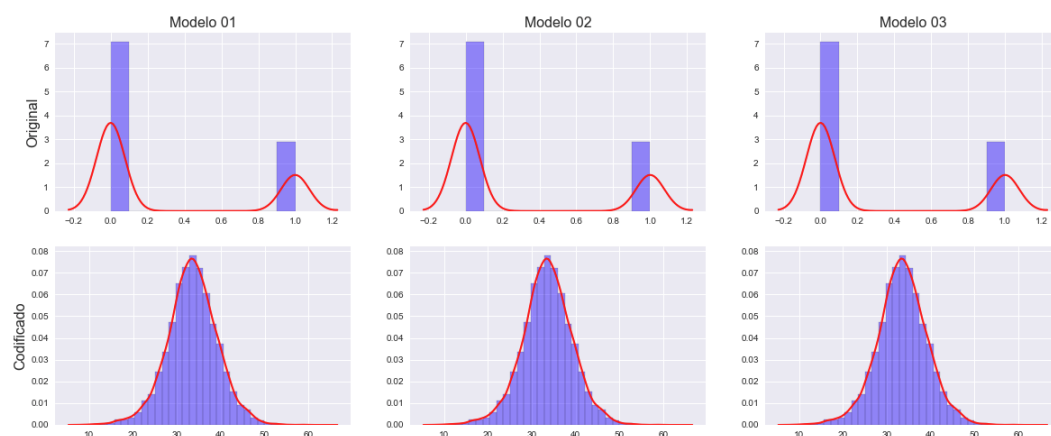


Figura 4.5: Histogramas do atributo original e codificado para três realizações de fraturas.

## 5

## Resultados

Ao final do processo de treinamento da rede, iniciou-se a etapa de ajuste de histórico utilizando o método ES-MDA. Foi escolhida uma dentre as 200 realizações de fratura disponíveis nesta etapa para representar o reservatório real ( $\mathbf{m}_{true}$ ). Portanto, o conjunto de modelos *a priori* que servirá de entrada para o processo de ajuste assistido foi constituído de 199 realizações.

O modelo  $\mathbf{m}_{true}$  foi então simulado por um período de produção de 3000 dias e seu resultado foi considerado uma representação dos dados históricos ( $\mathbf{d}_{hist}$ ) a serem ajustados, constituindo-se das vazões de óleo, água e gás dos 25 poços produtores, das vazões de água injetada pelos 9 poços injetores, além das pressões de fundo de todos os 34 poços.

O ajuste de histórico assistido considerou 6 etapas de assimilação para o método ES-MDA. Terminado o processo, a qualidade de ajuste pode ser quantificada calculando-se o erro quadrático médio (MSE) entre os dados simulados de cada modelo ( $\mathbf{d}$ ) e os dados históricos ( $\mathbf{d}_{hist}$ ):

$$MSE = \frac{1}{2N} \sum_{i=1}^{N_d} \left( \frac{d_i - d_{hist_i}}{\sigma_{hist_i}} \right)^2, \quad (5-1)$$

onde  $N_d$  é o número de dados observados e  $\sigma_{hist}$  o desvio padrão do ruído de medição (3).

A comparação entre os valores de MSE dos modelos iniciais (*a priori*) e dos modelos ajustados (*a posteriori*), apresentados na figura 5.1, atestam a melhora significativa da qualidade do ajuste global dos modelos.

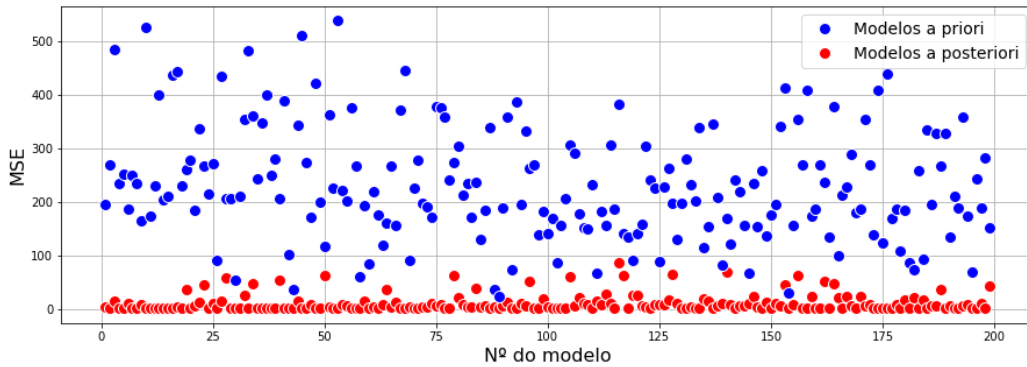


Figura 5.1: Comparação dos valores do erro quadrático médio (MSE) dos modelos *a priori* e *a posteriori*.

Detalhando a cada iteração para assimilação de dados (figura 5.2), pode-se observar uma redução sistemática do erro médio de todos os modelos ao longo do processo de ajuste.

