



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
E INSTITUTO DE GEOCIÊNCIAS

MARX VLADIMIR DE SOUSA MIRANDA

**UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA
PARA ACELERAR A OTIMIZAÇÃO DE
ESTRATÉGIAS DE PRODUÇÃO BASEADA EM
MODELOS**

CAMPINAS

2024

MARX VLADIMIR DE SOUSA MIRANDA

**UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA
PARA ACELERAR A OTIMIZAÇÃO DE
ESTRATÉGIAS DE PRODUÇÃO BASEADA EM
MODELOS**

Dissertação de Mestrado apresentada à Faculdade de Engenharia Mecânica e Instituto de Geociências da Universidade Estadual de Campinas como parte dos requisitos exigidos para obtenção do título de Mestre em Ciências e Engenharia de Petróleo, na área de Reservatórios e Gestão.

Orientador: Prof. Dr. Denis José Schiozer

Coorientador: Prof. Dr. Guilherme Daniel Avansi

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Marx Vladimir de Sousa Miranda e orientada pelo Prof. Dr. Denis José Schiozer.

CAMPINAS

2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

M672u Miranda, Marx Vladimir de Sousa, 1985-
Utilização de aprendizado de máquina para acelerar a otimização de estratégias de produção baseada em modelos / Marx Vladimir de Sousa Miranda. – Campinas, SP : [s.n.], 2024.

Orientador: Denis José Schiozer.
Coorientador: Guilherme Daniel Avansi.
Dissertação (mestrado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Mecânica.
Em regime multiunidades com: Instituto de Geociências.

1. Aprendizado de máquina. 2. Metodologia de superfície de resposta. 3. Otimização. 4. Reservatórios carbonáticos. 5. Simulação de reservatórios. I. Schiozer, Denis José, 1963-. II. Avansi, Guilherme Daniel, 1984-. III. Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia Mecânica. IV. Título.

Informações Complementares

Título em outro idioma: Use of machine learning to speed up model-base production strategy optimization

Palavras-chave em inglês:

Carbonate reservoir

Machine learning

Optimization

Reservoir simulation

Response surface methodology

Área de concentração: Reservatórios e Gestão

Titulação: Mestre em Ciências e Engenharia de Petróleo

Banca examinadora:

Denis José Schiozer [Orientador]

João Carlos Von Hohendorff Filho

Eduardo Camponogara

Data de defesa: 22-08-2024

Programa de Pós-Graduação: Ciências e Engenharia de Petróleo

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0009-0001-1060-1622>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4125414299578687>

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
E INSTITUTO DE GEOCIÊNCIAS

DISSERTAÇÃO DE MESTRADO ACADÊMICO

**UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA
PARA ACELERAR A OTIMIZAÇÃO DE
ESTRATÉGIAS DE PRODUÇÃO BASEADA EM
MODELOS**

Autor: Marx Vladimir de Sousa Miranda
Orientador: Prof. Dr. Denis José Schiozer
Coorientador: Prof. Dr. Guilherme Daniel Avansi

A Banca Examinadora composta pelos membros abaixo aprovou esta Dissertação:

Prof. Dr. Denis José Schiozer, Presidente
DEP / FEM / UNICAMP

Prof. Dr. João Carlos Von Hohendorff Filho
UNISIM / CEPETRO / UNICAMP

Prof. Dr. Eduardo Camponogara
Departamento de Automação e Sistemas / Universidade Federal de Santa Catarina

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Campinas, 22 de agosto de 2024.

DEDICATÓRIA

Dedico este trabalho aos meus pais, Alice e Naylor, cujos incansáveis apoio e incentivo foram a base de toda a minha jornada.

AGRADECIMENTOS

A conclusão deste trabalho só foi possível graças ao apoio e à colaboração de muitas pessoas, às quais expresso minha profunda gratidão:

Primeiramente, agradeço a Deus pela força e orientação em cada etapa da minha jornada.

Ao meu orientador, professor Denis, por todo suporte, experiência, paciência e pelos ensinamentos valiosos. Sou grato pela oportunidade de fazer parte do UNISIM, em especial do subgrupo MFM, uma equipe extremamente profissional que me proporcionou inúmeras oportunidades de aprendizado e crescimento, tanto pessoal quanto profissional.

Ao meu coorientador, professor Guilherme Avansi, pela constante disposição em ajudar, pela amizade e pelos valiosos ensinamentos.

Aos membros do grupo MFM – Célio Maschio, Eduardo Ramires, Fábio Bourdeux, Gabriel Cirac, Igor Vinícius, Jeanfranco Farfan, Leandro Danes, Luis Pineda, Mariana Marafon, Samuel Mello, Sharon Soler e Thiago Gomes – pelas discussões técnicas, companheirismo e contínua disposição em ajudar.

A todos os funcionários do CEPETRO, DEP e UNICAMP, especialmente a Paulo Drummond, Derek Vasconcelos e Guilherme Tonin, por todo o suporte prestado ao longo do mestrado. A Alice Obata e Leandro Fernandes, pela amizade e pelas conversas no café.

Aos professores da UNICAMP que, direta ou indiretamente, contribuíram para a realização deste trabalho, minha sincera gratidão.

Aos meus amigos do Departamento de Engenharia de Petróleo e da UNICAMP, pela amizade, companheirismo e pelas experiências compartilhadas.

Agradeço à, UNICAMP, UNISIM, FEM, CEPETRO por toda estrutura oferecida, assim como à CMG, Emerson, e Schlumberger pelas licenças dos *softwares*. Um agradecimento especial à CAPES e à Shell pelo suporte financeiro.

Aos meus familiares, pelo apoio incondicional durante toda a trajetória.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Esta dissertação foi realizada com apoio do projeto registrado sob o número ANP 21870-1 como “Desenvolvimento de uma Abordagem para Construção de Modelos Mutifidelidades para Reduzir Incertezas e Melhorar Previsão de Produção” (UNICAMP/Shell Brasil/ANP) financiado pela Shell Brasil, sob a taxa de P&D da ANP como “Compromisso de Investimentos com Pesquisa e Desenvolvimento”.

RESUMO

Os reservatórios de petróleo com litologia carbonática são caracterizados por apresentarem heterogeneidades complexas, as quais precisam ser adequadamente representadas nos modelos de simulação, que são utilizados para auxiliar tomadas de decisão. Como resultado, o tempo computacional necessário para rodar os modelos de simulação para estimar a produção pode ser muito grande. Durante um processo de otimização probabilístico de estratégias de produção, o tempo total é ainda maior de modo que, em alguns casos, pode ser necessário executar milhares de simulações para obter bons resultados. Devido à alta demanda computacional e ao seu impacto significativo no tempo de análise de decisão, são investigados métodos integrados de modelos de simulação e técnicas de inteligência artificial (IA) para reduzir a dependência do simulador no processo de otimização da produção. O estudo adota uma abordagem hierárquica, dividida em duas partes: (1) envolve a integração de modelos de simulação com o algoritmo de otimização hipercubo latino discreto iterativo (IDLHC) para desenvolver uma estratégia de produção nominal, otimizando variáveis de projeto a fim de obter a melhor configuração operacional e o máximo retorno econômico; (2) técnicas de IA são aplicadas para desenvolver um estimador mais rápido de função objetivo (FOFE) integrado ao IDLHC para ser utilizado nas partes mais custosas do processo de otimização. Os resultados confirmam que o FOFE é muito útil em determinadas fases do processo de otimização e deve ser avaliado com cautela em outras, oferecendo ao decisor a opção de escolher entre utilizar apenas modelos de fluxo, FOFE ou ambos. Além disso, nota-se que a FOFE é de grande valia na etapa de maior custo computacional, que é a determinação da posição dos poços. À medida que a otimização se torna mais refinada e a necessidade da precisão aumenta, a representatividade da FOFE em relação ao simulador tende a reduzir, ficando em uma margem de erro típica do processo de otimização, tornando-se, portanto, menos eficaz. Em suma, após a otimização com FOFE, o processo foi acelerado em mais de 50%, com uma redução de menos de 1% na qualidade dos resultados para os modelos de simulação (inferior ao erro esperado dos modelos comparados com o caso real). No âmbito da indústria de óleo e gás, a metodologia proposta permite que as empresas reduzam significativamente o tempo necessário para otimizar a produção, mantendo boa precisão nos resultados.

Palavras-Chave: aprendizado de máquina, estimadores mais rápidos de função objetivo, otimização de estratégias de produção, reservatórios carbonáticos, simulação numérica

ABSTRACT

Oil reservoirs with carbonate lithology are characterized by complex heterogeneities, which must be adequately represented in simulation models that aid decision-making. As a result, the computational time required to run simulation models to estimate production can be very large. During a probabilistic optimization process of production strategies, the total time is even greater, so in some cases, it may be necessary to run thousands of simulations to obtain good results. Due to the high computational demand and its significant impact on decision analysis time, integrated methods of simulation models and artificial intelligence (AI) techniques are investigated to reduce simulator dependency in the production optimization process. The study adopts a hierarchical approach, divided into two parts: (1) it involves integrating simulation models with the iterative discrete Latin hypercube (IDLHC) optimization algorithm to develop a nominal production strategy, optimizing design variables to obtain the best operational configuration and maximum economic return; (2) AI techniques are applied to develop a faster objective function estimator (FOFE) integrated with IDLHC to be used in the most costly parts of the optimization process. The results confirm that FOFE is very useful in certain phases of the optimization process and should be evaluated with caution in others, offering the decision-maker the option of choosing between using only flow models, FOFE, or both. In addition, FOFE is of great value in the stage with the highest computational cost, which is determining the position of the wells. As the optimization becomes more refined and the need for precision increases, the representativeness of the FOFE concerning the simulator tends to reduce, remaining within a margin of error typical of the optimization process, thus becoming less effective. In short, after optimization with FOFE, the process was accelerated by more than 50%, with a reduction of less than 1% in the quality of the results for the simulation models (lower than the expected error of the models compared to the real case). Within the oil and gas industry, the proposed methodology allows companies to significantly reduce the time needed to optimize production while maintaining good accuracy in the results.

Key Word: carbonate reservoir, faster objective function estimators, machine learning, numerical simulation, production strategy optimization

LISTA DE ILUSTRAÇÕES

Figura 1: 12 passos da metodologia de gerenciamento e desenvolvimento de campos em malha fechada.....	22
Figura 2: Desenvolvimento e Gerenciamento do Campo em Malha Fechada.	26
Figura 3: Variação da posição do poço para cada etapa de simulação.....	28
Figura 4: (a) Espaçamento dos poços produtores no método por varredura, (b) Distribuição dos poços produtores fixos, (c) Distribuição dos poços produtores e injetores fixos.....	28
Figura 5: Fluxograma da metodologia do IDLHC.	33
Figura 6: Representação esquemática do método de Floresta Aleatória, onde múltiplas árvores de decisão são combinadas para gerar uma previsão final por meio de agregação de previsões (votação por maioria ou média).....	36
Figura 7: Arquitetura do <i>Gradient Boosting Regressor</i> , onde múltiplas árvores de decisão (classificadores fracos) são treinadas sequencialmente em subconjuntos ponderados dos dados, com cada árvore corrigindo os erros residuais das anteriores. A previsão final é obtida pela combinação das previsões de todas as árvores, resultando em um classificador forte.....	37
Figura 8: Estrutura de uma RNA com uma camada oculta, ilustrando as conexões entre a camada de entrada, a camada oculta e a camada de saída.	39
Figura 9: Etapas de como o Gradiente Conjugado Escalonado é aplicado na Rede Neural. ...	40
Figura 10: Fluxograma da metodologia proposta na Parte 1 do trabalho.....	54
Figura 11: Fluxograma da metodologia proposta na Parte 2 do trabalho.....	57
Figura 12: Definição do número e posicionamento dos poços produtores para construção do mapa de qualidade.	63
Figura 13: Vista superior do reservatório destacando o espaço de busca avaliado para cada novo poço adicionado à estratégia.....	65
Figura 14: Corte bidimensional (XZ) do reservatório mostrando os intervalos de completção avaliados para não serem completados no ajuste fino para os poços produtores (100 m) e para os injetores (50 m). O eixo Z indica a posição em profundidade do reservatório, enquanto o eixo X ilustra a posição horizontal do modelo de fluxo.	65
Figura 15: Configuração da rede neural utilizada para realizar a estimativa da função objetivo.	69

Figura 16: Histograma da distribuição das estimativas do VPL das simulações para o cenário que gerou 1.000 amostras a cada iteração da otimização indicando a estratégia adotada a fim de evitar vies na seleção das amostras.	72
Figura 17: Definição das posições candidatas no reservatório.....	74
Figura 18: Curvas de vazão de produção total de líquidos, óleo e água dos quatro poços produtores já perfurados no campo, indicando o início do declínio da produção pela linha tracejada.....	74
Figura 19: Variação do IEC em função do número de poços ao longo de 14 iterações do processo de otimização. A estratégia de produção inicial (EPi) é representada por um triângulo vermelho invertido, definido como caso base. As duas estratégias com melhores indicadores econômicos (EPO1-1 e EPO1-2), alcançadas na primeira etapa do processo hierárquico, são destacadas por losangos preenchidos em preto e azul contornados em vermelho, respectivamente.....	75
Figura 20: Curvas de vazão de produção de líquidos, óleo e água das duas estratégias que retornaram o melhor IEC da otimização do número e posição dos poços na P1-F1.....	75
Figura 21: Comparação entre a distribuição dos poços na EPi e na EPO1.	76
Figura 22: Evolução do IEC ao longo das simulações nas diferentes iterações, com a melhor estratégia (EPO1) identificada na 6ª iteração.	77
Figura 23: Critério de parada indicando o momento de interrupção do processo.....	78
Figura 24: Indicação do deslocamento dos poços após o ajuste fino.	79
Figura 25: Comparação da saturação de água na região do poço antes e após o ajuste fino na posição dos poços adicionados à estratégia.	80
Figura 26: Comparação entre a EPO1 (linha contínua) e EPO2 (linha tracejada) indicando que o ajuste fino atuou de forma positiva no gerenciamento da produção de líquidos, promovendo uma redução na produção de água e maximização na produção de óleo.	80
Figura 27: Evolução do IEC ao longo das simulações nas diferentes iterações, com a melhor estratégia identificada na nona iteração.....	81
Figura 28: Evolução do IEC ao longo das iterações indicando o momento em que o processo deve ser interrompido para evitar realizar simulações desnecessárias.	81
Figura 29: Evolução da otimização do cronograma de abertura dos poços, comparando os resultados com a EPi, EPO1 e EPO2.....	83
Figura 30: Evolução do VPL ao longo das iterações do processo de otimização, destacando o momento de convergência e o fim da otimização nas iterações iniciais.	83

Figura 31: Comparação entre as métricas R^2 e SMAPE do treinamento, validação e teste nos três métodos de aprendizado de máquina avaliados, para identificar o que melhor captura o comportamento dos dados e realiza estimativas mais precisas.	86
Figura 32: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do número e posição dos poços.	87
Figura 33: Critério de parada indicando o momento de parada do processo de otimização do número e posição dos poços (F2-P1).....	88
Figura 34: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução do ajuste fino na posição e completção dos poços.	89
Figura 35: Evolução do valor da função objetivo ao longo da otimização, indicando o momento de interrupção do ajuste fino na posição e completção dos poços após encontrar a melhor resposta e não obter melhora nas três iterações subsequentes.....	90
Figura 36: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do cronograma de abertura dos poços, comparando-as com a Parte 1 da metodologia.	91
Figura 37: Evolução do valor da função objetivo ao longo da otimização, indicando o momento de interrupção do processo após violar o critério de parada na otimização do cronograma de abertura dos poços.	92
Figura 38: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do número e posição dos poços na P2-F1.....	93
Figura 39: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução do ajuste fino na posição e completção dos poços na P2-F2.	94
Figura 40: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização no cronograma de abertura dos poços na P2-F3.	95

LISTA DE TABELAS

Tabela 1: Classificação das variáveis de decisão.	30
Tabela 2: Principais parâmetros utilizados no treinamento de redes neurais aplicando a função <i>trainscg</i>	42
Tabela 3: Definições das premissas impostas para o Campo C.	61
Tabela 4: Parametrização das variáveis otimizadas.	62
Tabela 5: Valores fixos das variáveis do G2L.....	62
Tabela 6: Parametrização do algoritmo de IDLHC.....	63
Tabela 7: Intervalos de abertura avaliados para os poços injetores e produtores.....	66
Tabela 8: Descrição dos dados de entrada e de saída (FO) utilizados pelos algoritmos de aprendizado de máquina em cada etapa do processo hierárquico (Etapas 4.3, 4.4 e 4.5 do fluxograma apresentado na Figura 10).	67
Tabela 9: Modelo econômico do estudo de caso do Campo C.....	68
Tabela 10: Seleção dos hiperparâmetros utilizados na busca da melhor configuração para os algoritmos testados, baseado na literatura.	68
Tabela 11: Diferença do IEC e indicadores de produção (N_p , W_p e FR_o) entre a estratégia com nove e oito poços da P1-F1.	76
Tabela 12: Evolução das funções de objetivo ao longo das fases hierárquicas do processo de otimização comparadas com o caso base (E_{Pi}).....	84
Tabela 13: Comparação do tempo computacional em relação às execuções de simulação (realizadas no simulador mais as equivalentes), destacando o aumento de velocidade obtido com o uso do FOFE.	96
Tabela 14: Análise da qualidade do resultado comparando a otimização utilizando apenas RSM com a otimização utilizando RSM e FOFE gerando 1000 amostras (destacado na cor verde) e gerando 100 amostras (destacado na cor azul) nas três etapas do processo hierárquico.....	97

LISTA DE NOMENCLATURAS

Abreviaturas

Epi - Estratégia de produção inicial

EPO1 - Estratégia de produção otimizada na fase 1 do processo hierárquico

EPO2 - Estratégia produção otimizada na fase 2 do processo hierárquico

EPO3 - Estratégia de produção otimizada na fase 3 do processo hierárquico

FOFE 1000 - Estimador mais rápido de função objetivo que gera 1000 amostras por iteração

FOFE 100 - Estimador mais rápido de função objetivo que gera 100 amostras por iteração

G1 - Grupo de variáveis projeto

G2 - Grupo de variáveis de controle

G2L - Grupo de variáveis de controle para ciclo de vida

G2S - Grupo de variáveis a curto prazo

G3 - Grupo de variáveis futuras do projeto

Gp - Produção acumulada de gás (*Cumulative Gas Production*)

K US\$ - Mil dólares

MM US\$ - Milhões de dólares

Np - Produção acumulada de petróleo (*Cumulative Oil Production*)

nw - Número Máximo de Conexões de Poços

P1 - Parte 1 da metodologia

P2 - Parte 2 da metodologia

P1-F1 - Fase hierárquica 1 da parte 1 da metodologia

P1-F2 - Fase hierárquica 2 da parte 1 da metodologia

P1-F3 - Fase hierárquica 3 da parte 1 da metodologia

P2-F1 - Fase hierárquica 1 da parte 2 da metodologia

P2-F2 - Fase hierárquica 2 da parte 2 da metodologia

P2-F3 - Fase hierárquica 3 da parte 2 da metodologia

Wp - Produção acumulada de água (*Cumulative Water Production*)

Siglas

2P2K - Dupla Porosidade e Dupla Permeabilidade

BHP - Pressão de Fundo do Poço (*Bottom Hole Pressure*)

CiW - Capacidade de Injeção de Água

CLFDM - Desenvolvimento e Gerenciamento de Campo em Malha Fechada (*Closed Loop Field Development and Management*)

CMG - Computer Modelling Group

CNN - Rede Neural Convolutacional (*Convolutional Neural Network*)

CpG - Capacidade de Processamento de Gás

CpL - Capacidade de Processamento de Líquidos

CpO - Capacidade de Processamento de Óleo

CpW - Capacidade de Processamento de Água

cw - Poços Candidatos (*Candidate Wells*)

DECE - Deterministic Evolutionary Control of Exploration

DLHC - Hipercubo Latino Discreto (*Discrete Latin Hypercube*)

EDP - Equações Diferenciais Parciais

EDFM - *Embedded Discrete Fracture Model*

ENET - *Elastic Net*

EO - Enhanced Oil Recovery (Recuperação Melhorada de Óleo)