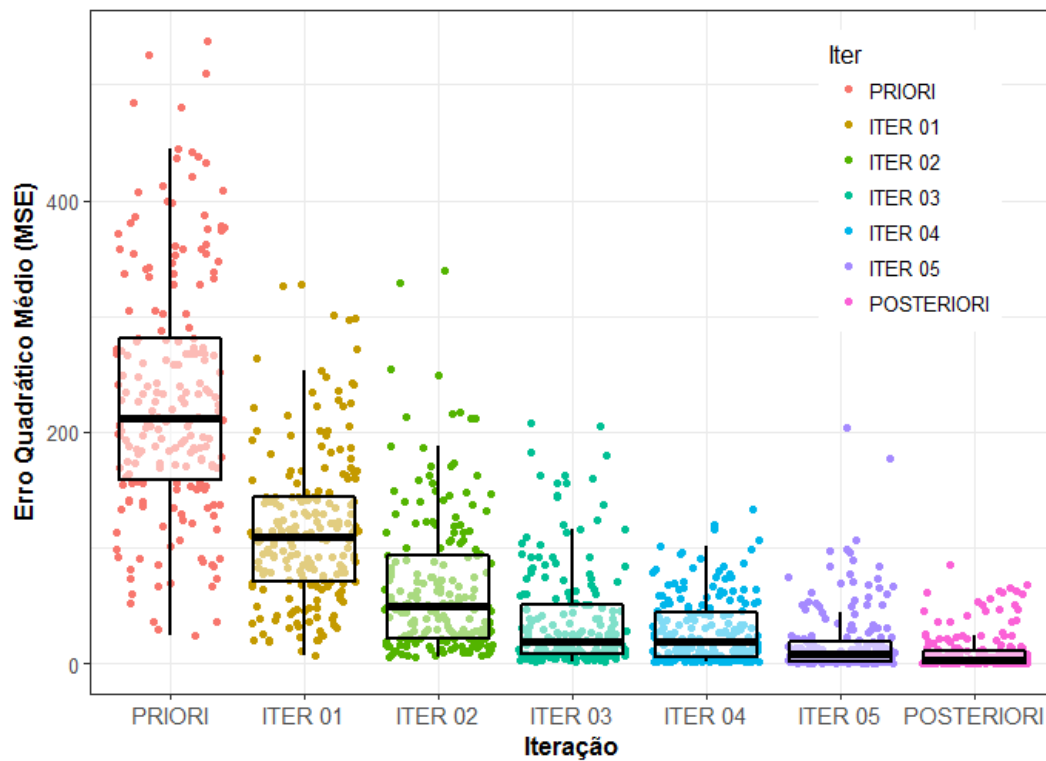


Figura 5.2: *Boxplot* do erro quadrático médio dos modelos a cada iteração.

## 5.1

### Dados de produção

A melhora no ajuste de histórico dos poços também pode ser verificada qualitativamente observando-se os dados de produção poço a poço.

As figuras 5.3 a 5.6 apresentam os resultados do ajuste de histórico de produção de óleo, água, gás e de pressão de fundo para quatro poços produtores e a figura 5.7 apresenta o ajuste de pressão de fundo de quatro poços injetores, sendo que os dados históricos estão representados em vermelho, os resultados simulados dos modelos iniciais em cinza, os modelos ajustados em azul claro e os 5 melhores modelos ajustados em azul escuro.

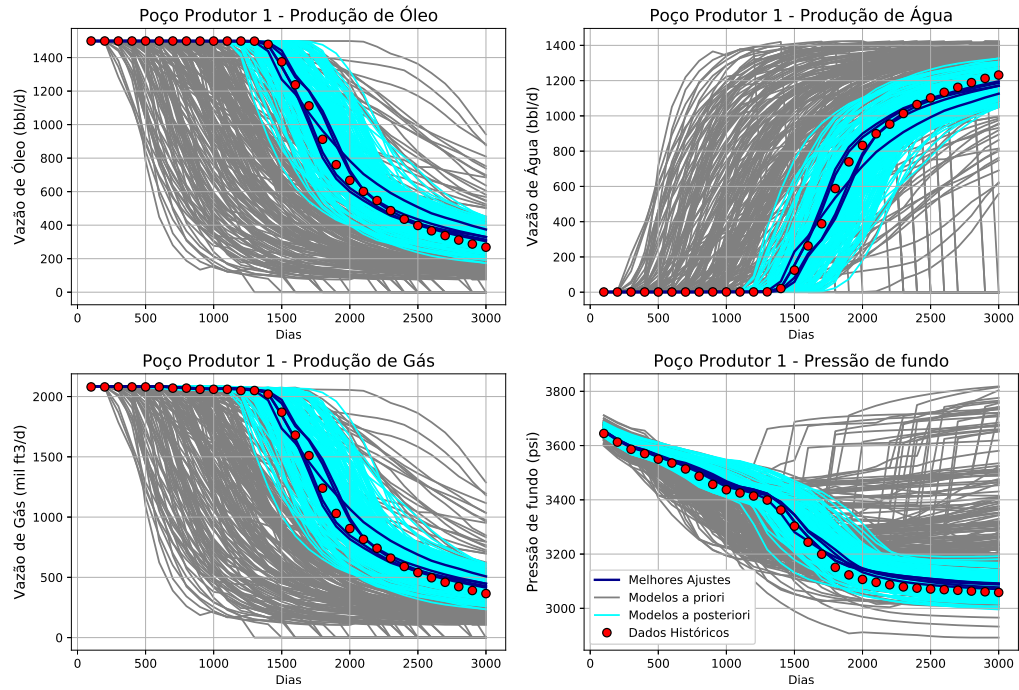


Figura 5.3: Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 1.

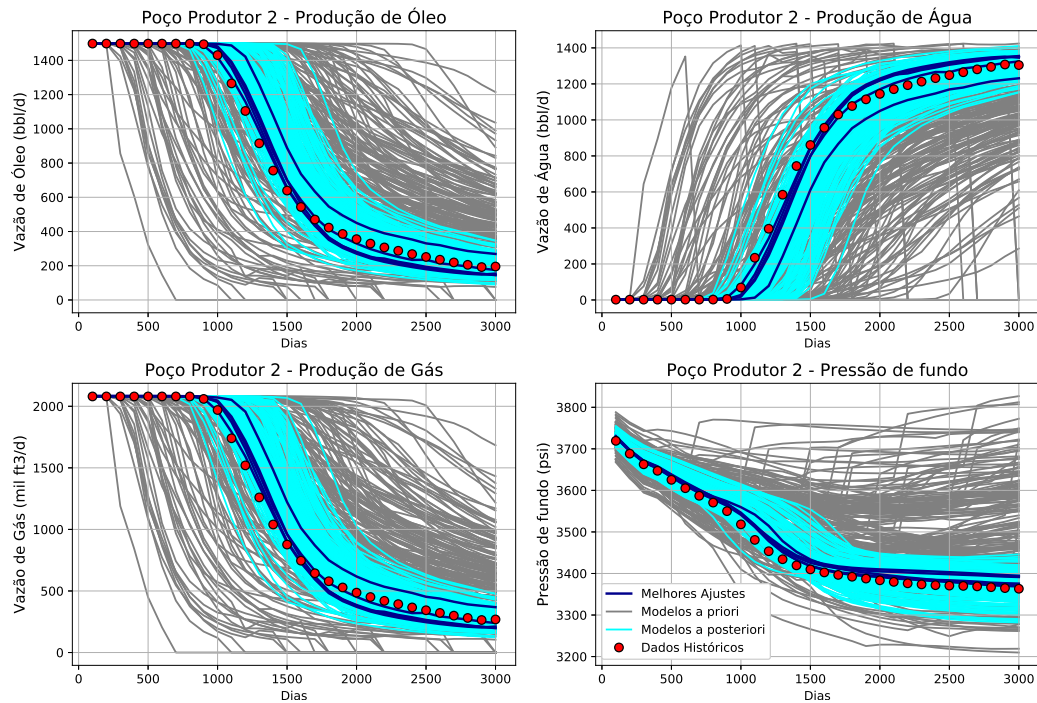


Figura 5.4: Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 2.



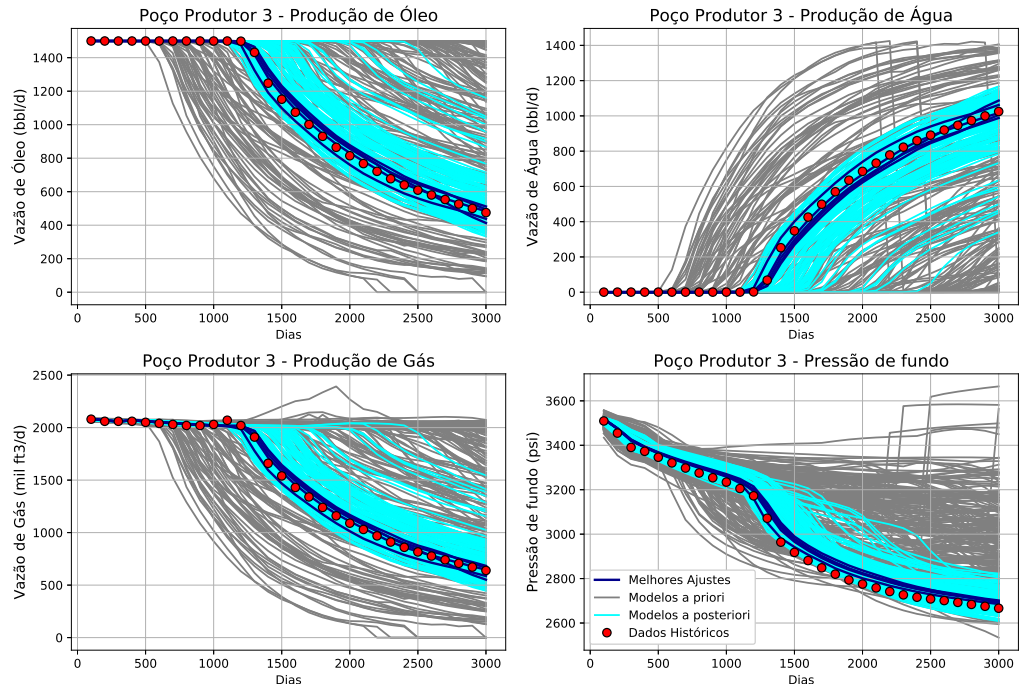


Figura 5.5: Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 3.