EP - Estratégia de Produção

FCL - Fluxo de Caixa Livre (Free Cash Flow)

FO - Função Objetivo

FOFE - Estimadores de Função Objetivo mais Rápidos (*Faster Objective Function Estimator*)

FRo - Fator de Recuperação de Óleo

GBR - Gradient Boosting Regressor

IA - Inteligência Artificial

IDLHC - Hipercubo Latino Discreto Iterativo (*Iterative Discrete Latin Hypercube*)

IEC - Indicador Econômico de Campo

KNN – *K-Nearest Neighbors*

KRR - Regressão de Kernel Ridge (*Kernel Ridge Regression*)

LHC - Hipercubo Latino (*Latin Hypercube*)

LSTM - Memória de Curto e Longo Prazo (*Long Short Term Memory*)

MLP - Perceptron Multicamada (*Multi-Layer Perceptron*)

PSO - Otimização por Enxame de Partículas (*Particle Swarm Optimization*)

PSOGA - Otimização por Enxame de Partículas com Algoritmo Genético (*Particle Swarm Optimization Genetic Algorithm*)

PVT - Pressão, Volume, Temperatura

RBF - Função de Base Radial (*Radial Basis Function*)

RSM - Modelos de Simulação de Reservatórios (*Reservoir Simulation Models*)

SCG - Gradiente Conjugado Escalar (*Scaled Conjugate Gradient*)

SMAPE - Erro Percentual Absoluto Médio Simétrico (*Symmetric Mean Absolute Percentage Error*)

SSA - Algoritmo de busca Sparrow (*Sparrow Search Algorithm*)

SVR - Máquinas de Vetores de Suporte (*Support Vector Regression*)

TPSO - Otimização de enxame de partículas com restrição de tempo (*Time constrained Particle Swarm Optimization*)

VPL - Valor Presente Líquido

wos - Sequência de Abertura dos Poços (*well opening schedule*)

wst - Intervalo de Abertura dos Poços (*well sequence time*)

SUMÁRIO

1	INTRODUÇÃO.....	19
1.1	Motivação	21
1.2	Objetivos.....	21
1.3	Premissas	22
1.4	Estrutura da dissertação	23
2	FUNDAMENTAÇÃO TEÓRICA	24
2.1	Simulação numérica de reservatórios.....	24
2.2	Otimização de estratégias de produção.....	25
2.3	Mapa de qualidade	27
2.4	Variáveis de otimização.....	29
2.5	Algoritmos de otimização	30
2.5.1	Particle Swarm Optimization (PSO).....	31
2.5.2	Designed Exploration and Controlled Evolution (DECE).....	31
2.5.3	Hipercubo Latino Discreto Iterativo (IDLHC)	32
2.6	Otimização de estratégias de produção utilizando Inteligência Artificial	34
2.7	Conceitos econômicos	42
2.7.1	Valor Presente Líquido (VPL).....	42
2.7.2	Indicador Econômico de Campo (IEC)	43
3	REVISÃO BIBLIOGRÁFICA	44
3.1	Otimização de estratégia de produção em reservatórios carbonáticos fraturados	44
3.2	Otimização de EP utilizando RSM e algoritmos de otimização	45
3.3	Inteligência artificial na otimização de estratégias de produção	48
3.4	Otimização de EP utilizando RSM, FOFE e algoritmos de otimização	48
4	METODOLOGIA.....	54
4.1	Parte 1	54
	Etapa 1: Caso base	55

Etapa 2: Parametrização.....	55
Etapa 3: Seleção do algoritmo de otimização	55
Etapa 4: Otimização da produção	55
4.2 Parte 2.....	57
4.2.1 Dados.....	58
4.2.2 Seleção do FOFE.....	58
4.2.3 Critérios para seleção do FOFE.....	58
4.2.4 Algoritmo de otimização	59
4.2.5 Geração de novas amostras.....	59
4.2.6 Selecionar amostras para validar no simulador e alimentar o banco de dados para atualizar o FOFE.....	59
4.2.7 Análise de conclusão do processo hierárquico	59
5 APLICAÇÕES	60
5.1 Parte 1 – Otimização utilizando RSM.....	61
5.1.1 Premissas estabelecidas para o Campo C	61
5.1.2 Parametrização das variáveis.....	61
5.1.3 Seleção e parametrização do algoritmo de otimização	62
5.1.4 Construção do mapa de qualidade	63
5.1.5 Definição das posições candidatas e completação dos poços produtores e injetores.	64
5.1.6 Ajuste fino na posição e completação dos poços.....	64
5.1.7 Configuração do cronograma de abertura dos poços	66
5.2 Parte 2 – FOFE + IDLHC	66
5.2.1 Seleção do conjunto de dados inicial para realizar o treinamento, validação e teste das técnicas de aprendizado de máquina propostas	66
5.2.2 Seleção do método de aprendizado de máquina	67
5.2.3 Critério de avaliação do algoritmo de previsão	70
5.2.4 Seleção e parametrização do algoritmo de otimização	70
5.2.5 Geração de novas amostras.....	71
5.2.6 Seleção de amostras para validação da FO prevista pelo FOFE no simulador.....	71
5.2.7 Critério de parada da otimização	71
5.2.8 Fases da otimização.....	72

6	RESULTADOS E DISCUSSÕES.....	73
6.1	Parte 1 - Otimização com RSM e IDLHC	73
6.1.1	Otimização da posição e número dos poços	73
6.1.2	Ajuste fino da posição dos poços.....	79
6.1.3	Otimização do cronograma de abertura dos poços	82
6.2	Parte 2 - Otimização com FOFE, RSM e IDLHC.....	85
6.2.1	Seleção do algoritmo de aprendizado de máquina.....	85
6.2.2	Otimização da posição e número dos poços FOFE – 1000 amostras	87
6.2.3	Ajuste fino na posição dos poços FOFE – 1000 amostras.....	89
6.2.4	Otimização do cronograma de abertura dos poços FOFE – 1000 amostras.....	90
6.2.5	Otimização da posição e número dos poços FOFE – 100 amostras	92
6.2.6	Ajuste fino na posição dos poços FOFE – 100 amostras.....	93
6.2.7	Otimização do cronograma de abertura dos poços FOFE – 100 amostras.....	94
6.3	Tempo computacional	95
7	CONCLUSÕES.....	98
	REFERÊNCIAS	101

1 INTRODUÇÃO

A seleção de estratégias de produção para campos petrolíferos é uma atividade bastante desafiadora, especialmente em ambientes complexos, como reservatórios carbonáticos com porosidade e permeabilidade dupla (Presho; Wo; Ginting, 2011; Rueda; Meije; Roehl, 2019; Xu; Chen; Leung, 2023). Nesses casos, o uso de modelos de simulação é fundamental não apenas para prever o potencial produtivo, mas também para otimizar estratégias de produção visando assim maximizar o retorno econômico do projeto (Li; Zhao; Li, 2016; Zheng, 2023). No entanto, a simulação desses modelos pode ser extremamente demorada, muitas vezes levando dias para realizar todas as simulações necessárias no processo de otimização (Bertini Junior *et al.*, 2019; Silva; Avansi; Schiozer, 2020). Isso pode causar atrasos operacionais, aumentando custos e riscos associados a incertezas durante o ciclo de vida dos projetos de exploração e produção (Santos *et al.*, 2017; Schiozer *et al.*, 2004; Suslick; Schiozer, 2004). Ao mesmo tempo, a indústria de petróleo e gás enfrenta desafios econômicos e operacionais que tornam a velocidade e a confiabilidade na tomada de decisão, fatores críticos para o sucesso. Diante desse cenário, técnicas que reduzam a dependência de simuladores tornam-se fundamentais para acelerar o processo de tomada de decisões e influenciar diretamente na viabilidade econômica dos projetos.

O processo de otimização em reservatórios carbonáticos, que integra os sistemas de matriz e fratura (2P2K), requer a construção de modelos geológicos mais fidedignos para representar os caminhos do fluxo de fluidos. Devido ao aumento no número de Equações Diferenciais Parciais (EDP) necessárias para representar esses sistemas, os modelos de simulação tornam-se complexos e computacionalmente custosos. A heterogeneidade dos reservatórios, caracterizada pela variação de porosidade, permeabilidade, e a distribuição complexa de poros (Mazzullo; Rieke; Chilingarian, 1996; Lucia, 2007; Guo *et al.*, 2019), intensifica a necessidade de uma caracterização mais precisa. Paiva (2012) destaca que em reservatórios naturalmente fraturados, as descontinuidades e diferenças de condutividade hidráulica e efeitos capilares entre matriz e fratura requerem o uso de funções de transferência para modelar a interação entre esses componentes heterogêneos, influenciando diretamente a eficiência das estratégias de recuperação.

O gerenciamento do campo tem sido conduzido como um processo contínuo, que busca otimizar a interação entre os dados e a tomada de decisão durante o ciclo de vida do campo (Fanchi, 2018). Entretanto, a alta demanda computacional exigida pelos modelos de

reservatórios carbonáticos 2P2K resulta em um aumento significativo no tempo de simulação e na quantidade de cenários que precisam ser analisados para otimizar as estratégias de produção. Em uma indústria onde decisões rápidas e precisas são extremamente importantes, esses atrasos podem comprometer a eficácia das operações e a rentabilidade dos investimentos. Esse desafio reforça a necessidade de obter previsões mais rápidas, destacando a importância de desenvolver soluções que possam não somente acelerar a simulação, mas manter uma qualidade aceitável dos resultados.

Os modelos de simulação de reservatórios (do inglês, *reservoir simulation models* (RSM)) têm sido extensivamente utilizados para avaliar o comportamento do reservatório sob diferentes condições físicas, operacionais e econômicas (Odeh, 1969; Stags; Herbeck, 1971). Essas ferramentas são essenciais para auxiliar no processo de previsão da produção e otimização de estratégias de produção, proporcionando uma análise mais detalhada do potencial produtivo do reservatório sob diferentes cenários. Devido à sua ampla aplicação na indústria, essa técnica se consolidou como uma das mais confiáveis para tal propósito em cenários complexos (Mattax; Dalton, 1990). Entretanto, durante a aplicação dos RSM na otimização de estratégias de produção, é importante considerar diferentes estratégias, a fim de cobrir um amplo espaço de busca, que por sua vez, podem envolver um elevado número de variáveis e incerteza (Botechia, 2012; Santos; Gaspar; Schiozer, 2018) s. Essa combinação resulta em uma elevada quantidade de modelos que, para serem simulados, demandam um tempo computacional significativo (Nasir; Yu; Sepehrnoori, 2020; Sarma *et al.*, 2006; Silva *et al.*, 2020; Silva *et al.*, 2023).

É importante destacar as metodologias de otimização aplicadas aos modelos de simulação de reservatórios. Métodos baseados em gradientes e algoritmos evolutivos, como os genéticos e o IDLHC, têm sido amplamente utilizados para otimizar a quantidade e o posicionamento de poços, maximizando o VPL com uma quantidade significativamente menor de simulações (Maschio, Nakajima e Schiozer, 2008; Nogueira e Schiozer, 2009; Hohendorff Filho, Maschio e Schiozer, 2016). Essas metodologias são fundamentais para lidar com a complexidade dos cenários que os reservatórios estão inseridos, permitindo explorar de maneira eficiente um amplo espaço de busca. Entretanto, apesar de suas vantagens, o alto custo computacional ainda é um obstáculo significativo, especialmente em reservatórios com alta heterogeneidade, como os carbonáticos. Deste modo, há uma crescente necessidade de integrar essas abordagens com técnicas avançadas.

Diante dos desafios impostos pela alta demanda computacional durante a otimização da produção utilizando RSM, torna-se essencial a busca por soluções que melhorem a eficiência

do processo mantendo a qualidade dos resultados. Nesse contexto, a aplicação de estimadores mais rápidos de função objetivo (do inglês *faster objective function estimator* (FOFE)) emerge como uma abordagem promissora na otimização de estratégias de produção. Os FOFEs têm sido utilizados, principalmente para estimar resultados de forma mais rápida através de modelos matemáticos analíticos, baseados na física ou híbridos (Avansi; Hohendorff Filho; Schiozer, 2020; Maschio; Avansi; Schiozer, 2023). Dentre a classificação dos FOFEs, o que se baseia em técnicas de inteligência artificial (IA), têm sido aplicadas com êxito para acelerar a obtenção de resultados (Cirac *et al.*, 2023a, b; Ertekin; Sun, 2019; Evensen; Womack, 2020; Sircar *et al.*, 2021) e é o objetivo de estudo deste trabalho.

Este estudo destaca a importância prática de novas abordagens para otimizar estratégias de produção na indústria de petróleo, propondo uma técnica que pode acelerar significativamente o processo de simulação e manter a precisão das soluções, resultando em decisões mais rápidas e eficientes ao incorporar técnicas avançadas com FOFE e inteligência artificial. Isso é importante para um setor onde as oportunidades de investimento e operações são frequentemente limitadas e onde decisões rápidas e precisas podem resultar em economias substanciais e aumentar a eficácia operacional.

1.1 Motivação

A previsão de produção de reservatórios é uma tarefa desafiadora no gerenciamento de campos de petróleo. Reservatórios complexos, como os carbonáticos 2P2K, requerem a implementação de modelos de simulação mais eficientes para tornar o processo viável. A simulação desses modelos é essencial, mas demanda um tempo computacional proibitivo (da ordem de dias) para executar a grande quantidade de simulações envolvidas na otimização. Por outro lado, os desafios econômicos e operacionais enfrentados pela indústria de óleo e gás colocam-na em um ambiente onde a velocidade e a confiabilidade na tomada de decisão tornam-se fundamentais para o sucesso do projeto. Nesse contexto, o desenvolvimento de técnicas que reduzam a dependência dos simuladores durante todas as etapas para otimizar uma estratégia de produção se torna crucial para tomar decisões mais rápidas que impactam na viabilidade econômica de um projeto.

1.2 Objetivos

O objetivo geral deste trabalho é desenvolver uma metodologia para acelerar o processo de otimização da estratégia de produção de reservatórios carbonáticos. Para atingir este objetivo

geral, o trabalho é dividido em dois objetivos específicos para uma otimização nominal de estratégia de produção.

Objetivos específicos:

1. Otimizar nominalmente uma estratégia de produção, por meio da aplicação de RSM, estabelecendo um caso de referência para avaliações futuras;
2. Acelerar o processo de otimização da estratégia de produção através da aplicação de FOFE, baseadas em técnicas de aprendizado de máquina.

1.3 Premissas

Este trabalho segue a metodologia de Schiozer *et al.* (2019) para o gerenciamento de campos de petróleo, que se baseia em 12 etapas de trabalho, conforme destacado na Figura 1. A otimização de estratégia e produção pode ser executada nas etapas 6 e 9 da metodologia. Entretanto, este trabalho é focado em otimizar uma estratégia de produção apenas para o caso base, caracterizado pelo passo 6.

1	Caracterização do reservatório
2	Construção do modelo
3	Calibração do modelo
4	Geração dos cenários
5	Redução dos cenários
6	Estratégia de produção – caso base
7	Análise de risco inicial
8	Modelos representativos
9	Estratégia de Produção para cada modelo representativo
10	Otimização robusta da produção
11	Decisão no ciclo de vida / Gerenciamento de ativos de infraestrutura
12	Análise de decisão / curto prazo

Figura 1: 12 passos da metodologia de gerenciamento e desenvolvimento de campos em malha fechada.

Fonte: Schiozer *et al.* (2019).

1.4 Estrutura da dissertação

Esta dissertação foi dividida em sete capítulos, conforme descritos a seguir:

Os principais conceitos e definições que circundam o tema de pesquisa são apresentados no Capítulo 2, tais como: otimização de estratégias de produção, algoritmos de otimização e FOFE.

Uma revisão da literatura é apresentada no Capítulo 3, destacando os trabalhos mais relevantes sobre otimização de estratégias de produção.

A metodologia do trabalho proposto é detalhada no Capítulo 4, descrevendo as etapas essenciais a serem seguidas.

No Capítulo 5, são detalhados o estudo de caso e a aplicação da metodologia proposta, incluindo os dados e premissas utilizados.

O Capítulo 6 apresenta os resultados obtidos com a aplicação da metodologia proposta, seguidos de uma análise e discussão desses resultados.

E, por fim, as conclusões obtidas neste trabalho são apresentadas no Capítulo 7, além de considerações e sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Simulação numérica de reservatórios

A simulação numérica de reservatórios é uma técnica que utiliza programas computacionais para resolver equações matemáticas que descrevem processos dinâmicos de fluxo de fluidos no meio poroso (Mattax; Dalton, 1990; Odeh, 1969). Esta técnica gera uma representação matemática computacional do reservatório para prever seu comportamento dinâmico. Esse modelo é geralmente estruturado em uma malha, onde as propriedades das rochas (como porosidade, saturação e permeabilidade) e as propriedades dos fluidos (como viscosidade e PVT) são especificadas para cada bloco. A quantidade e o tamanho dos blocos na malha definem a precisão com que as propriedades do reservatório serão representadas. Esse refinamento depende dos objetivos da simulação, e a configuração determinada tem forte impacto na velocidade de simulação do modelo (Mattax; Dalton, 1990; Weber; Geuns, 1990).

A simulação de reservatórios opera com base nos princípios de equilíbrio das três principais forças que atuam nas partículas fluidas (viscosas, gravitacionais e capilares), calculando o fluxo de fluido de um bloco para o próximo, de acordo com a lei de Darcy. A força que governa o fluxo de fluido é a diferença de pressão entre blocos da malha adjacente. O cálculo do fluxo é repetido em intervalos curtos de tempo, e ao final de cada intervalo, a nova saturação e pressão do fluido são calculadas para cada bloco da malha (Ertekin; Abou-Kassem; King, 2001).

Modelos de simulação de reservatórios

A construção de modelos de simulação de reservatório, pode ser realizada de acordo com as definições apresentadas nos trabalhos de Mattax e Dalton (1990) e Odeh (1969), que envolvem uma sequência de etapas:

- Antes de iniciar a construção de um modelo, é essencial definir os objetivos do estudo de simulação, que podem incluir a previsão futura da produção, otimização da recuperação ou estudo de diferentes estratégias de desenvolvimento;
- Coletar e preparar os dados geológicos e de produção, como propriedades das rochas (porosidade, permeabilidade), propriedades dos fluidos (viscosidade, pressão e temperatura), e histórico de produção, que são coletados e processados para serem usados no modelo;

- Configuração do reservatório, que é dividida em uma malha tridimensional composta por blocos menores, sendo que cada bloco é atribuído a propriedades que representam o reservatório;
- Definição de parâmetros e condições iniciais onde as propriedades da rocha e dos fluidos, bem como as condições iniciais de pressão e saturação, são atribuídas a cada bloco da malha;
- Seleção do método numérico para resolver as equações diferenciais parciais (por exemplo, o método de diferenças finitas), que é uma parte fundamental da construção do modelo;
- Validação do modelo, comparando os resultados simulados com dados históricos de produção, e ajustando-o até que os resultados simulados correspondam aos dados históricos;
- Aplicação do modelo na otimização de estratégias de produção.

Modelos de simulação de reservatórios naturalmente fraturados

Modelos de simulação de reservatórios naturalmente fraturados são ferramentas utilizadas na indústria de óleo e gás para prever o comportamento dos fluidos em formações geológicas que apresentam fraturas naturais. Esses modelos complexos utilizam conceitos de dupla porosidade e dupla permeabilidade para representar a interação entre a matriz rochosa, que possui porosidade e permeabilidade mais baixas, e as fraturas, que oferecem caminhos de alta permeabilidade, mas baixa capacidade de armazenamento (Nelson, 2001; Tiab; Donaldson, 2012). As simulações numéricas desses modelos são fundamentais para entender como o óleo, a água e o gás se movem através desses reservatórios heterogêneos, empregando funções de transferência que modelam a migração de fluidos entre a matriz e as fraturas (Paiva, 2012). Essa modelagem permite que os engenheiros otimizem métodos de recuperação de petróleo, ajustem estratégias de produção e melhorem as previsões de desempenho do reservatório, considerando a dinâmica das fraturas, que impactam significativamente a eficiência da produção e recuperação (Gugl *et al.*, 2022).