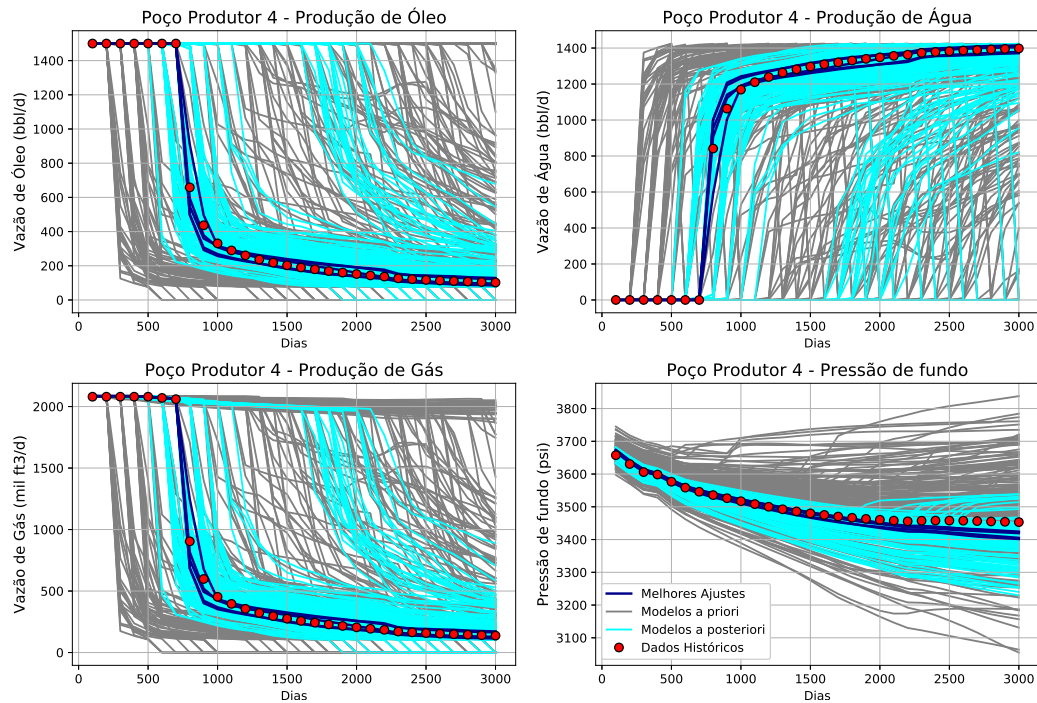


Figura 5.6: Ajuste de histórico das produções de óleo, água, gás e da pressão de fundo do poço produtor 4.

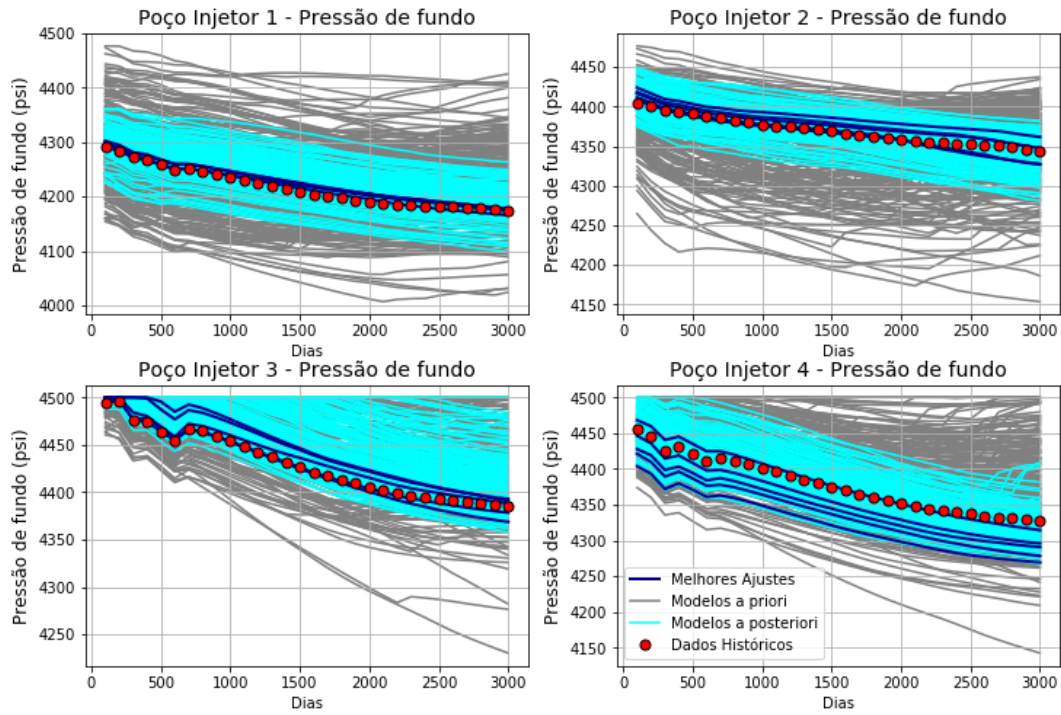


Figura 5.7: Ajuste de histórico da pressão de fundo de 4 poços injetores.

## 5.2

### Mapas de ocorrência de fraturas

Ao observarmos a propriedade de ocorrência de fraturas antes e após o ajuste (figura 5.8) podemos verificar que nos modelos finais as principais direções e a continuidade dos corredores de fraturas foram bem representadas ao compararmos com a imagem do modelo representativo do reservatório real ( $\mathbf{m}_{true}$ ).