2.2 Otimização de estratégias de produção

A otimização de estratégias de produção envolve a utilização de ferramentas e técnicas para identificar os métodos mais eficazes de exploração do reservatório. Os principais objetivos incluem minimizar os custos, mitigar os riscos e maximizar as receitas (Rosa; Carvalho; Xavier,

2006). Entretanto, as empresas podem indicar outros objetivos (Jahn; Cook; Graham, 2008). Com a aplicação dos RSM é possível analisar diferentes estratégias de produção a fim de identificar aquelas que alcancem os propósitos estabelecidos pela companhia de óleo e gás.

Há diferentes abordagens para otimizar as variáveis envolvidas no desenvolvimento e gerenciamento de um campo (por exemplo as apresentadas na Tabela 1), que vão desde as fundamentadas em modelos que representam a física dos reservatórios, até métodos que são orientados por dados (Ertekin; Abou-Kassem; King, 2001) como, por exemplo, aprendizado de máquina.

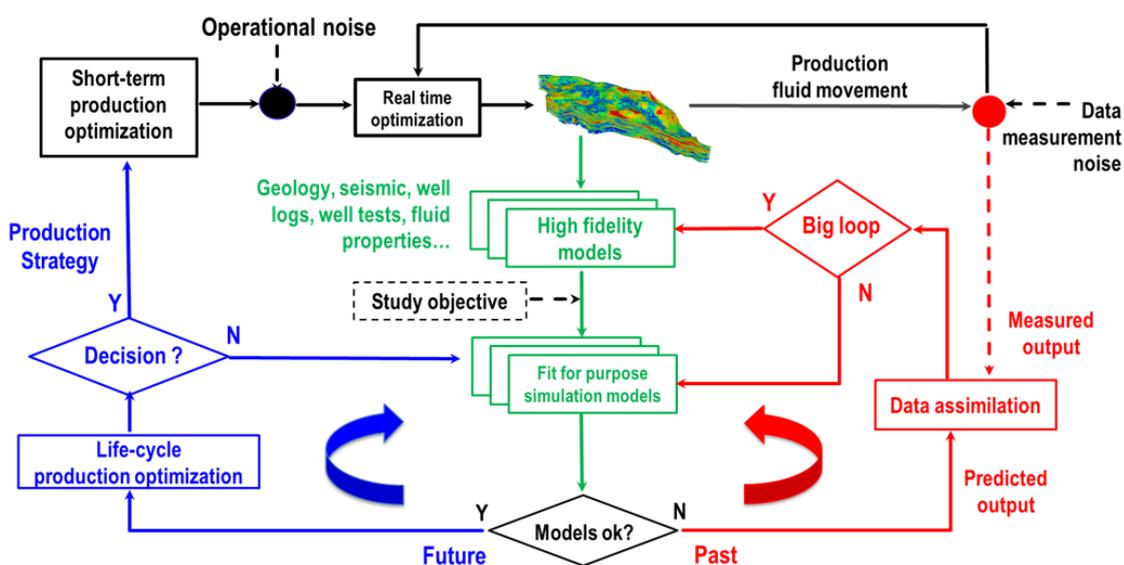


Figura 2: Desenvolvimento e Gerenciamento do Campo em Malha Fechada.

Fonte: Schiozer *et al.* (2019).

- Na fase verde, são coletados os dados e atribuídas as incertezas para construir o modelo. Nesta etapa, também é adaptado o grau de fidelidade do modelo (baixa, média ou alta) para equilibrar a qualidade dos resultados com o tempo computacional, de acordo com o propósito do estudo;
- Na fase vermelha, é realizada a assimilação de dados para redução de incertezas (ajuste de histórico). Todos os dados dinâmicos devem estar dentro de uma faixa de tolerância para selecionar os modelos que serão utilizados na fase azul da metodologia. Nesta fase, os modelos de simulação ou os modelos geológicos de alta fidelidade podem ser diretamente modificados (*big loop*);
- As decisões do ciclo de vida baseado em modelos incertos são realizadas na fase azul, onde também é definida a melhor estratégia de produção para o campo. Este trabalho está inserido nessa fase;

- Na fase preta, ocorre a implementação de decisões de longo prazo (normalmente baseadas em modelos) e decisões de curto prazo (normalmente baseadas em dados); definição do objetivo do estudo; e seleção do tipo de estudo (passado – assimilação de dados ou futura – análise de decisão).

Além da divisão em quatro cores, a metodologia é subdividida em 12 passos, de modo que cada um está representado na cor da fase a que pertence, conforme mostrado na Figura 1.

2.3 Mapa de qualidade

O mapa de qualidade, segundo Cruz (2000), é uma representação bidimensional do comportamento do reservatório, considerando incertezas. É uma ferramenta capaz de representar regiões com diferentes potenciais produtivos. Baseado nesse conceito, o mapa de qualidade auxilia na determinação de regiões para perfurar os poços, reduzindo o espaço de busca no processo de otimização do posicionamento e número de poços e, conseqüentemente, o esforço computacional. Cruz (2000) categorizou como possíveis aplicações para os mapas de qualidade:

- Localização de poços;
- Otimização da abordagem completa para determinar o melhor número de poços;
- Identificação de um cenário representativo;
- Classificação de cenários com suas incertezas;
- Caracterização e comparação de reservatórios.

O método de geração do mapa de qualidade proposto por Cruz (2000) baseia-se na execução do simulador de fluxo com apenas um poço produtor vertical, variando sua posição XY em cada execução, com o objetivo de realizar uma ampla varredura no reservatório, conforme mostra a Figura 3. Cada execução avalia, por um determinado tempo, a qualidade da célula horizontal onde o poço está localizado, e essa unidade de qualidade foi definida como produção acumulada de petróleo (N_p). O poço é completado em todas as camadas, definindo um limite para produção de água (ou gás) e um BHP mínimo e uma taxa de óleo mínima de acordo com as limitações esperadas dos poços durante a produção efetiva, para tentar produzir o máximo de cada posição do poço.

Dependendo do tamanho do reservatório, o processo de construção do mapa de qualidade pode se tornar computacionalmente custoso. Cruz (2000) propôs obter apenas alguns pontos para cada realização e, em seguida, interpolar os mapas através da *krigagem*. No entanto,

devido ao alto número de simulações envolvidas, Cavalcante Filho (2005), Martini, Schiozer e Nakajima (2005), Nakajima e Schiozer (2003) e Nakajima (2003) propuseram modificações para reduzir o esforço computacional e acelerar o processo de geração do mapa.

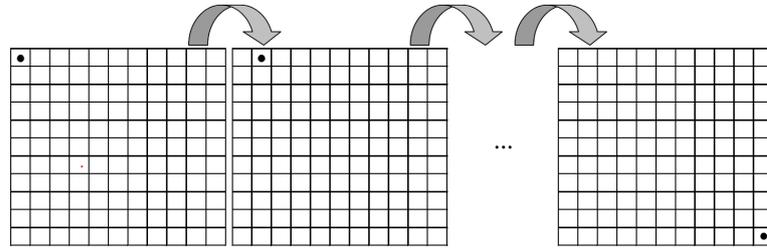


Figura 3: Variação da posição do poço para cada etapa de simulação.

Fonte: Cavalcante Filho (2005).

Uma das modificações propostas por Cavalcante Filho (2005) foi distanciar o número de blocos na configuração do método por varredura, como, por exemplo, de dois em dois, quatro em quatro e assim sucessivamente, conforme ilustrado na Figura 4a. Outra abordagem foi posicionar uniformemente todos os poços produtores fixos ao longo do reservatório e realizar a simulação de forma simultânea, como mostrado na Figura 4b. Além disso, para auxiliar na manutenção da pressão do reservatório, Cavalcante Filho (2005) intercalou poços produtores e injetores fixos, como exemplificado na Figura 4c. Essas modificações visam tornar o processo de geração do mapa de qualidade mais eficiente e menos computacionalmente intensivo.

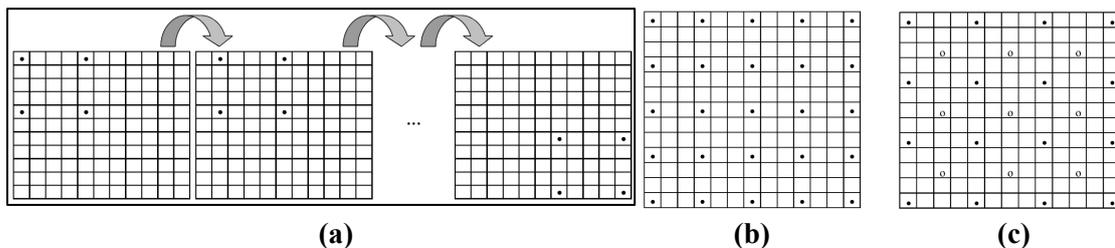


Figura 4: (a) Espaço dos poços produtores no método por varredura, (b) Distribuição dos poços produtores fixos, (c) Distribuição dos poços produtores e injetores fixos.

Fonte: Cavalcante Filho (2005).

Além dos mapas que dependem exclusivamente dos dados de simulação de reservatórios, existem os métodos analíticos. Nakajima (2003) propôs um mapa de qualidade baseado no modelo de Babu e Odeh (1989) que utiliza uma solução analítica para produtividade de poços horizontais. Os métodos analíticos podem incluir modelos matemáticos simplificados, relações empíricas ou até mesmo abordagens baseadas em dados de histórico. O processo de geração desses mapas é mais rápido e pode fornecer uma estimativa inicial útil da distribuição

da qualidade do reservatório, embora possam não ser tão precisos quanto os mapas baseados em simulação.

2.4 Variáveis de otimização

Para tomar decisões eficazes, é necessário entender e categorizar as variáveis envolvidas na otimização da estratégia de produção. Gaspar, Barreto e Schiozer (2016) classificaram as variáveis de otimização em três grupos distintos, de modo que cada um deles desempenha um papel único em diferentes fases de um processo de otimização iterativo:

- Grupo 1 (G1) - As variáveis de projeto representam as especificações da infraestrutura do campo antes do início do desenvolvimento. Ao definir a configuração fundamental do campo, as variáveis G1 influenciam criticamente as perspectivas econômicas e são fundamentais para evitar atrasos no processo de otimização da estratégia de produção.
- Grupo 2 (G2) - As variáveis de controle desempenham um papel fundamental no gerenciamento dinâmico e na otimização estratégica dos processos operacionais do campo durante sua vida útil, a um custo reduzido. Diferentemente das variáveis de projeto, elas não controlam diretamente o equipamento, mas são essenciais para a avaliação do sistema e a previsão de condições futuras. Devido às incertezas intrínsecas de sua operação, essas variáveis influenciam substancialmente o projeto, embora em menor grau em relação às variáveis de projeto.
- Após Schiozer *et al.* (2019) terem descrito a otimização da produção a curto prazo e do ciclo de vida em um fluxo de trabalho geral para aplicações em malha fechada, Mirzaei-Paiaman, Santos e Schiozer (2021) sugeriram a divisão do G2 em duas subclasses: G2L para otimização do ciclo de vida e G2S para otimizações a curto prazo.
- Grupo 3 (G3) - As variáveis futuras do projeto representam opções futuras, e sua integração pode afetar as escolhas relacionadas às variáveis do projeto. As variáveis dessa categoria exigem investimentos adicionais. Devido à imprevisibilidade da previsão, as estratégias de exploração estão sujeitas a mudanças. Portanto, fazer provisões para possíveis evoluções e opções futuras pode ajudar a reduzir o impacto dessa incerteza inevitável.

Seguindo a classificação proposta por Gaspar, Barreto e Schiozer (2016), a Tabela 1 apresenta exemplos de variáveis para cada um dos três grupos definidos no processo de otimização da estratégia.

Tabela 1: Classificação das variáveis de decisão.

Grupo 1 (G1) Variáveis de projeto	Grupo 2 (G2L e G2S) Variáveis de operação e regras de controle	Grupo 3 (G3) Variáveis de Revitalização
Unidade de produção (Capacidade da plataforma)	Operações do poço (controle de <i>chokes</i> e injeção)	<i>Infill drilling</i>
Número, tipo e posição dos poços	Método de rateio	Recompletação
Número e localização das válvulas de controle de fluxo (ICV)	Operação das ICVs	Conversão de poços (de produtor para injetor)
Sequência de perfuração	Ciclos de WAG	Intervenção do poço
Método de recuperação		Estimulação do poço
Cronograma de abertura dos poços		

2.5 Algoritmos de otimização

Os algoritmos de otimização desempenham um papel fundamental na otimização de estratégias de produção de petróleo. São ferramentas computacionais projetadas para encontrar a melhor solução possível para um determinado problema, levando em consideração diferentes variáveis, restrições e objetivos. Existem diversos tipos de algoritmos de otimização, cada um com suas próprias características e aplicações. Alguns dos algoritmos mais comuns incluem:

- **Algoritmos Genéticos:** É um método de otimização estocástica baseado no processo evolutivo da seleção natural e da dinâmica genética. O método procura combinar a sobrevivência do mais apto entre as estruturas de cadeia com uma troca de informações aleatórias, mas estruturadas, para formar uma solução ideal. Inspirados no processo de seleção natural, os algoritmos genéticos operam através de uma combinação de seleção, crossover e mutação para evoluir uma população de soluções em direção ao ótimo global (El-Mihoub *et al.*, 2006);
- **Algoritmos de Enxame:** É um método de otimização estocástica baseado em população, motivado pelo comportamento coletivo inteligente de alguns animais, como pássaros ou formigas. Esses algoritmos envolvem a simulação de múltiplos agentes que interagem entre si para encontrar soluções otimizadas (Wang; Tan; Liu, 2018);
- **Algoritmos de Descida de Gradiente:** É um dos algoritmos mais populares para realizar a otimização e, de longe, a forma mais comum de otimizar redes neurais. Esses algoritmos utilizam a derivada da Função Objetivo (FO) para encontrar a direção de maior declive e, então, ajustam iterativamente os parâmetros na direção que minimiza a FO. Esses algoritmos, no entanto, são frequentemente usados como otimizadores de caixa preta, pois é difícil encontrar explicações práticas sobre seus pontos fortes e fracos (Ruder, 2016);

- Algoritmos de Busca Aleatória: Algoritmo que utiliza algum tipo de aleatoriedade ou probabilidade (normalmente sob a forma de um gerador de números pseudoaleatórios) na definição do método e, na literatura, pode ser designado por método de Monte Carlo, algoritmo estocástico ou metaheurística. Simples e diretos, os algoritmos de busca aleatória exploram o espaço de soluções de forma aleatória até encontrarem uma solução que satisfaça os critérios de otimização (Zabinsky, 2009);
- Algoritmo de Recozimento Simulado: É um algoritmo de otimização inspirado no processo de recozimento térmico dos metais. Ele é eficaz para encontrar soluções aproximadas em problemas com muitos mínimos locais, utilizando uma estratégia de busca que aceita soluções subótimas em certas condições para escapar de mínimos locais (Kirkpatrick; Gelatt; Vecchi, 1983). No contexto da indústria de petróleo é usado para otimização de processos de injeção, planejamento de desenvolvimento de campo, e ajuste de modelos de reservatórios (Azamipour; Assareh, 2017; Beckner; Song, 1995; Yang; Zhang; Gu, 2003).

2.5.1 Particle Swarm Optimization (PSO)

É um algoritmo de otimização simples e eficaz, inspirado no comportamento social de grupos animais, como bandos de aves. Ele combina características de algoritmos evolutivos e programação evolucionária, utilizando processos estocásticos para explorar soluções em um espaço de busca. O PSO se destaca por sua simplicidade e capacidade de otimizar uma ampla variedade de funções, operando de forma similar aos algoritmos genéticos, mas com uma abordagem distinta de "navegar" através do espaço de soluções em direção a melhores respostas, o que permite uma exploração eficiente e robusta do domínio do problema (Kennedy; Eberhart, 1995).

2.5.2 Designed Exploration and Controlled Evolution (DECE)

É um método de otimização desenvolvido pela Computer Modelling Group (CMG), amplamente aplicado em problemas de ajuste histórico e otimização da produção de reservatórios de petróleo. Este método é implementado no software CMOST e segue um processo iterativo, dividido em duas etapas principais. Na primeira etapa, conhecida como exploração projetada, são aplicadas técnicas de design experimental e busca Tabu (técnica de otimização meta-heurística utilizada para resolver problemas complexos de otimização combinatória) para explorar o espaço de soluções de forma abrangente. Em seguida, na etapa

de evolução controlada, realiza-se uma análise estatística dos resultados das simulações, permitindo a identificação e exclusão de valores de parâmetros que não contribuem significativamente para a melhoria das soluções (CMG, 2021).

2.5.3 Hipercubo Latino Discreto Iterativo (IDLHC)

O IDLHC é um algoritmo de otimização desenvolvido por Maschio e Schiozer (2016), que se baseia na aplicação do Hipercubo Latino Discreto (DLHC) em sucessivas iterações. O DLHC, por sua vez, é uma evolução do Hipercubo Latino (LHC) proposto por McKay, Beckman, Conover, (1979), uma técnica de amostragem na qual o espaço de busca é dividido em n intervalos equiprováveis, e cada intervalo contempla uma amostra. Uma das grandes diferenças do DLHC em relação ao LHC é a capacidade de considerar não apenas variáveis aleatórias discretas, mas também variáveis categóricas (Maschio; Schiozer, 2016).

A metodologia do IDLHC proposta por Maschio e Schiozer (2016) inicia com a parametrização dos atributos, discretizando-os em um número específico de níveis para lidar com incertezas categóricas e variáveis codificadas. Em seguida, os modelos iniciais são gerados pelo método DLHC, que proporciona uma amostragem eficiente da distribuição anterior. Após isso, são realizadas simulações de fluxo para cada modelo. O desajuste normalizado entre os resultados simulados e os dados observados é então calculado para avaliar a qualidade da correspondência. Com base nesse desajuste, os modelos são selecionados, eliminando-se as piores combinações e tornando a redução da incerteza dos atributos mais eficiente. Posteriormente, é estimada a distribuição de probabilidade marginal não paramétrica para cada atributo. Este processo é iterativo, retornando à geração de amostras pelo DLHC até que o número definido de iterações seja alcançado, conforme representado na Figura 5. Essa abordagem integrada permite uma busca eficiente e refinada por soluções ótimas em problemas de otimização complexos na otimização de estratégias de produção (Hohendorff Filho; Maschio; Schiozer, 2016).

Maschio e Schiozer (2016) constataram que o IDLHC auxilia na redução de incertezas dos atributos dos reservatórios ao refinar iterativamente o processo de amostragem com base na qualidade de ajuste de histórico. A abordagem demonstrou melhorar a eficiência do ajuste de histórico ao reduzir significativamente a variabilidade da FO (desvio normalizado médio) ao longo de múltiplas iterações. Deste modo, vários autores utilizaram o IDLHC em etapas do desenvolvimento de campos de petróleo. Hohendorff Filho, Maschio e Schiozer (2016) aplicaram o IDLHC para otimizar estratégias de produção no benchmark UNISIM-I-D (Avansi;

Schiozer, 2015), utilizando a metodologia para maximizar a FO VPL em problemas de otimização de estratégia de produção, com um elevado número de variáveis aleatórias discretas do G1, em espaços de busca descontínuos que apresentam muitos ótimos locais. A robustez e eficácia da técnica foi validada através de uma comparação com a ferramenta de otimização DECE. Maschio e Schiozer (2018) utilizaram o IDLHC no ajuste de histórico, combinando o método com a técnica de simulação de recozimento simulado para manter a diversidade das soluções e melhorar a previsão de produção com modelos ajustados. Loomba, Botechia e Schiozer (2022) implementaram o IDLHC em um fluxo de trabalho de otimização robusta, considerando as variáveis do G1 e G2, para melhorar a eficiência do processo de otimização de desenvolvimento de um campo de petróleo complexo e com grandes proporções. A otimização foi aplicada em um modelo composicional do campo inteiro, alcançando resultados promissores em termos de eficiência.

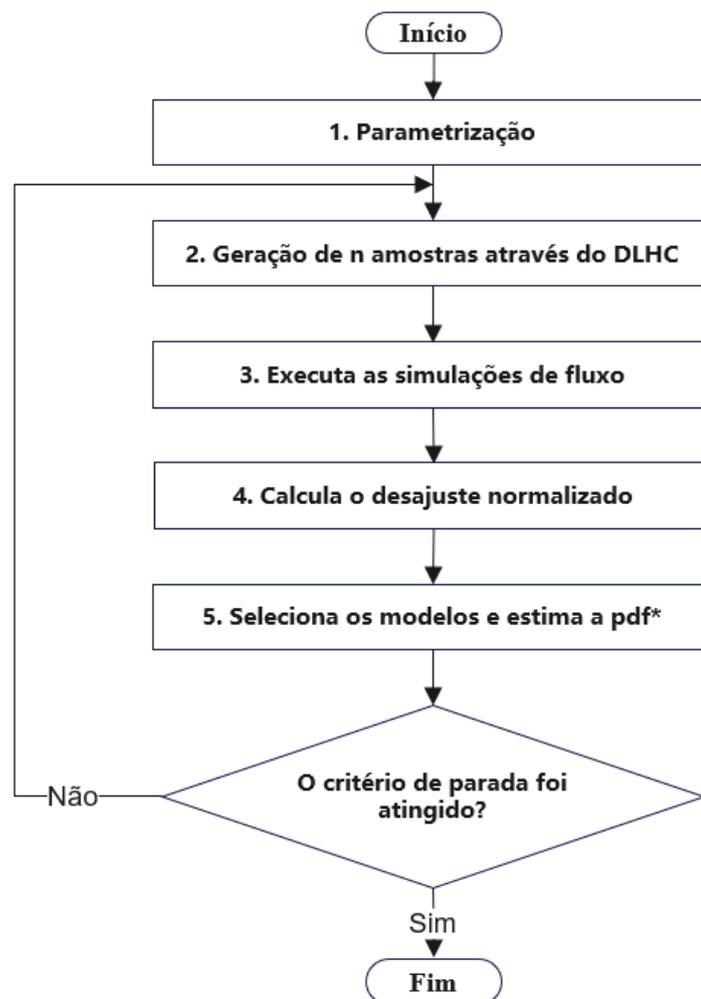


Figura 5: Fluxograma da metodologia do IDLHC.
Fonte: Modificada de Maschio e Schiozer (2016).

2.6 Otimização de estratégias de produção utilizando Inteligência Artificial

Para aprimorar a eficácia e eficiência na produção de petróleo, modelos computacionais avançados que baseados em técnicas de inteligência artificial têm sido desenvolvidos e aplicados com sucesso para otimizar estratégias de produção.

Estimadores mais rápidos de função objetivo

Os FOFE são ferramentas que permitem a obtenção de estimativas rápidas de funções objetivo (Avansi, Hohendorff Filho e Schiozer, 2020), e têm se mostrado fundamentais no contexto de otimização de estratégias de produção, acelerando as avaliações essenciais durante o processo. Essa abordagem se destaca em relação aos métodos tradicionais por consumir menos recursos computacionais ao integrar métodos físicos, analíticos e híbridos, proporcionando, assim, estimativas rápidas e precisas. Com essa característica dos FOFE, é possível explorar um maior número de cenários e variáveis, o que é fundamental em ambientes com elevado nível de incertezas e complexidade. Avansi, Hohendorff Filho e Schiozer (2020) classificaram FOFE em quatro tipos:

- **Proxys:** São modelos analíticos, também conhecidos como modelos substitutos ou metamodelos. Usam funções analíticas para fornecer uma estimativa da FO a partir do modelo de simulação. Exemplos incluem metodologia de superfície de resposta, regressão polinomial, krigagem ordinária, redes neurais artificiais (RNA), funções de base radial etc. A qualidade depende da abordagem matemática, dos dados de entrada e da complexidade do sistema modelado (Nwachukwu *et al.*, 2018; Bruyelle; Guérillot 2019; Kwon *et al.*, 2021);
- **Emuladores:** São aproximações estatísticas de uma FO que fornecem uma estimativa e uma declaração de incerteza sobre essa estimativa. Permitem avaliações muito mais rápidas que os modelos tradicionais e ajudam a identificar regiões compatíveis com os dados observados e incertezas mapeadas, por exemplo, emuladores bayesianos, modelos de krigagem universal, processos gaussianos etc. (Ferreira *et al.*, 2014; Moreno *et al.*, 2018; Wilson, 2016);
- **Modelos Híbridos Baseados em Física e Orientados por Dados:** São modelos que combinam fenômenos físicos e observações de dados, utilizando algoritmos de aprendizado de máquina e IA, por exemplo, Redes Neurais Informadas por Física (PINNs). Estes modelos são empregados na simulação de reservatórios e sistemas de produção, combinando a interpretabilidade de modelos baseados em física com a

eficiência computacional e capacidade de identificação de padrões dos modelos baseados em dados (Kaneko *et al.*, 2022; Michael, 2023; Wang *et al.*, 2023);

- **Modelos Híbridos Baseados em Física e Análise:** São modelos que integram fenômenos físicos com formulações matemáticas para prever interações físicas complexas ao longo do tempo em simulações. Combinam o aprendizado do comportamento físico com formulações analíticas para obter estimativas mais precisas das funções objetivo e das incertezas associadas, por exemplo, modelos semianalíticos, análise de declínio de produção etc. (Hu *et al.*, 2022; Mayerhofer *et al.*, 2018; Mogollon *et al.*, 2022).

A otimização de estratégias de produção utilizando técnicas de IA é uma abordagem inovadora que visa melhorar a eficiência e a rentabilidade das operações de produção de petróleo e gás. A IA compreende uma variedade de técnicas computacionais que permitem sistemas e processos automatizarem a aprendizagem e a tomada de decisões com base em dados (Mohaghegh, 2016; Yeturu, 2020). O conceito de IA tem se tornado cada vez mais importante na otimização de estratégias de produção devido à sua capacidade de lidar com grandes volumes de dados, identificar padrões complexos e adaptar-se dinamicamente às mudanças no ambiente operacional (Panja *et al.*, 2018).

A IA pode complementar ou substituir partes custosas do processo de otimização de estratégias de produção, proporcionando soluções mais eficientes e econômicas. Também pode ser utilizada para automatizar tarefas repetitivas que envolvem uma quantidade intensa de dados, reduzindo significativamente o tempo e os recursos necessários para essas operações. Além disso, a IA pode auxiliar na inserção de novos dados para melhorar o poder de previsão de uma determinada estratégia. Por exemplo, na previsão da produção de campos de petróleo com base em dados históricos de produção e parâmetros operacionais (Panja *et al.*, 2018; Rahmanifard; Plaksina, 2018; Staff *et al.*, 2020; Wang; Feng; Haynes, 2015) e na otimização de estratégias de produção a partir de dados em tempo real (Al Jawhari *et al.*, 2023).

Algoritmos de aprendizado de máquina aplicados na otimização

Os algoritmos de aprendizado de máquina têm sido amplamente aplicados na otimização de estratégias de produção devido à sua capacidade de lidar com grandes volumes de dados e extrair padrões complexos. Neste trabalho, foram abordados três dos principais algoritmos: Floresta Aleatória, *Gradient Boosting Regressor* e Redes Neurais Artificiais.

Floresta Aleatória

A Floresta Aleatória é um algoritmo de aprendizado de máquina que utiliza uma combinação de modelos (*ensemble*) de múltiplas árvores de decisão para melhorar a precisão da previsão. É um algoritmo teoricamente simples, mas muito poderoso (Géron, 2019). Cada árvore na floresta é treinada com um subconjunto aleatório de dados de treinamento, e a média das previsões das árvores individuais é utilizada como a previsão final do modelo, conforme ilustra a Figura 6. Este método é robusto a *overfitting* e pode lidar bem com dados complexos e de alta dimensionalidade (Breiman, 2001).

No contexto de otimização de estratégias de produção, a Floresta Aleatória pode ser utilizada para prever variáveis a partir de um conjunto de parâmetros operacionais. O método é adequado para modelar cenários operacionais (por exemplo, em campos de petróleo em que vários fatores interagem de maneira não linear e complexa) em função da capacidade de lidar com dados complexos e multivariados (Hastie; Tibshirani; Friedman, 2009; Li *et al.*, 2022).

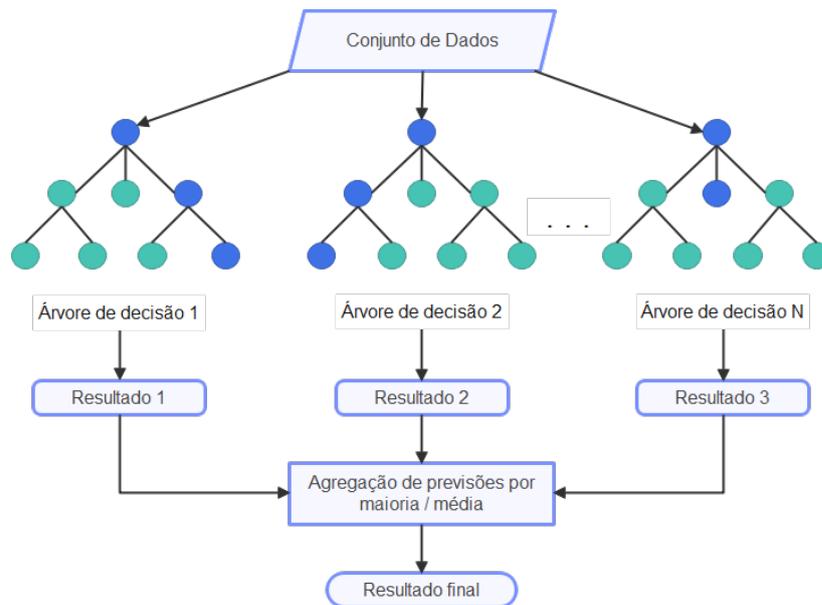


Figura 6: Representação esquemática do método de Floresta Aleatória, onde múltiplas árvores de decisão são combinadas para gerar uma previsão final por meio de agregação de previsões (votação por maioria ou média).

Fonte: Adaptada de Khan *et al.* (2021).

Gradient Boosting Regressor (GBR)

O GBR é um método também baseado em *ensemble* que constrói o modelo de previsão adicionando sequencialmente árvores de decisão. A ideia central do GBR é que cada nova árvore é treinada para corrigir os erros residuais das árvores anteriores, resultando em uma melhoria contínua do modelo (Ilhan; Turali; Kozat, 2023; Shi; Li; Li, 2019; Xiang; Zhang; Xia,

2020), conforme demonstrado na Figura 7. Diferente de métodos mais simples que apenas agregam modelos independentes, o GBR é muito eficaz em capturar relações não lineares entre as variáveis (Breiman, 2004; Friedman; Hastie; Tibshirani, 2000).

Adicionalmente, o *Gradient Boosting* pode ser combinado com técnicas de regularização, como *shrinkage* (redução do tamanho dos passos) (Agapitos; Brabazon; O’neill, 2017) e sub-amostragem, para prevenir *overfitting* e melhorar a generalização do modelo em novos dados (Zhou; Hooker, 2018). Esses ajustes permitem que o GBR mantenha um bom equilíbrio entre viés e variância, resultando em modelos preditivos robustos e mais confiáveis.

O *Gradient Boosting* pode ser aplicado na otimização de estratégias de produção ajustando iterativamente os parâmetros do modelo, conforme proposto por Friedman (2001), com o objetivo de minimizar os erros de previsão e melhorar a acurácia das previsões. Esse ajuste iterativo é desejável em cenários de otimização, onde a precisão na previsão de variáveis pode ter um impacto significativo nas decisões estratégicas.

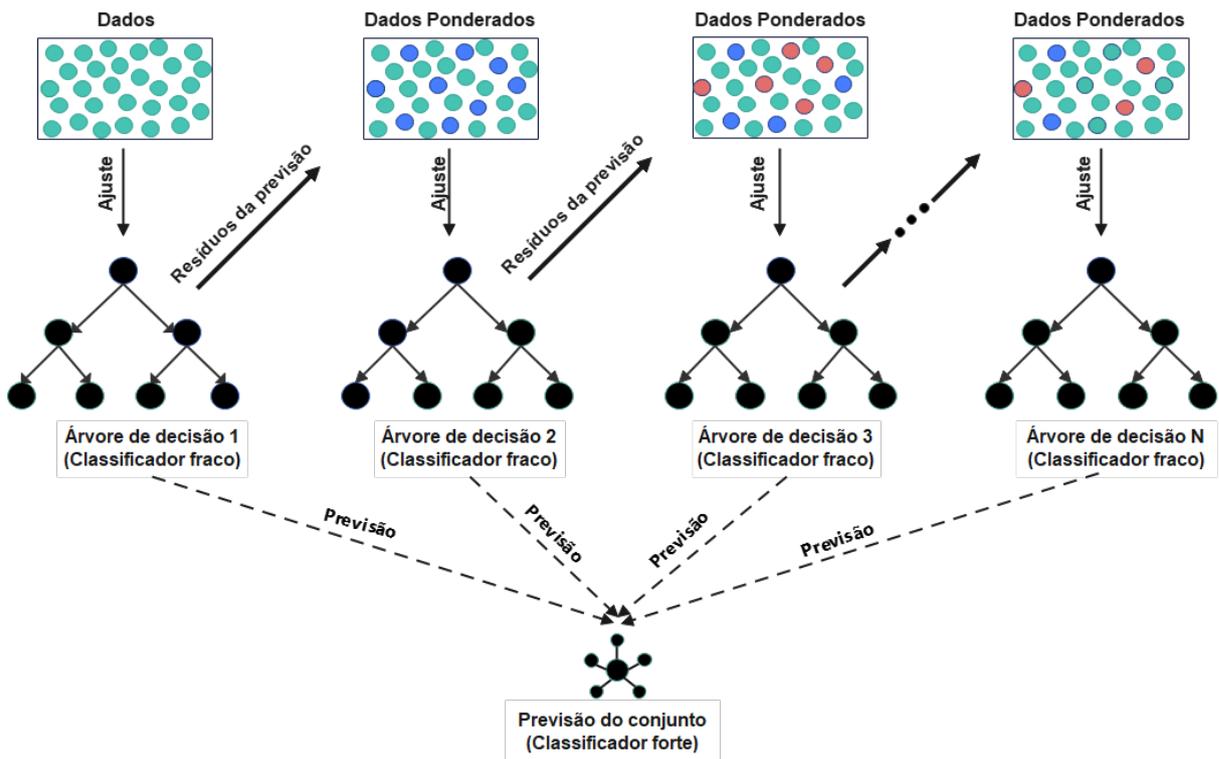


Figura 7: Arquitetura do *Gradient Boosting Regressor*, onde múltiplas árvores de decisão (classificadores fracos) são treinadas sequencialmente em subconjuntos ponderados dos dados, com cada árvore corrigindo os erros residuais das anteriores. A previsão final é obtida pela combinação das previsões de todas as árvores, resultando em um classificador forte.

Fonte: Adaptada de Deng *et al.* (2021).

Redes Neurais Artificiais (RNA)

As RNA são técnicas computacionais que apresentam um modelo matemático inspirado no funcionamento do cérebro humano, compostas por camadas de nós (neurônios) que processam e transmitem informações (Haykin, 2008). Essenciais para inteligência artificial, as RNA são capazes de aprender padrões complexos e não lineares a partir de grandes volumes de dados, desempenhando papel fundamental na previsão e classificação (Pandey *et al.*, 2020), principalmente quando não há soluções numéricas ou analíticas simples (Ashena; Thonhauser, 2015). No contexto da otimização de estratégias de produção, as RNA podem ser treinadas para prever a produção de petróleo com base em dados históricos e em tempo real, permitindo ajustes dinâmicos nas operações para maximizar a eficiência e a rentabilidade.

As RNA são constituídas por dois componentes principais: unidades de processamento e interconexões ou sinapses. A Figura 8 ilustra a estrutura de um Perceptron Multicamadas (MLP), que é um tipo específico de RNA. Os elementos de processamento, também conhecidos como neurônios ou nós, são responsáveis pelo tratamento e processamento da informação. Na Figura 8, esses nós estão representados nas três camadas principais: a camada de entrada, a camada oculta e a camada de saída:

- Camada de Entrada: composta por nós que representam as variáveis de entrada (*Input 1*, *Input 2* e *Input 3*). Esses nós recebem os dados brutos ou características do conjunto de dados que está sendo processado pela rede.
- Camada Oculta: composta por nós que realizam o processamento intermediário dos dados recebidos da camada de entrada (*Hidden Node 1* e *Hidden Node 2*). Esses nós aplicam funções de ativação para extrair características e padrões complexos.
- Camada de Saída: composta por nós que produzem as previsões ou classificações finais da rede neural (*Output 1*, *Output 2* e *Output 3*). Esses nós aplicam uma função de ativação final para gerar o resultado desejado.

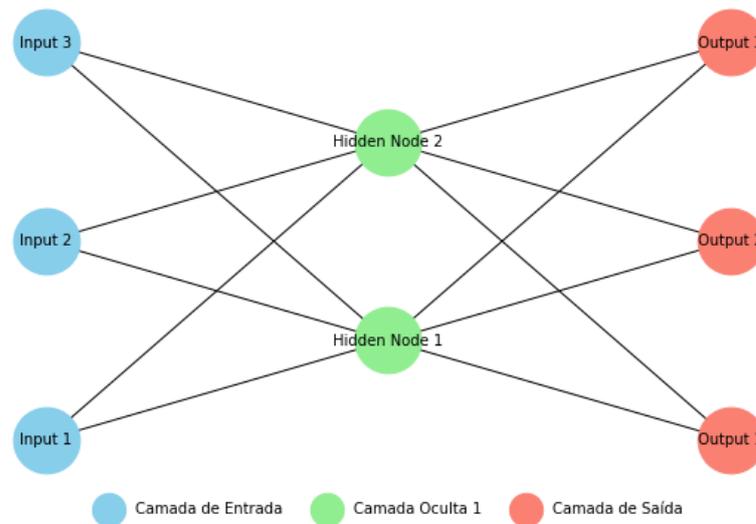


Figura 8: Estrutura de uma RNA com uma camada oculta, ilustrando as conexões entre a camada de entrada, a camada oculta e a camada de saída.

As sinapses estabelecem conexões entre os neurônios. Ainda na Figura 8, essas conexões são representadas pelas linhas que ligam os nós entre as camadas. Essas conexões têm pesos associados que são ajustados durante o processo de treinamento da rede neural. Os pesos determinam a influência que uma entrada tem sobre a ativação de um nó na camada seguinte.

Os modelos de redes neurais mais utilizados são a Função de Base Radial (RBF) e o Perceptron Multicamadas (MLP). A principal diferença entre estes dois modelos está relacionada aos neurônios utilizarem métodos diferentes para processar a informação (Hemmati-Sarapardeh *et al.*, 2020). A Figura 8 representa a estrutura simples de um MLP, onde a informação é processada através de camadas ocultas utilizando funções de ativação, o que é típico para este modelo.

Estas características tornam as RNA poderosas ferramentas na otimização de estratégias de produção, proporcionando a modelagem de sistemas complexos e a adaptação contínua a novos dados e condições operacionais.

Algoritmo do Gradiente Conjugado Escalar

O algoritmo do gradiente conjugado escalonado (SCG), desenvolvido por Møller (1993), é uma técnica de otimização utilizada no treinamento de RNA, especificamente para ajustar os pesos das conexões (sinapses) entre os neurônios (unidades de processamento). Ele é utilizado durante o processo de *backpropagation* (retropropagação) para minimizar a função de custo da rede neural, ou seja, a diferença entre as previsões da rede e os valores reais dos dados de treinamento.

O treinamento de uma rede neural envolve várias etapas principais (conforme mostrado na Figura 9). Na etapa 5, o processo de atualização dos pesos pode ser realizado de duas formas distintas:

- **Gradiente Conjugado:** Ao invés de usar um simples gradiente descendente, o gradiente conjugado é utilizado para encontrar a direção de descida mais eficiente. Ele considera a direção do gradiente anterior para acelerar a convergência.
- **Atualização Escalar:** O algoritmo ajusta os pesos de maneira escalar, considerando tanto a magnitude quanto a direção dos gradientes calculados.