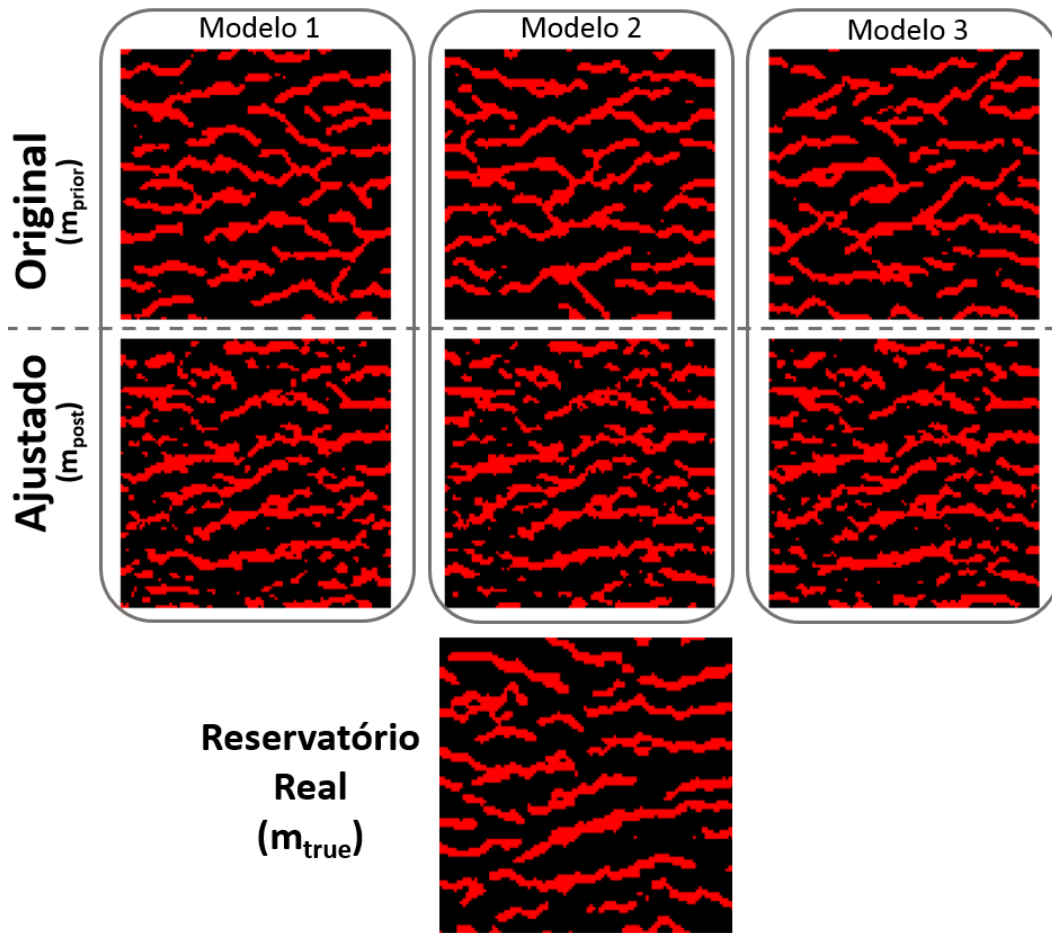


Figura 5.8: Comparação entre 3 modelos de fraturas antes e após o ajuste em relação ao reservatório real ( $m_{\text{true}}$ ).

### 5.3

#### Mapas de saturação de água

Os bons resultados são ainda mais aparentes quando comparamos os mapas de saturação de água dos modelos ao final do período de produção (3000 dias). A figura 5.9 mostra que os modelos originais não representam adequadamente os caminhos preferenciais de água observados no reservatório real. Após o ajuste (figura 5.10) os mesmos modelos apresentam um comportamento de canalização de água muito próximos ao esperado.



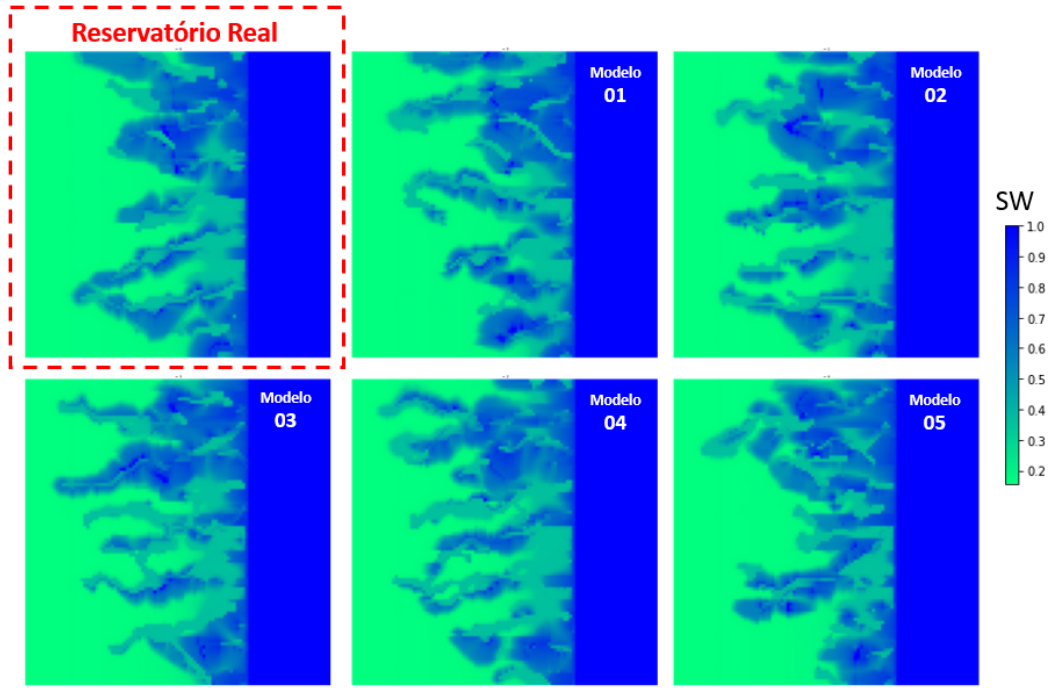


Figura 5.9: Comparação entre mapas saturação de água do reservatório real ( $\mathbf{m}_{true}$ ) e dos modelos originais (*a priori*) após 3000 dias de produção.