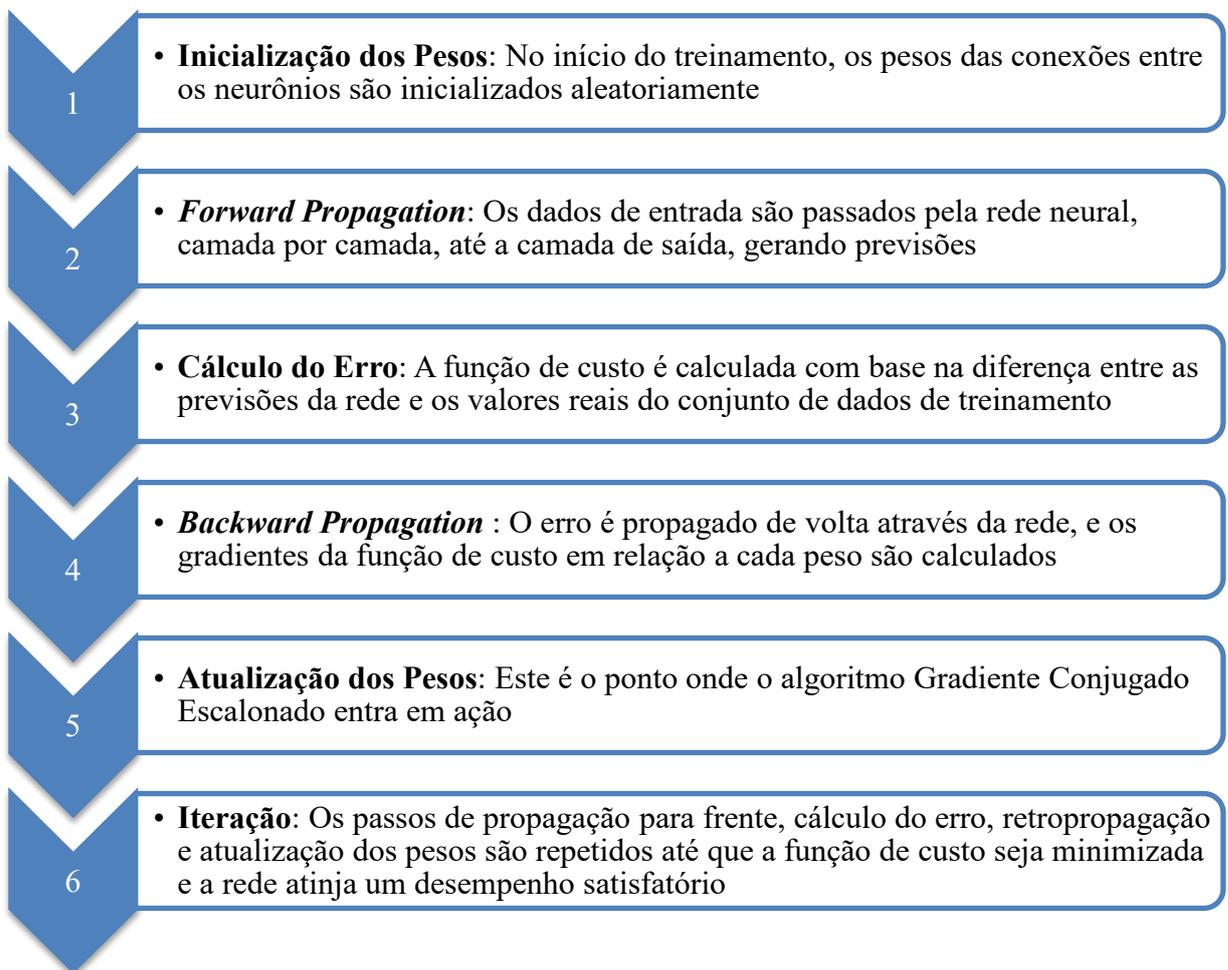


Figura 9: Etapas de como o Gradiente Conjugado Escalonado é aplicado na Rede Neural.

A equação do SCG incorpora técnicas de otimização avançadas para melhorar a eficiência e a convergência. A atualização dos pesos e vieses é realizada de acordo com a Equação 1.

$$y_k + 1 = y_k + \alpha_k p_k \quad \text{Equação 1}$$

onde:

y_k representa o vetor de parâmetros no passo atual (vetor de pesos e vieses atuais);

α_k é a taxa de aprendizado ou tamanho do passo na direção do gradiente, e p_k é a direção conjugada no passo k .

A direção conjugada p_k é calculada utilizando a Equação 2:

$$p_{k+1} = g_{k+1} + \beta_k p_k \quad \text{Equação 2}$$

onde:

g_{k+1} é o gradiente no passo $k + 1$;

β_k é um fator calculado para garantir que as direções sejam conjugadas, e

p_k é a direção conjugada no passo anterior.

O fator de conjugação pode ser calculado usando diferentes fórmulas, dependendo do método específico de gradiente conjugado utilizado (Andrei, 2007; Faramarzi; Amini, 2019). Duas fórmulas comuns são demonstradas na Equação 3 e Equação 4, respectivamente, Método de Fletcher-Reeves (FR) (Fletcher; Reeves, 1964) e Método de Polak-Ribiere (PR)(Polak; Ribiere, 1969).

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \quad \text{Equação 3}$$

$$\beta_k^{PR} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \quad \text{Equação 4}$$

No SCG, o valor de λ_k é ajustado dinamicamente para garantir que o passo seja eficiente. Isso é feito por meio de uma técnica que evita a necessidade de uma busca linear completa a cada iteração. Esse ajuste é feito de acordo com a curvatura local da função de custo, de modo que a atualização no SCG leva em consideração tanto a magnitude quanto a direção dos gradientes calculados, com ajustes adicionais para melhorar a eficiência da convergência, diferindo assim do gradiente descendente.

Trainscg

A função `trainscg` do MATLAB é utilizada para treinar redes neurais empregando o método do SCG. Este método é uma técnica de otimização eficiente que ajusta os pesos e as polarizações da rede neural para minimizar a função de custo, que representa a diferença entre as previsões da rede e os valores reais dos dados de treinamento (The Mathworks Inc., 2017b). As características principais do *Trainscg* são:

- Atualização de Pesos e Polarizações: A `trainscg` ajusta os valores de peso e polarização segundo o método do SCG;
- Treinamento Eficiente: O SCG é eficiente em termos de memória e tempo de computação, tornando-o adequado para redes neurais grandes.

Parâmetros de treinamento

Para configurar corretamente o treinamento, é fundamental entender os principais parâmetros que controlam esse processo. Os parâmetros padrão usados no treinamento de redes neurais com a função `trainscg` estão listados na Tabela 2. O número máximo de épocas, a frequência das exibições de progresso e os critérios de parada são ajustáveis por meio desses parâmetros. Para garantir que o treinamento funcione bem e que a rede neural tenha o desempenho esperado, é fundamental configurar esses parâmetros corretamente.

Tabela 2: Principais parâmetros utilizados no treinamento de redes neurais aplicando a função `trainscg`.

Parâmetro	Descrição	Valor Padrão
<code>net.trainParam.epochs</code>	Número máximo de épocas para treinar	1.000
<code>net.trainParam.show</code>	Número de épocas entre exibições	25
<code>net.trainParam.showCommandLine</code>	Gera saída no terminal	<code>false</code>
<code>net.trainParam.showWindow</code>	Mostra a GUI de treinamento	<code>true</code>
<code>net.trainParam.goal</code>	Meta de desempenho	0
<code>net.trainParam.time</code>	Tempo máximo para treinar em segundos	∞
<code>net.trainParam.min_grad</code>	Gradiente de desempenho mínimo	1E-6
<code>net.trainParam.max_fail</code>	Número máximo de falhas de validação	6
<code>net.trainParam.mu</code>	Parâmetro de ajuste de Marquardt	0.005
<code>net.trainParam.sigma</code>	Determina a mudança no peso para a aproximação da segunda derivada	5,0E-5
<code>net.trainParam.lambda</code>	Parâmetro para regular a indefinição do Hessiano	5,0E-7

2.7 Conceitos econômicos

Na otimização de estratégias de produção de petróleo e gás, os conceitos econômicos são essenciais para a tomada de decisão. O Valor Presente Líquido (VPL) é um dos principais indicadores econômicos, fundamental para avaliar a viabilidade financeira de projetos e estratégias de produção.

2.7.1 Valor Presente Líquido (VPL)

O VPL é uma métrica amplamente utilizada para avaliar a rentabilidade de um projeto de investimento, considerando o valor do dinheiro no tempo (Xavier, 2004). O VPL é calculado

descontando os fluxos de caixa futuros ao seu valor presente, utilizando uma taxa mínima de atratividade. A fórmula do VPL pode ser calculada através da Equação 5.

$$VPL = \sum_{i=1}^N \frac{FCL_{t_i}}{(1+r)^{t_i}} \quad \text{Equação 5}$$

onde:

FCL_{t_i} representa o fluxo de caixa no período t_i ;

i é o índice de identificação de tempo;

r é a taxa mínima de atratividade;

N é o número total de períodos de tempo, e

t_i é o tempo no passo i relacionado à data de análise $t_0 = 0$.

Os investidores podem avaliar um projeto usando esta métrica. Quando o VPL é positivo, significa que o valor atual dos fluxos de caixa futuros esperados é maior do que o custo inicial do projeto. Em outras palavras, o projeto deve gerar mais valor do que o custo inicial, tornando-se viável e recomendado. Por outro lado, se o VPL for negativo, indica que o valor presente dos fluxos de caixa futuros esperados é menor do que o custo inicial. Isso significa que o projeto não recuperará seu investimento inicial, não sendo recomendável para os investidores.

2.7.2 Indicador Econômico de Campo (IEC)

O Indicador Econômico de Campo é utilizado como uma FO inicial nas etapas preliminares do planejamento e otimização de campos petrolíferos. Este indicador é considerado uma versão análoga do VPL. Entretanto, assume que todos os poços são abertos simultaneamente, uma suposição impraticável devida aos altos custos associados à necessidade de uma sonda para perfurar cada poço, mas que simplifica o modelo nas fases iniciais. O IEC permite uma avaliação e comparação justas do comportamento dos poços sob condições uniformes, facilitando decisões estratégicas iniciais sobre o número e a posição dos poços. Após esta fase, com a definição do cronograma de abertura dos poços, a FO é alterada para o VPL, refletindo uma abordagem mais realista e detalhada que considera os tempos específicos de abertura e operação de cada poço, permitindo uma análise financeira mais precisa dos fluxos de caixa futuros (Gaspar *et al.*, 2014).

3 REVISÃO BIBLIOGRÁFICA

3.1 Otimização de estratégia de produção em reservatórios carbonáticos fraturados

Na otimização de estratégia de produção em reservatórios naturalmente fraturados, Nie *et al.* (2012) utilizaram modelos de dupla porosidade e permeabilidade para maximizar a produção. No estudo, foi desenvolvido um modelo matemático para poços horizontais que incorpora o fluxo de dupla permeabilidade, facilitando assim o transporte de fluidos entre os sistemas de matriz e de fratura. Foram utilizadas simulações numéricas para obter curvas de tipo log-log, analisando o comportamento transiente do fluxo por meio de métodos avançados, como a transformação integral de Laplace e a separação de variáveis. As vantagens do modelo de dupla permeabilidade, como o aumento do suprimento de energia e a redução da resposta de matriz-fratura, são demonstradas nessas análises. Apesar dos avanços, o estudo admite limitações, como simplificações, que podem comprometer as previsões em alguns cenários.

Agada *et al.* (2013) exploraram reservatórios carbonáticos complexos por meio de simulações numéricas detalhadas, utilizando afloramentos análogos em alta resolução 3D. O estudo busca entender a influência das heterogeneidades geológicas no desempenho do reservatório e na recuperação secundária de petróleo. Ao utilizar respostas de pressão sintéticas em testes de poços numéricos, foi constatado que a maioria das características estruturais e diagenéticas não possui uma assinatura de pressão única, devido à alta heterogeneidade da matriz. A pesquisa destaca o papel significativo dessas heterogeneidades na canalização dos fluxos de fluidos e sugere que a otimização do posicionamento de poços e das estratégias de injeção são soluções-chave para melhorar a recuperação de óleo. Essa abordagem representa um avanço significativo na utilização de análogos de alta resolução para melhor prever e gerenciar processos de fluxo em reservatórios carbonáticos fraturados.

Kharrat *et al.* (2023) realizaram uma revisão sobre a caracterização de fraturas e seu impacto na produção de petróleo em reservatórios naturalmente fraturados. O estudo destaca a importância de realizar uma caracterização precisa das fraturas e identificar como isso afeta as etapas de recuperação de petróleo, incluindo a manutenção de pressão e os métodos de recuperação melhorada de óleo (EOR). Aplicando uma metodologia que integra dados de geociências e de engenharia de reservatórios, foi desenvolvido um modelo híbrido de fraturas que combina modelagem contínua e discreta para representar de forma mais realista as fraturas. Os resultados destacam a necessidade de atualizações contínuas do modelo com novos dados de campo, essenciais para melhorar a descrição das fraturas e a eficácia da recuperação. As

limitações incluem a complexidade e os custos da integração de dados, além dos desafios de generalização dos resultados devido às simplificações realizadas na modelagem.

3.2 Otimização de EP utilizando RSM e algoritmos de otimização

Maschio, Nakajima e Schiozer (2008) apresentaram uma abordagem inovadora para otimizar a estratégia de produção de campos de petróleo, combinando algoritmos genéticos e mapas de qualidade para localizar e determinar o número de poços. A metodologia aproveita a capacidade dos algoritmos genéticos de lidar com espaços de solução complexos, integrando mapas de qualidade para restringir o espaço de busca. Isso resultou em melhorias significativas no VPL, permitindo encontrar configurações ótimas com um número significativamente reduzido de simulações. Além disso, os autores concluíram que a divisão do processo de otimização em etapas auxilia na redução significativa do número de simulações. Ainda assim, o número de simulações no processo mais custoso chegou a 583 realizações.

Nogueira e Schiozer (2009) introduziram uma metodologia para otimização de estratégia de produção em campos petrolíferos, usando algoritmos genéticos, com foco na quantidade e no posicionamento dos poços. O processo foi dividido em três rotinas independentes, de modo que uma foi utilizada para otimizar a localização dos poços e as outras duas para otimizar a quantidade de poços. Nos campos analisados, a metodologia mostrou eficiência ao otimizar a produção em estratégias tanto com poços verticais quanto com horizontais. Este trabalho demonstrou que a abordagem utilizada permitiu explorar múltiplas estratégias, reduzir o número de simulações e refinar configurações para maximizar a produção e o VPL, contribuindo para a otimização do projeto. Foram necessárias 765 simulações para otimizar a estratégia de produção com poços verticais e 1.272 simulações para otimizar a estratégia de produção com poços horizontais.

Emerick *et al.* (2009) propuseram uma ferramenta para otimização da quantidade, localização e trajetória de poços produtores e injetores, com o objetivo de maximizar o fator de recuperação de óleo e o VPL. A metodologia desenvolvida é baseada em um algoritmo genético, projetado para problemas de otimização numérica com restrições não lineares. O trabalho envolve a parametrização das características dos poços, incluindo localização inicial e final, tipo e *status*, permitindo a otimização simultânea dessas variáveis. A metodologia adota algumas restrições, como o tamanho da malha, comprimento máximo dos poços e distância mínima entre eles. A ferramenta foi aplicada a três campos reais e um caso sintético, e os resultados mostraram um aumento considerável no VPL e no fator de recuperação em relação

aos cenários propostos inicialmente pelos engenheiros de reservatórios. Em um campo real, a otimização resultou em um aumento de 28% no VPL e de 13% na produção acumulada de óleo. No caso sintético, a otimização aumentou o VPL em 31% e a produção acumulada de óleo em 13,6%. Em ambos os casos, o número de simulações necessárias para otimizar o processo foi elevado.

Gaspar, Barreto e Schiozer (2016) propuseram uma metodologia de otimização assistida aplicada na seleção de estratégias de exploração para campos de petróleo marítimos. O processo integra análises de engenharia com o algoritmo genético DECE, para otimizar o número e o posicionamento de poços, além de capacidades de produção e cronogramas de abertura. A metodologia foi aplicada ao benchmark UNISIM-I-D, um campo sintético baseado no Campo de Namorado, localizado na Bacia de Campos, no Brasil. A abordagem determinística refinou as estratégias iniciais, proporcionando até 11,48% de melhoria na produção, na otimização da posição dos poços e economia significativa no número de simulações necessárias. A metodologia comprova a eficiência de uma abordagem sistemática para maximizar a produção e otimizar os custos em cenários de alta complexidade, destacando a eficiência da otimização do posicionamento dos poços.

Hohendorff Filho, Maschio e Schiozer (2016) aplicaram o IDLHC para otimizar a estratégia de produção, buscando maximizar o VPL. A técnica reduz gradualmente o espaço de busca a cada iteração, podendo lidar com elevado número de variáveis discretas e distribuindo adequadamente suas frequências. O método se destaca por explorar novas combinações de forma eficiente com base no conhecimento adquirido ao longo das iterações. Ao aplicar a metodologia, os autores compararam as otimizações entre o IDLHC proposto e o método DECE, disponível em software comercial, que também é um processo iterativo que aplica uma sequência de estágios de exploração e evolução controlada. O IDLHC considera as distribuições de frequência posteriores dos níveis de variáveis discretas para maximizar a FO, que não precisa ser necessariamente monótona, em espaços de busca descontínuos, permitindo encontrar vários ótimos locais. Isso foi feito com sucesso no caso aplicado, comparando-o a uma metodologia de otimização consolidada. O IDLHC encontrou soluções otimizadas de VPL com um número significativamente menor de simulações (900), comparado ao DECE, que precisou de 2.936 simulações.

Rios, Avansi e Schiozer (2020) propuseram um fluxo de trabalho prático para otimizar o desempenho numérico em modelos de simulação de reservatórios que consomem muito tempo, utilizando submodelos e intervalos de tempo reduzidos. O objetivo é melhorar a eficiência dos modelos de simulação de reservatórios, estabelecendo procedimentos para

selecionar submodelos representativos e reduzir o tempo total gasto na otimização numérica. O método inicia com a simulação do caso base, utilizando parâmetros numéricos padrão, seguida de uma avaliação no modelo para identificar áreas de convergência problemáticas. A partir dessa análise, decide-se se é necessário construir um submodelo. O submodelo é extraído das regiões identificadas com problemas de convergência. Posteriormente, é realizada uma análise de sensibilidade e *tuning* numérico nos parâmetros mais críticos do modelo, para reduzir o tempo de simulação. Aplicado em dois casos distintos, o método comprovou sua eficácia ao reduzir o tempo de simulação em 44% e 39%, respectivamente.

Asadian-Pakfar, Sedaei e Nakhaee (2023) propuseram uma abordagem inovadora que utiliza algoritmos de busca inteligentes, como *Particle Swarm Optimization* (PSO) e uma versão combinada de PSO com Algoritmo Genético (PSOGA). A abordagem visa maximizar o VPL, considerando variáveis como localização dos poços, perfuração e taxas de fluxo. Os autores introduziram variantes aprimoradas, TPSO e TPSOGA, alterando a condição de término da otimização para o tempo total, em vez do número máximo de iterações, e ajustando o tamanho da população para melhorar a convergência. A aplicação dessa metodologia em um modelo de reservatório heterogêneo mostrou que, limitando a otimização a 24 horas ou uma redução de 70% no tempo de execução, com uma população de 30 agentes, os mesmos valores de VPL podem ser alcançados. Entretanto, a otimização tanto com o PSO quanto com o PSOGA, demandou aproximadamente 84 horas para otimizar um cenário.