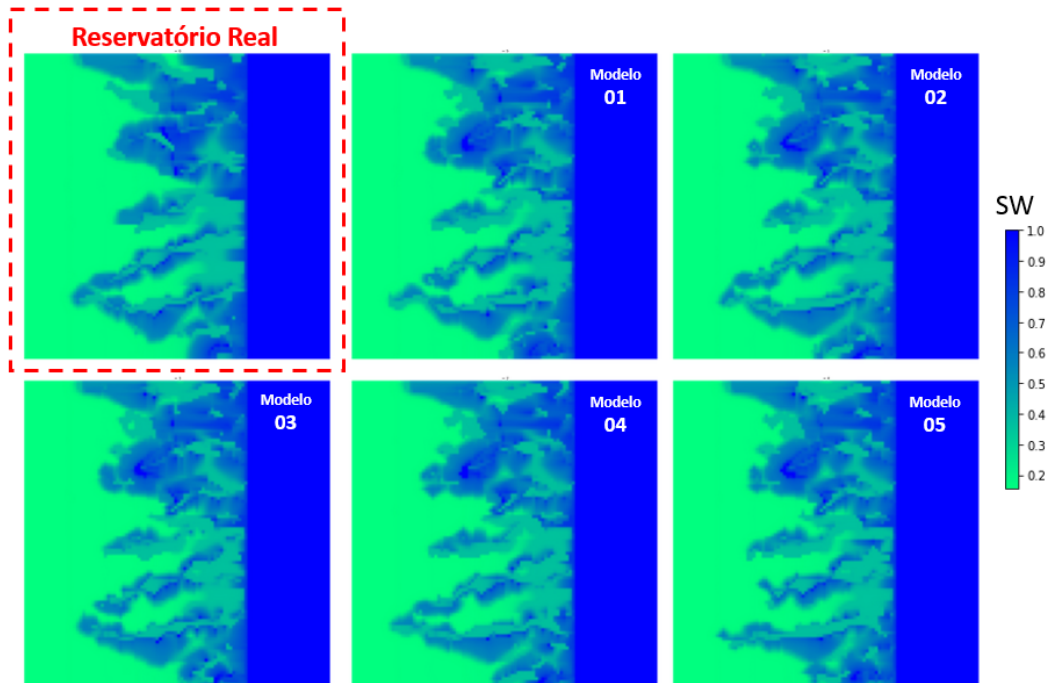


Figura 5.10: Comparação entre mapas saturação de água do reservatório real ( $\mathbf{m}_{true}$ ) e dos modelos ajustados (*a posteriori*) após 3000 dias de produção.

## 6

### Conclusão

A incorporação de uma rede convolucional *autoencoder* ao processo de ajuste de histórico utilizando o método baseado em filtro de Kalman *ensemble smoother with multiple data assimilation* (ES-MDA) foi bem sucedida em viabilizar uma alternativa que permita considerar atributos discretos (ocorrência/ausência de fraturas) como parâmetros de entrada do processo.

Os resultados obtidos após a aplicação da metodologia ao modelo proposto demonstram a boa capacidade do método de fornecer modelos representativos aderentes às características geológicas do reservatório real e cujas respostas simuladas estejam ajustadas aos dados históricos de produção, conferindo assim uma maior confiabilidade à previsão de produção destes modelos.

A opção por adotar uma metodologia baseada em *deep learning* de forma a contornar as limitações do ES-MDA se mostrou altamente positiva, devido a sua capacidade de, através do aprendizado não supervisionado das características de um conjunto de dados de treinamento, realizar transformações complexas e tratar os dados em múltiplos níveis de abstração, permitindo a codificação e a reconstrução dos dados de entrada com um alto grau de fidelidade.

Apesar dos bons resultados alcançados para esta aplicação, a arquitetura de rede adotada neste trabalho poderá não ser escalável para tratar modelos de reservatórios muito grandes, uma vez que a necessidade de se incluir uma camada *fully connected* ao final do processo de codificação (*encoder*) aumenta o número de parâmetros treináveis da rede proporcionalmente ao número de células do modelo de simulação, podendo atingir assim o limite de memória disponível das placa gráficas (GPU) utilizadas na etapa de treinamento da rede *autoencoder*.

Contudo, uma alternativa viável seria não utilizar a técnica de *localization* durante o processo de ajuste de histórico, permitindo que a saída da rede *encoder* possa gerar um atributo codificado com dimensões inferiores ao atributo original. Essa abordagem, além da redução dos parâmetros treináveis da rede, ainda traria a vantagem de reduzir o número de parâmetros de modelo a serem ajustados pelo método ES-MDA.

Portanto, como sugestão para trabalhos futuros pode-se propor a avali-

ação de arquiteturas de rede alternativas que possam se adequar às especificidades de outros modelos de reservatório. Uma vez que a área de *deep learning* encontra-se em crescente desenvolvimento, existe um amplo conjunto de novas soluções que podem ser aplicadas para melhorar o desempenho do processo de ajuste de histórico.

## Referências bibliográficas

- [1] JOHANSEN, K. **Statistical Methods for History Matching**. FRYDENBERG AS, Copenhagen (DK), 2008.
- [2] EMERICK, A. A.; REYNOLDS, A. C. **Ensemble smoother with multiple data assimilation**. *Computers & Geosciences*, 2013.
- [3] SILVA, V. L. S. **Ajuste de histórico e otimização da produção de petróleo sob incertezas - aplicação do gerenciamento de reservatórios em malha fechada**. Dissertação de mestrado, Universidade Federal do Rio de Janeiro - UFRJ, Rio de Janeiro, 2016.
- [4] CANCHUMUNI, S.A.; EMERICK, A. A.; PACHECO, M. A. **Integration of ensemble data assimilation and deep learning for history matching facies models**. *Offshore Technology Conference*, 2017.
- [5] ZHAO, Y.; FOROUZANFAR, F.; REYNOLDS, A. C. **Assisted history matching for multi-facies channelized reservoir using es-mda with common basis DCT**. *ECMOR XV*, 2016.
- [6] SEBACHER, M. B. **Data Assimilation under Geological Constraints**. Tese de doutorado, Delft University of Technology, Delft (NL), 2014.
- [7] XAVIER, C. R.; SANTOS, E. P.; VIEIRA, V. F.; SANTOS, R. W. **Genetic algorithm for the history matching problem**. *Procedia Computer Science*, 2013.
- [8] CHITHRA CHAKRA, N. C.; SARAF, D. N. **History matching of petroleum reservoirs employing adaptive genetic algorithm**. *Journal of Petroleum Exploration and Production Technology*, 2015.
- [9] MONFARED, A. D.; HELALIZADEH, A.; PARVIZI, H. **Automatic history matching using the integration of response surface modeling with a genetic algorithm**. *Petroleum Science and Technology*, 2011.
- [10] WU, A. **A new integrated approach of natural fracture modelling to improve historical matching and prediction for devonian car-**

- bonate reservoir in Nenetskoye oil field, Russia. PETROVIETNAM JOURNAL, 2018.
- [11] OLIVER, D. S.; REYNOLDS, A. C.; LIU, N. **Inverse Theory for Petroleum Reservoir Characterization and History Matching**. Cambridge University Press, Cambridge (UK), 2008.
- [12] NEELAKANTAN, P. **Comparison of multi-model history matching methods**. Dissertação de mestrado, Delft University of Technology, Delft (NL), 2017.
- [13] JANSEN, J. D.; DOUMA, S. D.; BROUWER, D. R.; VAN DEN HOF, P. M. J.; BOSGRA, O. H.; HEEMINK, A. W. **Closed-loop reservoir management**. SPE 119098, 2009.
- [14] RODRIGUES, J. R. P. **Calculating derivatives for automatic history matching**. Computational Geosciences, 10:119–136, 2006.
- [15] KENNEDY, J.; EBERHART, R. **Particle swarm optimization**. IEEE ICNN, 1995.
- [16] MOHAMED, L.; CHRISTIE, M.; DEMYANOV, V.; ROBERT, E.; KACHUMA, D. **Application of particle swarms for history matching in the brugge reservoir**. SPE 135264, 2010.
- [17] REYNOLDS, A. P.; ABDOLLAHZADEH, A.; CORNE, D. W.; CHRISTIE, M.; DAVIES, B.; WILLIAMS, G. **Guide objective assisted particle swarm optimization and its application to history matching**. PPSN, 2012.
- [18] AGUIRRE, L. A. **Introdução à identificação de sistemas: técnicas lineares e não lineares: teoria e aplicação**. Editora UFMG, Belo Horizonte, 2015.
- [19] SKJERVHEIM, J.; EVENSEN, G. **An ensemble smoother for assisted history matching**. SPE 141929, 2011.
- [20] FOSSUM, K. **Assessment of Sequential and Simultaneous Ensemble-based History Matching Methods for Weakly Non-linear Problems**. Tese de doutorado, University of Bergen, Bergen (NO), 2015.
- [21] 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, 2011.

- [22] EMERICK, A. A.; REYNOLDS, A. C. **History matching a field case using the ensemble Kalman filter with covariance localization.** SPE Reservoir Evaluation & Engineering, 2011.
- [23] KILLOUGH, J. E. **Ninth SPE comparative solution project: A reexamination of black-oil simulation.** SPE 29110, 1995.
- [24] TAKAFUJI, E. H. M.; ROCHA, M. M.; RAMOS, G. Z. **Estudo comparativo dos métodos de simulação SNESIM e FILTERSIM – aplicado a um modelo sintético de cobre.** Boletim Paranaense de Geociências, Vol. 74 (UFPR), 2018.
- [25] REMY, N.; BOUCHER, A.; WU, J. **Applied Geostatistics with SGeMS: A User’s Guide.** Cambridge University Press, Cambridge (UK), 2009.
- [26] STREBELLE, S. **Conditional simulation of complex geological structures using multiple-point statistics.** Mathematical Geology, Vol. 34, No. 1, 2002.
- [27] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning.** The MIT Press, Massachusetts (USA), 2016.