Tabatabaei, Asadian-Pakfar e Sedaei (2023) aplicaram o *Sparrow Search Algorithm* (SSA) e o PSO para otimizar a localização dos poços e as taxas de fluxo de produção e injeção em um reservatório heterogêneo do sudoeste do Irã, com auxílio de mapas de qualidade. Com o objetivo de maximizar o VPL, a metodologia ajustou os parâmetros do SSA através de análises de sensibilidade, mantendo o tamanho da população e o número de iterações fixos para otimizar o tempo de execução. O SSA mostrou-se mais eficaz do que o PSO, aumentando o VPL em até 18,21% em alguns cenários, fornecendo *insights* valiosos sobre a localização dos poços. No entanto, o SSA requer significativamente mais tempo de execução e recursos computacionais do que o PSO, mas, ainda sim, a eficácia do SSA se destaca como uma ferramenta de otimização para a localização de poços, oferecendo uma abordagem promissora para maximizar o VPL no desenvolvimento do projeto.

Uma análise comparativa entre técnicas que utilizam algoritmos de otimização e métodos tradicionais demonstra que os algoritmos, têm um desempenho superior ao lidar com problemas não lineares e de alta dimensionalidade. A revisão dos trabalhos apresentados destaca a relevância dos algoritmos de otimização na busca por estratégias mais eficazes para

produção de petróleo, abordando principalmente questões como o posicionamento e a quantidade ideal de poços. Está claro que os algoritmos são essenciais para a melhoria da otimização das estratégias de produção. Ainda assim, enfrentam limitações significativas quanto ao tempo computacional necessário para alcançar soluções otimizadas.

Nesse contexto, a integração de novas técnicas, como a Inteligência Artificial, surge como uma solução promissora para superar essas limitações, potencializando a capacidade de explorar espaços de soluções mais complexos em menor tempo.

3.3 Inteligência artificial na otimização de estratégias de produção

A introdução de técnicas de Inteligência Artificial (IA) em reservatórios de petróleo representa um avanço significativo para superar as limitações dos métodos tradicionais na otimização de estratégias de produção (Ertekin; Sun, 2019). Com o aumento da complexidade dos campos de petróleo e a necessidade de análises mais precisas, os métodos tradicionais de otimização, na maioria dos casos, são inadequados para lidar com a alta dimensionalidade e os problemas não lineares inerentes à indústria do petróleo (Evensen; Womack, 2020). Nesse contexto, a IA surge como uma alternativa, possibilitando a análise de grandes conjuntos de dados e a identificação de padrões que podem resultar em respostas mais rápidas, favorecendo a tomada de decisões mais eficazes (Kuang *et al.*, 2021). Técnicas como aprendizado de máquina, RNA e otimização baseada em algoritmos evolutivos têm sido aplicadas com sucesso para otimizar a produção de campos petrolíferos (Ertekin; Sun, 2019; Kuang *et al.*, 2021). Essas abordagens permitem explorar, de forma mais rápida, espaços de busca complexos, encontrar configurações ótimas e, assim, melhorar a eficiência operacional e maximizar a FO de interesse (Evensen; Womack, 2020).

3.4 Otimização de EP utilizando RSM, FOFE e algoritmos de otimização

Nwachukwu *et al.* (2018) desenvolveram uma metodologia para avaliação rápida de localizações de poços em modelos de reservatórios heterogêneos, baseada em um método de aprendizado de máquina. A metodologia utilizou o XGBoost (*Extreme Gradient Boosting*), um algoritmo baseado em árvores de decisão, que, neste caso, foi projetado para prever as respostas do reservatório com base nas localizações dos poços injetores. O trabalho aborda cinco estudos de caso sintéticos que incluem diferentes cenários de injeção e produção, utilizando modelos heterogêneos. Para cada estudo de caso, foram necessárias diferentes quantidades de simulações para otimizar a resposta, variando de 20 a 1.000 observações de treinamento. O tempo

necessário para otimizar a estratégia de produção estava relacionado ao tamanho do conjunto de treinamento e do modelo, variando de 37,8 segundos até 10.950 segundos. O *proxy* proposto permite identificar a localização ótima do poço com até 95% menos custo computacional em comparação com métodos tradicionais, além de possibilitar a geração de uma superfície de resposta mais precisa e rápida, com uma precisão de até 99% em relação ao método de otimização tradicional.

Bruyelle e Guérillot (2019) otimizaram a localização dos poços em reservatórios de petróleo, utilizando técnicas de inteligência artificial para maximizar o VPL e minimizar riscos financeiros. A metodologia foi aplicada ao Campo Sintético Brugge, avaliando as funções objetivo VPL e o índice de *Sharpe*, com RNA substituindo o simulador de reservatórios. Ao aplicar o plano experimental *Box-Behnken* e a estratégia evolutiva de adaptação da matriz de covariância (CMA-ES), o treinamento de um Perceptron multicamadas proporcionou uma redução significativa no tempo de computação e no número de simulações necessárias. A técnica resultou em um aumento de 2,88% no VPL em comparação ao caso base, demonstrando a eficácia da abordagem inovadora, embora a qualidade do modelo *proxy* ainda dependa da qualidade dos dados de treinamento.

Santos *et al.* (2020) desenvolveram uma metodologia que integra técnicas de aprendizado de máquina com o algoritmo de otimização IDLHC para reduzir o número de simulações necessárias na otimização de controle de poços de longo prazo. O estudo enfatiza a eficiência computacional ao incorporar múltiplas técnicas de aprendizado de máquina, que preveem cenários com o melhor VPL a partir de dados iniciais de simulação, permitindo uma redução de até 27% no número de simulações sem comprometer a precisão dos resultados. A abordagem proposta demonstra ser uma alternativa viável e eficiente, destacando o potencial do aprendizado de máquina para otimizar estratégias de produção em reservatórios, mantendo a qualidade das decisões sob restrições de tempo e recursos computacionais.

Kwon *et al.* (2021) desenvolveram uma metodologia para integrar um método de aprendizado de máquina no processo de otimização a fim de determinar o posicionamento ideal de um poço produtor de petróleo em um reservatório sintético, considerando as incertezas geológicas. A metodologia utiliza uma Rede Neural Convolucional (CNN) integrada com otimização robusta. A CNN é treinada com um conjunto de simulações para aprender a correlacionar dados de propriedades petrofísicas com a produtividade de hidrocarbonetos. O modelo é então utilizado para prever a produtividade em todas as localizações possíveis do poço para todas as realizações consideradas. A abordagem é testada em um conjunto de realizações que descrevem o reservatório, considerando o aporte de um aquífero ao invés de poços injetores.

O uso do método proposto, que integra CNN com otimização robusta, permite identificar a localização ótima do poço com apenas 4,6% do custo computacional necessário para avaliar exaustivamente todos os cenários possíveis. Entretanto, a CNN pode subestimar a variabilidade em cenários específicos devido às limitações inerentes a proxies de produtividade.

Bertini Junior *et al.* (2022) compararam diferentes modelos de aprendizado de máquina para acelerar a otimização de estratégia de produção, permitindo a avaliação de muitas combinações possíveis de posicionamento de poços. A metodologia consistiu na construção de modelos substitutos baseados em aprendizado de máquina para prever o VPL a partir de dados binários que indicam a presença ou ausência de poços em uma estratégia de produção. Seis algoritmos diferentes foram comparados: *Elastic Net* (ENET), *K-Nearest Neighbors* (KNN), *Multi-Layer Perceptron* (MLP), *Kernel Ridge Regression* (KRR), *Support Vector Regression* (SVR) e *Gradient Tree Boosting* (GTB). A otimização foi realizada utilizando algoritmos genéticos para identificar a estratégia de produção ótima, seguida de uma validação no simulador numérico. O estudo de caso foi baseado em um modelo de campo sintético, o UNISIM-II-D, construído com características análogas aos campos do Pré-sal brasileiro e do campo Ghawar, na Arábia Saudita. O estudo não apresentou uma avaliação quantitativa do VPL obtido através das técnicas propostas, mas sim do erro de previsão das mesmas. A melhoria em relação ao uso direto do simulador foi a capacidade de explorar rapidamente uma grande quantidade de estratégias de produção para encontrar a melhor. O MLP foi o modelo que melhor previu o VPL.

Li *et al.* (2022) propuseram um modelo de previsão de produção baseado em dados reais de produção de um poço de petróleo de Folhelho no campo de Panke, bacia de Ordos. Foi aplicada uma abordagem híbrida que combina redes neurais convolucionais (CNN) e redes neurais de memória de curto e longo prazo (LSTM), otimizada pelo algoritmo PSO. O objetivo é superar as limitações dos modelos de rede neural única para problemas de previsão de séries temporais, melhorando a precisão da previsão de produção dos reservatórios. A metodologia aplicada consiste em três etapas principais: 1) implementação de uma CNN para extrair características espaciais dos dados de produção; 2) implementação de uma LSTM para extrair características temporais dos dados; 3) otimização dos hiperparâmetros do modelo CNN-LSTM utilizando o algoritmo PSO para melhorar a precisão da previsão de produção. Foram realizadas simulações para comparar a precisão da previsão antes e depois da otimização pelo PSO, bem como para avaliar o desempenho do modelo em relação a outras abordagens, como CNN, LSTM e CNN-LSTM sem otimização pelo PSO. Em comparação com o modelo CNN simples ou o modelo de rede neural LSTM simples, o modelo CNN-LSTM híbrido pode caracterizar

melhor as correlações espaciais e temporais, o que faz com que o modelo de previsão CNN-LSTM se ajuste melhor à produção real, sendo considerada uma das abordagens orientadas por dados mais promissoras para enfrentar o desafio da previsão de séries temporais. Entretanto, o modelo proposto apresenta obstáculos-chave na gestão em tempo real, incluindo escassez de dados e custos computacionais.

Kim e Durlafsky (2023) desenvolveram o modelo substituto EDFM-CNN-RNN, que incorpora técnicas avançadas de aprendizado profundo, como redes neurais convolucionais-recorrentes (CNN-RNN). O modelo foi projetado para prever respostas de fluxo e otimizar estratégias de produção, visando melhorar o VPL. A metodologia baseou-se em Modelos de Fraturas Discretas Incorporadas (EDFM), representando a geometria e a conectividade das fraturas juntamente com as propriedades da matriz rochosa. O modelo foi treinado utilizando 300 execuções de simulação, com diferentes perfis de pressão de fundo de poço (BHP) em 20 realizações distintas e então aplicado à otimização de estratégias de produção usando o algoritmo de otimização PSO. Os resultados mostraram uma melhoria significativa no VPL, cerca de 18% em três execuções de otimização, com uma aceleração superior a 500 vezes em comparação com métodos tradicionais. Apesar dos resultados promissores, o modelo requer um número substancial de simulações para o treinamento (300 simulações), que equivale a 160 horas de simulação em série, o que pode limitar sua aplicabilidade em modelos de reservatórios maiores.

Musayev, Shin e Nguyen-Le (2023) otimizaram, em termos de eficiência e precisão, o posicionamento dos poços de injeção de CO₂ e produção de salmoura, utilizando um modelo *proxy* baseado em RNA. A metodologia aplicada consistiu em desenvolver vários modelos de RNA, com diferentes características de entrada, sendo o melhor modelo usado em conjunto com um algoritmo genético para otimizar a localização dos poços no modelo de reservatório completo para avaliar a maximização da eficiência do armazenamento de CO₂ em um reservatório sintético. Comparando a solução baseada no modelo *proxy* com a simulação tradicional, constatou-se que a simulação tradicional exigiu 622 execuções para convergir para o ótimo. Por outro lado, a otimização baseada em RNA alcançou uma solução comparável com apenas 120 execuções, mostrando que a metodologia proposta foi capaz de alcançar resultados semelhantes com aproximadamente 81% menos simulações, reduzindo significativamente o tempo necessário para otimizar o caso.

Son *et al.* (2023) desenvolveram uma *proxy* para otimizar o posicionamento dos poços em um reservatório. O estudo apresenta uma solução baseada em um esquema de amostragem em dois estágios para aprimorar a precisão das previsões do VPL e minimizar a necessidade de

retreinamento durante a otimização. Aplicado aos *benchmarks* sintéticos 3D, modelo PUNQ-S3, que é um modelo retirado de um estudo de engenharia de reservatório em um campo real realizado pela *Elf Exploration Production* (Barker; Cuypers; Holden, 2001) e modelo Egg (Zandvliet *et al.*, 2007), que é uma versão determinística utilizada frequentemente para testar algoritmos de otimização. O método emprega modelos *proxy* baseados em redes neurais convolucionais (CNNs) para prever com precisão a produção e otimizar o posicionamento dos poços. O uso do esquema de amostragem permitiu reduzir em até 48% o tempo de computação, com um erro médio de previsão do VPL inferior a 2,12%. Esses resultados demonstram que a abordagem proposta pode melhorar significativamente a eficiência e a precisão da previsão e otimização, minimizando a necessidade de retreinamento. Apesar do sucesso, a técnica pode encontrar limitações em relação à qualidade dos dados de entrada e ao modelo *proxy*, indicando a necessidade de um planejamento cuidadoso na seleção de dados e de um modelo adequado.

Os estudos revisados evidenciam uma tendência crescente na integração de técnicas avançadas de aprendizado de máquina na otimização de estratégias de produção em reservatórios de petróleo. Essas ferramentas têm sido utilizadas para construir FOFE que podem prever com precisão as respostas dos reservatórios, minimizando a necessidade de executar repetidamente simulações numéricas. Isso reduz significativamente a dependência do simulador e, conseqüentemente, o tempo de computação necessário para otimizar as estratégias de produção, acelerando assim o processo. Os FOFE baseados em aprendizado de máquina têm sido capazes de melhorar a precisão das previsões em relação aos métodos tradicionais de otimização. Isso é especialmente importante em reservatórios carbonáticos heterogêneos, onde as relações entre as propriedades do reservatório e a produção de hidrocarbonetos podem ser complexas e não lineares.

No entanto, apesar desses avanços, persistem desafios e lacunas na literatura que necessitam de abordagem para solucionar o problema de maneira mais eficiente. Dentre os principais desafios estão (1) a necessidade de obtenção de dados de treinamento com qualidade e representatividade para garantir a precisão dos modelos de aprendizado de máquina; e (2) o custo computacional significativo associado ao treinamento, validação e teste de modelos complexos, especialmente em casos reais, como o deste trabalho, onde número de variáveis e incertezas são elevados. Além disso, há uma carência de estudos que integrem eficientemente técnicas de aprendizado de máquina com métodos de otimização hierárquica em ambientes de reservatórios reais que validem as repostas obtidas pelo FOFE em simuladores comerciais.

Esta dissertação contribui para a literatura ao propor uma metodologia que integra o FOFE baseado em aprendizado de máquina com técnicas de otimização hierárquica visando

reduzir o número de simulações necessárias sem comprometer a qualidade dos resultados. Espera-se, assim, demonstrar a eficácia e eficiência da metodologia, oferecendo uma solução prática para a otimização de estratégias de produção em reservatórios heterogêneos, preenchendo as lacunas identificadas na literatura.

4 METODOLOGIA

A metodologia proposta neste trabalho está estruturada em duas partes fundamentais. Na Parte 1, a estratégia de produção é otimizada nominalmente para um caso base específico, utilizando RSM e um algoritmo de otimização. O resultado dessa otimização estabelecerá uma referência essencial para comparações com a Parte 2, que utiliza o FOFE para acelerar partes mais custosas (em número de execuções do simulador de fluxo) do processo de otimização.

4.1 Parte 1

A metodologia específica da Parte 1 é estruturada conforme ilustrado na Figura 10. O processo inicia com a definição de um modelo base, que serve como ponto de partida para otimização. Após essa definição, é estabelecida a parametrização das variáveis de otimização. Em seguida, é definido um algoritmo de otimização. Na fase final, ocorre a implementação da otimização, que é subdividida em cinco etapas. Cada etapa tem como objetivo obter a melhor otimização possível para o caso base. As caixas do fluxograma da Figura 10 que estão destacadas na cor verde representam etapas que foram otimizadas utilizando o mesmo algoritmo de otimização e são aplicadas nas Partes 1 e 2 da metodologia, através de uma abordagem hierárquica, em três fases subsequentes.

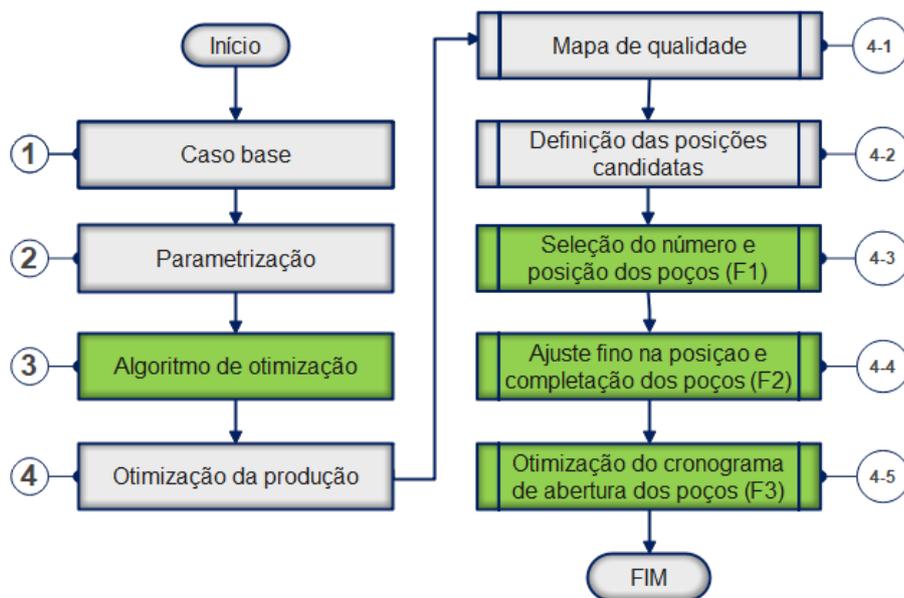


Figura 10: Fluxograma da metodologia proposta na Parte 1 do trabalho.

Etapa 1: Caso base

Antes de otimizar uma estratégia de produção, é necessário estabelecer um caso base. Este processo envolve a análise e filtragem de múltiplos cenários de simulação, resultando na escolha de um modelo que melhor representa as condições iniciais para a otimização da produção. Após definir o caso base, é gerada uma EP inicial que servirá de base para comparações com a otimização desenvolvida neste trabalho.

Etapa 2: Parametrização

Após definir o caso base, o projeto avança para a fase de parametrização. Essa etapa é essencial para definir as variáveis que serão otimizadas com o uso de um algoritmo. O projeto pode incluir a otimização das variáveis G1, G2 e G3 de forma individual, simultânea ou em uma combinação de dois grupos, como G1 e G2, por exemplo.

Etapa 3: Seleção do algoritmo de otimização

O algoritmo de otimização é selecionado e implementado em cada fase do processo hierárquico de otimização da produção, conforme visualizado na Figura 10, destacado pela cor verde. A ideia é aplicar o algoritmo de otimização para aumentar a velocidade de convergência das respostas ao maximizar ou minimizar uma FO selecionada, de acordo com o objetivo de estudo, a fim de atender às metas econômicas e operacionais do projeto.

Etapa 4: Otimização da produção

Nessa fase, são aplicadas técnicas para reduzir o número de execuções de simulação necessárias para o desenvolvimento da estratégia de produção mais eficaz. Nesse contexto, uma técnica essencial empregada para tal propósito foi a geração de um mapa de qualidade para auxiliar na redução do espaço de busca. Os mapas auxiliaram a eliminar posições candidatas para poços produtores em áreas com baixo potencial produtivo, enquanto os injetores foram alocados de forma a suprir a perda de pressão ocasionada pela produção. Em seguida, foi realizada uma otimização hierárquica em três estágios, respeitando um critério de parada para assegurar que, uma vez encontrada a melhor solução, o processo seja finalizado sem realizar simulações adicionais desnecessárias. O critério consiste em interromper o processo após três iterações consecutivas sem melhora na função objetivo (FO) avaliada.

Etapa 4.1: Mapa de qualidade

O mapa de qualidade aplicado neste trabalho foi construído com base na modificação proposta por Cavalcante Filho (2005) da metodologia original de Cruz (2000). O método utilizado consiste em avaliar a produção acumulada de óleo (N_p) e produção acumulada de gás (G_p) (convertida posteriormente em N_p equivalente), em várias posições, buscando cobrir grande parte da área do reservatório. Para as regiões onde não são calculados a N_p e a G_p , é então realizada uma interpolação para estimar a propriedade nas áreas adjacentes.

Etapa 4.2: Definição das posições candidatas

As posições candidatas, tanto para os poços produtores quanto para os injetores, foram definidas a partir do mapa de qualidade gerado. O critério utilizado para definir as posições dos poços produtores, foi evitar posicionar os poços em regiões com baixa qualidade, mantendo um distanciamento mínimo de três blocos entre eles, o que equivale a aproximadamente 450 m. Os poços injetores foram posicionados margeando os produtores, respeitando o mesmo espaçamento mínimo dos poços produtores.

Etapa 4.3: Seleção do número e posição dos poços (P1-F1)

A aplicação combinada de RSM com um algoritmo de otimização é fundamental para varrer o espaço de busca e determinar o número e posicionamento ideais dos poços. Esse processo é a primeira etapa da otimização hierárquica, e foi realizado de forma iterativa, eliminando sistematicamente as posições e quantidades de poços menos favoráveis para as estratégias avaliadas com base nos valores da FO definida, buscando a convergência para uma solução ótima.

Etapa 4.4: Ajuste fino na posição e completção dos poços (P1-F2)

Após definir o número e as posições ideais dos poços, a próxima etapa do processo de otimização hierárquica consiste em realizar um ajuste fino nas posições x e y previamente determinadas e na completção dos poços. Isso inclui avaliar posições mais distantes do contato água-óleo, conforme o que for definido nas premissas do projeto. Assim como no processo anterior, a ideia é que o algoritmo otimize a posição dos poços, bem como a completção das camadas, visando evitar a produção de água e maximizar a produção de óleo para os poços produtores. Além disso, é necessário definir posições para os poços injetores que auxiliem na manutenção da pressão e contribuam para a varredura do óleo.

Etapa 4.5: Otimização do cronograma de abertura dos poços (P1-F3)

Nas etapas anteriores, o cronograma de abertura de todos os poços estava definido para colocá-los em operação no tempo inicial do projeto, simplificando as avaliações. Entretanto, nesta etapa, foi definido um intervalo de datas candidatas para a abertura dos poços produtores e injetores, com o objetivo de encontrar a melhor configuração para maximizar a FO avaliada.

4.2 Parte 2

A Parte 2 da metodologia consiste em desenvolver uma estratégia de produção com a mesma parametrização definida na Parte 1, buscando reduzir o esforço computacional para encontrar uma estratégia de produção ótima. A metodologia empregada, representada pela Figura 11 envolveu a seleção de dados para o treinamento, validação e teste a serem aplicados em três tipos de FOFE, dos quais foi selecionado um para estimar uma FO, como, por exemplo, o VPL.

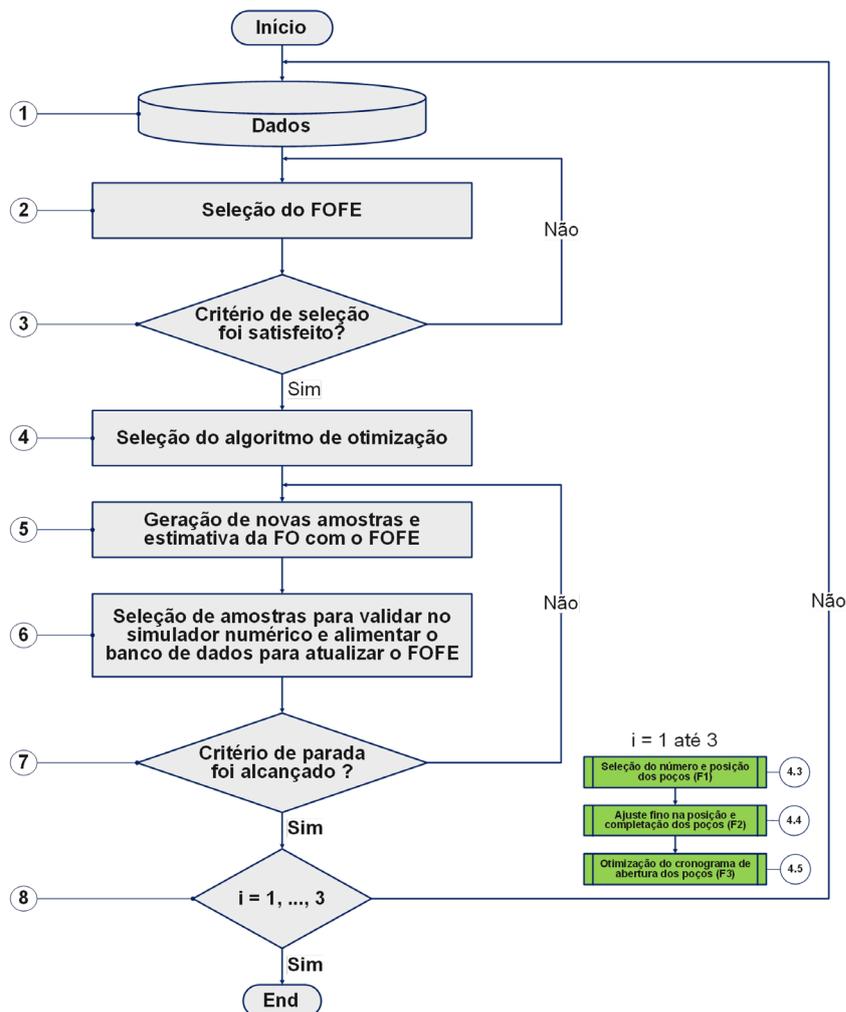


Figura 11: Fluxograma da metodologia proposta na Parte 2 do trabalho.

Ao aplicar o algoritmo de otimização, em cada iteração, é gerado um número específico de novas estratégias de produção. A função objetivo avaliada tem seu valor estimado pelo FOFE, e parte dessas amostras é selecionada para validar a estimativa utilizando o simulador numérico. As amostras escolhidas para validação são então incorporadas ao banco de dados inicial. Após cada iteração, um novo FOFE é gerado com um conjunto expandido de amostras, tornando o processo de treinamento do algoritmo dinâmico. O objetivo da inclusão de novos dados em cada iteração é aumentar a eficiência na estimativa da FO em avaliação. Um critério de parada é estabelecido caso a otimização não alcance os resultados esperados.

4.2.1 Dados

Este primeiro período da Parte 2 da metodologia consiste em coletar os dados que serão utilizados no treinamento, validação e teste dos métodos de aprendizado de máquina a serem avaliados. Esses dados, por sua vez, compreendem um conjunto de amostras contendo a entrada e a saída. As entradas são as variáveis de acordo com cada fase do processo hierárquico (por exemplo, na fase 1, número e posições dos poços), e a saída é uma função objetiva definida pelo usuário (VPL, por exemplo).

4.2.2 Seleção do FOFE

Nesse momento, são realizados o treinamento, a validação e o teste dos FOFEs selecionados pelo usuário, que apresentam potencial para identificar o algoritmo que melhor estime a FO a partir do conjunto de dados disponível. Além disso, é determinado o número de amostras para treinamento, validação e teste, que são importantes para a criação do modelo analítico (FOFE). O desempenho de cada método é avaliado com base em uma ou mais métricas estabelecidas, garantindo a seleção do modelo mais eficaz e preciso para prever a FO.

4.2.3 Critérios para seleção do FOFE

Os critérios definidos no código fazem parte do processo de treinamento de um FOFE e determinam que o treinamento deve continuar até que um determinado nível de desempenho, avaliado por métricas, seja atingido para garantir que o algoritmo tenha um desempenho satisfatório, ou que um número máximo de iterações seja realizado para que o treinamento não consuma muito tempo.

4.2.4 Algoritmo de otimização

Para essa etapa, devemos selecionar o algoritmo de otimização que será utilizado com o FOFE durante a otimização da produção. O algoritmo selecionado é aplicado nas mesmas fases do processo hierárquico da Parte 1 da metodologia – Seleção do número e posição dos poços (P2-F1), Ajuste fino na posição e completação dos poços (P2-F2) e Otimização do cronograma de abertura dos poços (P2-F3).

4.2.5 Geração de novas amostras

O algoritmo de otimização selecionado é usado para gerar um novo conjunto de amostras. Como o custo computacional da estimativa do valor da função objetiva com o FOFE (função analítica, como por exemplo, $y = ax + b$) é significativamente menor, o número de amostras geradas pode ser maior, pois o cálculo ocorre em milésimos de segundos. Assim, o espaço de busca pode ser explorado com uma quantidade maior de avaliações, que seriam proibitivas com o uso do simulador de fluxo.

4.2.6 Selecionar amostras para validar no simulador e alimentar o banco de dados para atualizar o FOFE

Depois de gerar o novo conjunto de amostras, elas devem ser validadas usando um simulador numérico para avaliar a precisão e capacidade do modelo analítico em encontrar soluções melhores. Como esse processo é iterativo, as amostras validadas podem ser incorporadas dinamicamente ao conjunto de dados de treinamento inicial, visando melhorar a capacidade de estimativa do FOFE.

4.2.7 Análise de conclusão do processo hierárquico

Assim como na Parte 1 da metodologia, a otimização será realizada de forma hierárquica, contemplando as mesmas três fases P2-F1, P2-F2 e P2-F3. Essa etapa determina que, quando o processo atinge a terceira fase, conforme indicado pelo contador, a otimização é considerada concluída e o processo é finalizado.

5 APLICAÇÕES

Este estudo é baseado em um reservatório marítimo denominado Campo C, localizado na Bacia de Santos, na região do pré-sal brasileiro. O reservatório apresenta uma litologia carbonática complexa, com fraturas tanto na direção horizontal quanto na vertical, conferindo uma natureza de 2P2K. Essas características geológicas promovem um fluxo de fluido dentro do reservatório significativamente influenciado por falhas conectivas e corredores de fraturas, o que aumenta a produtividade do campo. O reservatório contém um depósito de óleo leve, classificado como *Black-Oil*, com a fase gasosa dispersa na fase oleosa. A injeção de água é definida nas premissas como o método de recuperação. Atualmente, o Campo C está em fase inicial de desenvolvimento. Essa etapa é marcada por incertezas, tanto geológicas quanto de fluidos, devido às características do reservatório, com informações provenientes exclusivamente de dados de testes de poços. O reservatório contempla cinco poços perfurados, sendo quatro produtores e um injetor. O modelo de simulação compreende aproximadamente 2,4 milhões de células distribuídas espacialmente em $79 \times 173 \times 228$ nas direções x, y e z, respectivamente, com 125.625 células ativas (células que contribuem para o fluxo e são incluídas no cálculo das EDP. Apenas uma fração das células ativas totais é considerada nesses cálculos). Esse modelo de alta resolução, com um tamanho médio de célula de $150 \times 150 \times 5$ m, permite uma análise detalhada do comportamento do reservatório.

Caso base

O estudo de caso utilizado neste trabalho foi definido por Danes *et al.* (2024) através da aplicação da metodologia de seleção de modelos representativos proposta por Meira *et al.* (2020). Durante todo o processo de seleção conduzido por Danes *et al.* (2024), foi realizada uma análise de dados para reduzir a incerteza, resultando em uma nova distribuição de variáveis incertas, a partir da qual foi gerado um conjunto de 300 realizações. Esses novos cenários foram simulados posteriormente e filtrados para selecionar apenas os modelos que representavam com precisão mais próxima do histórico dos testes de poços dentro de uma faixa de aceitação pré-definida. Após essa avaliação, 195 modelos de simulação foram considerados adequados para análise posterior e foram utilizados como dados de entrada para a seleção do caso base, que fosse representativo para as propriedades do reservatório na média de um conjunto das funções-objetivo NPV, N_p , W_p e ORF. Desses 195 casos filtrados, foram pré-selecionados nove modelos representativos com diferentes probabilidades de ocorrência. O modelo escolhido por

Danes *et al.* (2024) como caso base representa um cenário com probabilidades médias em todas as funções objetivas, exceto pelo fator de recuperação.

Estratégia de Produção (EP) inicial

Após a definição do caso base, Mello *et al.* (2024) otimizaram uma estratégia de produção inicial (EPI). A metodologia empregada consistiu na aplicação de técnicas de engenharia de reservatórios para selecionar manualmente uma estratégia de produção. A EPI foi utilizada na Parte 1 da metodologia deste trabalho para avaliar o desempenho da otimização com RSM e o algoritmo de otimização.

5.1 Parte 1 – Otimização utilizando RSM

5.1.1 Premissas estabelecidas para o Campo C

Devido a questões estratégicas operacionais, o campo em estudo possui algumas restrições físicas que impactam o processo de otimização da estratégia de produção. Essas restrições são baseadas nas condições de operação real do campo e são apresentadas de forma detalhada na Tabela 3.

Tabela 3: Definições das premissas impostas para o Campo C.

Número de poços		15 (no máximo)	
Tipo dos poços (produtores ou injetores)		Poços verticais	
Posicionamento dos poços (candidatos)		Posições I, J, K	
Tempo total de simulação		11.888 dias (32 anos e seis meses)	
Configuração da plataforma (capacidade máxima da unidade de produção), seleção de injeção de água como método de recuperação avançada	Processamento da produção de:	Líquido (CpL)	15.200 m ³ /dia
		Óleo (CpO)	15.200 m ³ /dia
		Gás (CpG)	4.200.000 m ³ /dia
		Água (CpW)	12.000 m ³ /dia
	Processamento da injeção de:	Água (CiW)	18.200 m ³ /dia
Máximo de conexões de poços (nw)		15	
Mínimo entre perfuração, completação e conexão dos poços		30 dias para cada poço	
Método de recuperação		Injeção de água	
Elevação mínima da completação em relação ao aquífero		50 m	

5.1.2 Parametrização das variáveis

Este estudo é focado apenas na otimização das variáveis do G1, conforme detalhado na Tabela 4. Foi empregada a abordagem de poços candidatos (cw) para definir o número e posições ideais para os poços, atribuindo 0 aos poços não incluídos na estratégia e 1 aos poços

incluídos. Para os poços candidatos produtores (nwp), foram definidas 50 posições e 52 posições candidatas para poços injetores (nwi). Definimos o intervalo de abertura dos poços em 30 dias, no mínimo, que é o tempo necessário para uma sonda de perfuração perfurar, completar e colocar o poço em operação. Por fim, foi atribuído ao algoritmo de otimização determinar o tempo e a sequência de abertura dos poços.

Tabela 4: Parametrização das variáveis otimizadas.

	Variável	Descrição	Intervalo
G1	cw	Incluir ou não na estratégia de produção	[0; 1]
	nwp	Número de poços candidatos produtores	[0, 1, ... , 50]
	nwi	Número de poços candidatos injetores	[0, 1, ... , 52]
	wst	Intervalo de abertura entre os poços	[30]
	wos	Sequência de abertura dos poços produtores e injetores	[P - I]

As variáveis do G2L são essenciais para o processo de otimização, pois definem restrições operacionais e de monitoramento com o objetivo de alcançar eficiência e segurança das operações. Neste trabalho, o foco foi dado nas variáveis que mais impactam em uma otimização da produção de óleo para um campo em estágio inicial de desenvolvimento e, portanto, as variáveis do G2L foram consideradas fixas e configuradas de acordo com os valores apresentados na Tabela 5.

Tabela 5: Valores fixos das variáveis do G2L.

	Tipo do poço	Descrição	Tipo da restrição	Valores
G2L	Produtor	Mínimo BHP	Operacional	288,3 kgf/cm ²
		Máxima vazão de líquidos	Operacional	6.600 m ³ /dia
		Mínima vazão de óleo	Monitoramento	79,5 m ³ /dia
		Máximo corte de água	Monitoramento	95%
	Injetor	Máximo BHP	Operacional	750 kgf/cm ²
		Máxima vazão de líquidos	Operacional	8.300 m ³ /dia

5.1.3 Seleção e parametrização do algoritmo de otimização

O algoritmo IDLHC foi selecionado e aplicado em ambas as partes da metodologia do projeto, nas três etapas do processo hierárquico devido sua eficiência demonstrada em problemas de otimização (Hohendorff Filho; Maschio; Schiozer, 2016; Mirzaei-Paiaman; Santos; Schiozer, 2023; Santos *et al.*, 2020). A abordagem foi configurada para reduzir gradativamente o espaço de busca a cada iteração, utilizando 20% das melhores amostras para redistribuir as probabilidades da amostragem das iterações subsequentes. Essa configuração auxilia no aprimoramento progressivo da solução ótima, buscando maximizar o IEC/VPL (a

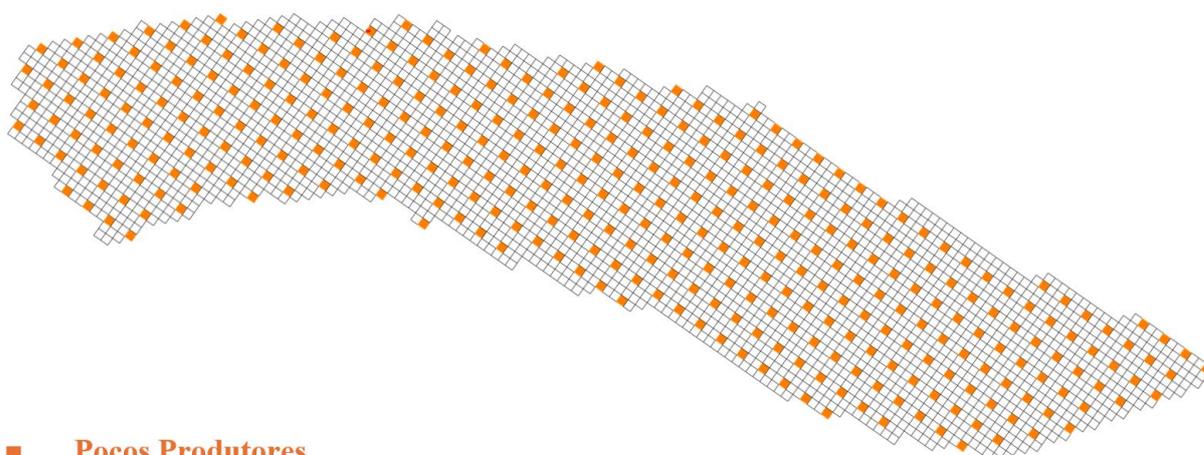
dependem da etapa hierárquica) até que um dos critérios de parada estabelecidos seja atingido: quando não ocorrer evolução da FO avaliada após três iterações consecutivas – quando o número mínimo de níveis for igual a 1 – quando atingir a 20ª iteração. A Tabela 6 mostra a parametrização do IDLHC em ambas as partes do processo de otimização.

Tabela 6: Parametrização do algoritmo de IDLHC.

Parte da metodologia	1	2
Tipo de otimização	Maximização	Maximização
Função objetivo	IEC/VPL	IEC/VPL
Número de amostras	100	100/1000
% de aproveitamento	20	20
Número máximo de iterações	20	20
Número mínimo de níveis	1	1

5.1.4 Construção do mapa de qualidade

Na construção do mapa de qualidade 351 poços produtores fixos foram distribuídos ao longo do reservatório, mantendo um intervalo de dois blocos entre cada poço, que equivale a aproximadamente 300 m, conforme ilustrado na Figura 12. Foi realizada apenas uma execução do modelo no simulador de fluxo *Black-Oil* (IMEX 2019), desenvolvido pela CMG, e calculada a produção acumulada de óleo acrescida da produção acumulada de gás (convertida em barril de óleo equivalente). A soma dessas duas produções conferiu a qualidade de cada posição referente aos 351 poços. Nas regiões adjacentes, foi realizada a extrapolação utilizando Ponderação Inversa da Distância, através de um algoritmo desenvolvido no MATLAB 2017 para estimar a qualidade na região entre os poços.



■ Poços Produtores

Figura 12: Definição do número e posicionamento dos poços produtores para construção do mapa de qualidade.

5.1.5 Definição das posições candidatas e completção dos poços produtores e injetores

Para a otimização da posição e número dos poços, foram consideradas 102 posições candidatas, além de cinco fixas (poços já perfurados). Essas posições foram definidas com o auxílio do mapa de qualidade. Foram definidas 50 posições candidatas para poços produtores, utilizando como critério para seleção das posições, evitar posicionar os poços em regiões de baixa qualidade, garantindo um espaçamento mínimo de três blocos (equivalente a aproximadamente 450 m), de modo a minimizar a interferência na produção entre poços vizinhos além de questões operacionais (Ajisafe *et al.*, 2017; Feder, 2019; Guo *et al.*, 2019; Pichon *et al.*, 2018) . Simultaneamente, 52 posições candidatas para poços injetores foram estabelecidas próximas aos produtores, com o objetivo de equilibrar a perda de pressão.