## A

### Algoritmo ES-MDA

O algoritmo para a aplicação do método ES-MDA, adaptado de SILVA (3), é descrito a seguir:

#### 1. Inicialização:

- ▷ Definir o conjunto de realizações (modelos) *a priori*  $\{\mathbf{m}_j^0\}_{j=1}^{N_m}$ , sendo:

$$\mathbf{m}_j^0 = [m_{j,1}^0 \quad m_{j,2}^0 \quad \dots \quad m_{j,N_p}^0]^T, \quad (\text{A-1})$$

onde  $N_m$  é o número de modelos *a priori* e  $N_p$  é o número de parâmetros do modelo;

- ▷ Definir vetor de dados históricos medidos  $\mathbf{d}_{hist}$ , com dimensão  $N_d$  (número total de dados):

$$\mathbf{d}_{hist} = [d_{hist_1} \quad d_{hist_2} \quad \dots \quad d_{hist_{N_d}}]^T; \quad (\text{A-2})$$

- ▷ Definir a matriz covariância do ruído nos dados medidos  $\mathbf{C}_D$ . Esta é uma matriz em que os elementos de sua diagonal principal correspondem às variâncias do ruído em cada um dos dados históricos medidos. Se for assumido que os ruídos nos dados são independentes, os demais elementos da matriz são nulos:

$$\mathbf{C}_D = \begin{bmatrix} \sigma_{d_{hist_1}}^2 & 0 & \dots & 0 \\ 0 & \sigma_{d_{hist_2}}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{d_{hist_{N_d}}}^2 \end{bmatrix}; \quad (\text{A-3})$$

- ▷ Definir o número de iterações para assimilação de dados  $N_a$ ; e
- ▷ Definir os fatores de inflação  $\alpha_\ell$ , para  $\ell = 1, 2, \dots, N_a$ , condicionados a seguinte restrição:

$$\sum_{\ell=1}^{N_a} \frac{1}{\alpha_\ell} = 1. \quad (\text{2-7})$$

**2. Função ES-MDA( $\mathbf{m}_j^0, \mathbf{d}_{hist}, \mathbf{C}_D, N_a, \alpha_\ell$ ):**

**Para**  $\ell = 1$  **até**  $N_a$ , **faça**:

- ▷ Rodar os modelos do conjunto  $\{\mathbf{m}_j^{\ell-1}\}_{j=1}^{N_m}$ , do tempo zero até o final do histórico de dados, para gerar o vetor de dados simulados:

$$\mathbf{d}_j^{\ell-1} = \mathbf{g}(\mathbf{m}_j^{\ell-1}), \quad \text{para } j = 1, 2, \dots, N_m. \quad (\text{A-4})$$

- ▷ Estimar as matrizes de covariância  $\mathbf{C}_S$  e  $\mathbf{C}_{MS}$ , a partir do conjunto de modelos e seus respectivos dados simulados:

$$\mathbf{C}_S^\ell = \frac{1}{N_m - 1} \sum_{j=1}^{N_m} (\mathbf{d}_j^{\ell-1} - \bar{\mathbf{d}}^\ell)(\mathbf{d}_j^{\ell-1} - \bar{\mathbf{d}}^\ell)^T, \quad (\text{A-5})$$

$$\mathbf{C}_{MS}^\ell = \frac{1}{N_m - 1} \sum_{j=1}^{N_m} (\mathbf{m}_j^{\ell-1} - \bar{\mathbf{m}}^\ell)(\mathbf{d}_j^{\ell-1} - \bar{\mathbf{d}}^\ell)^T, \quad (\text{A-6})$$

onde

$$\bar{\mathbf{d}}^\ell = \frac{1}{N_m} \sum_{j=1}^{N_m} \mathbf{d}_j^{\ell-1} \quad \text{e} \quad \bar{\mathbf{m}}^\ell = \frac{1}{N_m} \sum_{j=1}^{N_m} \mathbf{m}_j^{\ell-1}. \quad (\text{A-7})$$

- ▷ Perturbar o vetor de dados históricos  $\mathbf{d}_{hist}$  com um ruído aleatório  $\boldsymbol{\varepsilon}_j$  da seguinte forma:

$$\boldsymbol{\varepsilon}_j^\ell = \sqrt{\alpha_\ell} \mathbf{C}_D^{1/2} \mathbf{z}_j, \quad (\text{A-8})$$

$$\mathbf{d}_{hist,j}^\ell = \mathbf{d}_{hist} + \boldsymbol{\varepsilon}_j^\ell, \quad (\text{A-9})$$

para  $j = 1, 2, \dots, N_m$  e onde  $\mathbf{z}_j$  é um vetor coluna aleatório com dimensão  $N_d$ , amostrado com uma distribuição normal de média zero e desvio padrão um ( $\mathcal{N}(0, 1)$ ).

- ▷ Atualizar o vetor de parâmetros do modelo  $\mathbf{m}_j^\ell$ , para  $j = 1, 2, \dots, N_m$ :

$$\mathbf{m}_j^\ell = \mathbf{m}_j^{\ell-1} + \mathbf{C}_{MS}^\ell (\mathbf{C}_S^\ell + \alpha_\ell \mathbf{C}_D)^{-1} (\mathbf{d}_{hist,j}^\ell - \mathbf{d}_j^{\ell-1}). \quad (2-8)$$

**Fim Para**

**Fim Função**

Ao final, temos o conjunto de modelos *a posteriori* ajustados  $\{\mathbf{m}_j^{N_a}\}_{j=1}^{N_m}$ .