As variáveis possuem dois níveis (0 e 1) e o IDLHC, através do sorteio, definirá, além dos cinco poços já perfurados, até dez poços (produtores e injetores) adicionais, respeitando o número máximo de 15 poços que a plataforma suporta (Tabela 3). Durante as três fases do processo hierárquico, em ambas as partes da metodologia, foi utilizada a ferramenta OTM do programa de otimização *in-house* MERO (2022).

Em relação à completção dos poços, foi seguida a restrição estabelecida para os poços produtores, completando-os desde a camada superior até a distância mínima de 50 m acima do contato água-óleo. Os poços injetores foram completados em todas as camadas do reservatório, garantindo uma abordagem abrangente nessa primeira etapa do processo hierárquico. As simulações numéricas foram realizadas com o IMEX (2019) em todas as etapas da simulação. A FO avaliada teve seu valor calculado, em todo o projeto, através da ferramenta OTM do programa MERO (2022).

5.1.6 Ajuste fino na posição e completção dos poços

Uma vez definidos o número e a posição dos poços, foi proposto um ajuste fino para encontrar melhores posições e completções. Nesta fase subsequente, foram explorados 80 blocos adjacentes à posição de cada novo poço perfurado, distribuídos uniformemente, com o objetivo de refinar as posições pré-estabelecidas dos novos poços adicionados à estratégia, conforme ilustrado na Figura 13, na qual o retângulo vermelho hachurado destaca novas posições a serem exploradas.

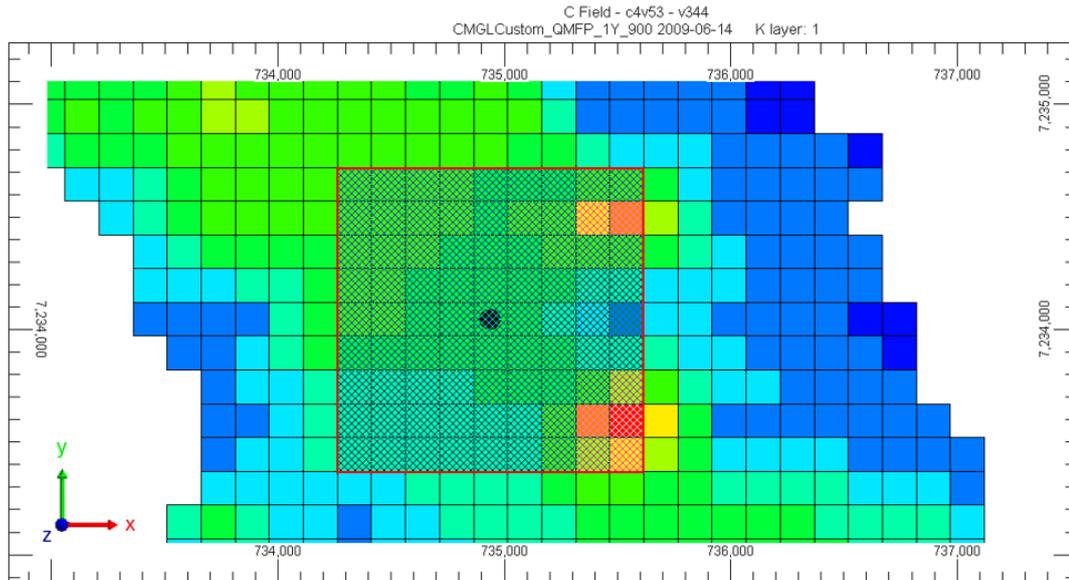


Figura 13: Vista superior do reservatório destacando o espaço de busca avaliado para cada novo poço adicionado à estratégia.

Em relação à completção dos poços, inicialmente, a premissa estabelecia uma análise para não completar os poços produtores em um intervalo de 50 m acima do contato óleo-água; porém, nesta nova fase do estudo, foi avaliado um intervalo considerando até 100 m. Foi analisada também a possibilidade de não completar os poços injetores em um intervalo de até 50 m abaixo do topo do reservatório, conforme indicado na Figura 14,. As abordagens propostas têm o objetivo de mitigar o avanço precoce da água na produção dos poços.

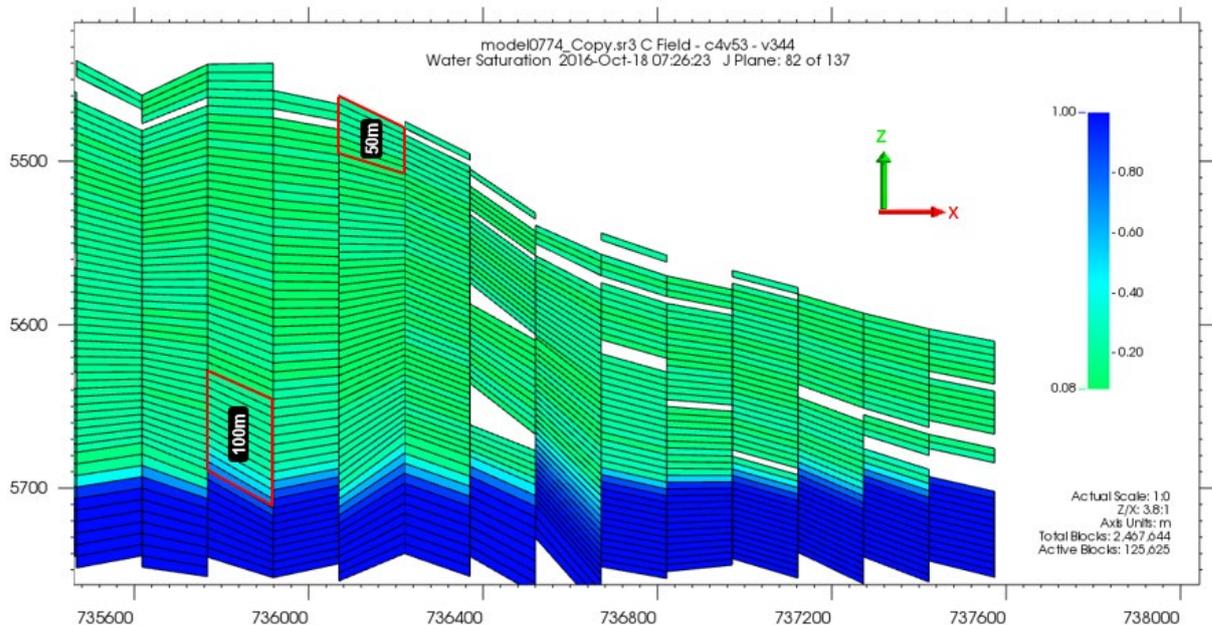


Figura 14: Corte bidimensional (XZ) do reservatório mostrando os intervalos de completção avaliados para não serem completados no ajuste fino para os poços produtores (100 m) e para os injetores (50 m). O eixo Z indica a posição em profundidade do reservatório, enquanto o eixo X ilustra a posição horizontal do modelo de fluxo.

5.1.7 Configuração do cronograma de abertura dos poços

A definição do cronograma de abertura dos poços foi baseada em estudos preliminares. As avaliações consideraram o comportamento da produção dos poços já perfurados e o impacto da produção de novos poços na capacidade de processamento da plataforma. Dessa forma, a abertura dos poços produtores foi avaliada no período inicial subsequente ao tempo de referência do projeto (3.105 dias), enquanto os poços injetores foram extrapolados por um período de dias maior, que compreende desde o tempo referência até dois terços do tempo de avaliação do projeto. que ó período em que a queda de pressão é mais acentuada (Figura 18), conforme destacado na Tabela 7.

Tabela 7: Intervalos de abertura avaliados para os poços injetores e produtores.

Tipo do poço	Intervalos de Abertura (em dias)
Injetor	3136, 3167, 3195, 3226, 3256, 3287, 3317, 3348, 3379, 3409, 3440, 3470, 3652, 3835, 4018, 4201, 4413, 4566, 4809, 5052, 5296, 5540, 5752, 5997, 6239, 6362, 6604, 6848, 7093, 7335, 7578
Produtor	3136, 3440, 4475

5.2 Parte 2 – FOFE + IDLHC

Nesta segunda parte do trabalho, buscou-se definir o tipo de FOFE a ser aplicado na otimização da estratégia de produção. O objetivo principal não é desenvolver novos FOFE, mas utilizar o estado da arte para selecionar o FOFE mais adequado para solucionar o problema proposto. A meta foi estabelecer um mecanismo de estimativa da FO preciso e eficiente que pudesse contribuir na redução do tempo e custo computacional em ambas as fases do processo de otimização hierárquico (P2-F1, P2-F2 e P2-F3).

O espaço de busca avaliado em cada etapa hierárquica dessa parte da metodologia é exatamente o mesmo que foi avaliado na Parte 1, conforme descrito nas etapas 5.1.5, 5.1.6 e 5.1.7, para garantir uma comparação justa.

5.2.1 Seleção do conjunto de dados inicial para realizar o treinamento, validação e teste das técnicas de aprendizado de máquina propostas

Durante a execução da Parte 2, o primeiro passo consistiu em selecionar as amostras para realizar o treinamento, validação e teste das técnicas de aprendizado de máquina. Para esse procedimento, foram utilizadas 100 amostras, originárias de um conjunto de 100 simulações com respectivos IECs/VPLs, obtidos de uma primeira iteração realizada utilizando os modelos de fluxo. Esses dados representam as variáveis de entrada e de saída, respectivamente,

necessárias para o treinamento da rede. Os dados foram pré-processados para garantir a compatibilidade com a estrutura da rede neural. As amostras foram divididas em 60% para treinamento, 20% para validação e 20% para teste.

Em cada etapa do processo hierárquico (P2-F1, P2-F2 e P2-F3), a natureza dos dados de entrada dessas amostras variava:

- P2-F1: os dados de entrada consistiam em um conjunto de 100 estratégias, representadas por valores binários, que indicavam se um poço seria ou não adicionado à estratégia. Os dados de saída eram os 100 IECs correspondentes a cada uma dessas estratégias;
- P2-F2: os dados de entrada incluíam 100 estratégias com variações na localização espacial do poço no reservatório e na completção das camadas a serem produzidas. Os dados de saída eram os respectivos IECs para cada uma dessas estratégias;
- P2-F3: os dados de entrada referiam-se a um conjunto de estratégias com variações do número de dias em que cada poço poderia ser aberto. Os dados de saída eram os VPLs correspondentes a cada uma dessas estratégias, conforme apresentado na Tabela 8.

Tabela 8: Descrição dos dados de entrada e de saída (FO) utilizados pelos algoritmos de aprendizado de máquina em cada etapa do processo hierárquico (Etapas 4.3, 4.4 e 4.5 do fluxograma apresentado na Figura 10).

Fase da Otimização	Dados de entrada	Dado de Saída (FO)
Otimização do número e posição dos poços (F1)	Número e posição dos poços	IEC
Ajuste fino nas posições e completção dos poços (F2)	Posições x, y, z inicial e z final dos poços pertencentes à estratégia	IEC
Otimização do cronograma de abertura dos poços (F3)	Intervalos de abertura dos poços	VPL

As estratégias iniciais utilizadas no treinamento, teste e validação da rede neural foram geradas aplicando a técnica de amostragem Hipercubo Latino, com auxílio do MERO (2022). O IEC/VPL foi obtido a partir de simulações numéricas de cada estratégia gerada, utilizando o IMEX (2019). Com os dados de produção, foi utilizada a ferramenta OTM do MERO (2022) para calcular o IEC/VPL, baseando-se no modelo econômico apresentado na Tabela 9.

5.2.2 Seleção do método de aprendizado de máquina

Para selecionar o algoritmo de aprendizado de máquina utilizado na otimização, foram realizados testes com três métodos: Floresta Aleatória, GBR e RNA. Antes de avaliar cada uma das técnicas propostas, foi realizado um *grid search* a fim de encontrar a melhor parametrização para cada técnica, otimizando os hiperparâmetros descritos na Tabela 10.

Tabela 9: Modelo econômico do estudo de caso do Campo C.

Descrição		Valor	Unidade
Preço do petróleo		314,5	US\$/m ³
Preço do gás		0	US\$/m ³
Taxa de desconto		9	%
Royalties		10	%
Impostos especiais sobre a receita bruta		9,25	%
Impostos corporativos		34	%
Custo de abandono do campo		0	-
Custo de investimento do campo		8,16E+08	US\$
Plataforma	Custo fixo	417	MM US\$
	Capacidade de processamento de líquidos	3,15	M m ³ /dia
	Capacidade de processamento de petróleo	12,2	M m ³ /dia
	Capacidade de processamento de água	3,15	M m ³ /dia
	Capacidade de injeção de água	3,15	M m ³ /dia
	Capacidade de processamento de gás	-	M m ³ /dia
	Capacidade de injeção de gás	-	M m ³ /dia
	Capacidade do número de poços	0,1	-
Custo do poço	Perfuração	4,78E+07	US\$
	Completação	5,50E+07	US\$
	Conexão	2,73E+07	US\$
	Recompletação	2,25E+07	US\$
	Conexão (poços perfurados)	2,73E+07	US\$
Lista de custos	Produção de petróleo	62,89	US\$/m ³
	Produção de água	6,29	US\$/m ³
	Produção de gás	-	US\$/m ³
	Injeção de água	6,29	US\$/m ³
	Injeção de gás	-	US\$/m ³
	Abandono	20	% do investimento no poço
Data de referência		31/12/2017 00:00	-

Tabela 10: Seleção dos hiperparâmetros utilizados na busca da melhor configuração para os algoritmos testados, baseado na literatura.

Algoritmo	Hiperparâmetros	Valores		
Floresta Aleatória	<i>n estimators</i>	100	200	300
	<i>max depth</i>	<i>None</i>	10	20
	<i>min samples split</i>	2	4	6
	<i>min samples leaf</i>	1	2	3
	<i>max features</i>	<i>sqrt</i>	<i>log2</i>	-
	<i>bootstrap</i>	<i>True</i>	<i>False</i>	-
Gradient Boosting Regressor	<i>n estimators</i>	100	200	300
	<i>learning rate</i>	0,05	0,1	0,5
	<i>max depth</i>	4	6	8
	<i>min samples split</i>	2	4	6
	<i>min samples leaf</i>	2	4	6
	<i>subsample</i>	0,6	0,8	1,0
	<i>max features</i>	<i>sqrt</i>	<i>log2</i>	-

	<i>loss</i>	<i>squared error</i>	<i>absolute error</i>	<i>huber</i>
Rede Neural Artificial (trainsgd)	<i>epochs</i>	100	150	200
	<i>goal</i>	0	0,01	0,10
	<i>max fail</i>	5	10	15
	<i>min grad</i>	0	1	10
	<i>sigma</i>	1E-05	5E-05	1E-04
	<i>lambda</i>	0,1	0,5	1

Parametrização da Rede Neural

A rede neural *Feedforward* foi configurada com uma arquitetura composta por uma camada oculta com 20 neurônios. Para o treinamento da rede, foi utilizada a função *trainsgc*.

A Figura 15 ilustra a configuração da rede neural aplicada na P2-F1 processo de otimização hierárquica, que inclui uma camada de entrada com 107 nós representando as variáveis de otimização (102 poços candidatos e cinco fixos), uma camada oculta com 20 neurônios e 1 camada de saída com um neurônio fornecendo o valor da previsão do IEC.

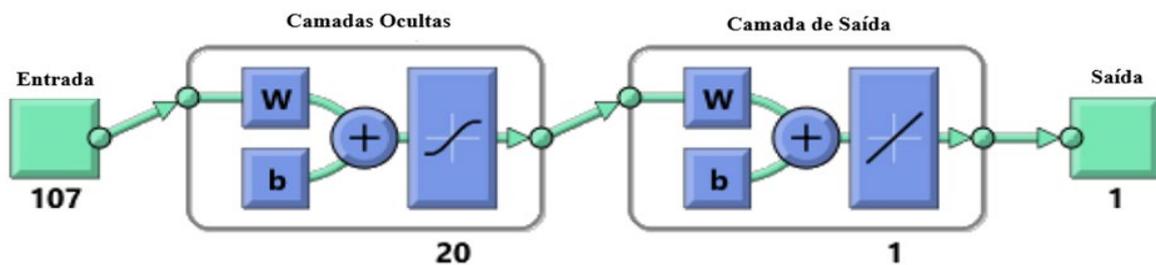


Figura 15: Configuração da rede neural utilizada para realizar a estimativa da função objetivo.

Fonte: MATLAB (2017).

Para o treinamento da rede, utilizou-se a função *trainsgc*, que implementa o algoritmo de Gradiente Conjugado Escalonado. Esse algoritmo é eficiente para redes neurais de grande porte, uma vez que realiza uma atualização dos pesos baseada em uma aproximação de segunda ordem, sem necessidade de cálculo explícito da Hessiana (Møller, 1993). A parametrização da rede neural é fundamental para garantir a convergência e a precisão do modelo no processo de otimização hierárquica aplicado ao problema em questão. O problema de otimização abordado neste trabalho pode ser formalizado matematicamente como uma minimização da função de erro quadrático médio (MSE - *Mean Squared Error*) entre as saídas previstas pela rede neural.

$$\min_{W,b} f(W,b) = MSE(W,b) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i(W,b))^2 \quad \text{Equação 6}$$

onde:

$f(W,b)$: função objetivo que representa o erro quadrático médio (MSE) entre as previsões da rede neural e os valores reais, e deve ser minimizada durante o treinamento.

W : pesos da rede neural;

b : bias da rede neural;

n : número de amostras no conjunto de dados;

y_i : valor real da i -ésima amostra;

$\hat{y}_i(W, b)$: é a previsão feita pela rede neural para a i -ésima amostra.

Restrições nas Variáveis de Decisão:

As restrições associadas ao algoritmo de treinamento *trainscg* podem ser expressas como $\|\nabla f(W, b)\| \geq \text{min_grad}$, onde *min_grad* define o limite inferior para o gradiente, abaixo do qual o treinamento é interrompido (Tabela 10).

Condições de Parada:

O treinamento é interrompido quando uma das seguintes condições é atendida:

- Meta de MSE: O treinamento para quando o MSE atinge um valor abaixo de uma meta específica:

$$f(W, b) \leq \text{goal}$$

- Número máximo de épocas: O treinamento para quando o número de épocas atinge o valor máximo permitido:

$$\text{epoch} \leq \text{max_epochs}$$

- Número máximo de falhas de validação: O treinamento para quando o número de falhas consecutivas na validação atinge o limite:

$$\text{val_fail} \leq \text{max_fail}$$

5.2.3 Critério de avaliação do algoritmo de previsão

Para avaliar o desempenho de cada técnica, foram utilizadas as métricas Coeficiente de Determinação (R^2) e o Erro Percentual Absoluto Médio Simétrico (SMAPE). O método que apresentou as melhores métricas para o treinamento, validação e teste foi a rede neural, sendo escolhido para ser utilizada na previsão da FO.

5.2.4 Seleção e parametrização do algoritmo de otimização

Nesta Parte 2 da metodologia, também foi selecionado o algoritmo de otimização IDLHC devido à sua efetividade em otimizar problemas com alta dimensionalidade. O IDLHC foi aplicado com a mesma parametrização utilizada na P1-F1. No entanto, na P2-F2 e P2-F3,

como não há garantia de que os poços e as posições selecionadas serão as mesmas da Parte 1, a parametrização do algoritmo poderá sofrer alterações em comparação com a Parte 1 da metodologia proposta.

5.2.5 Geração de novas amostras

O algoritmo IDLHC foi utilizado para gerar a estratégia inicial de treinamento da rede, assim como as estratégias de cada iteração subsequente. A geração das amostras foi conduzida de duas maneiras. Na primeira, aplicando a mesma metodologia da Parte 1, foram geradas 100 amostras a cada iteração. Na segunda, foi realizada uma modificação na parametrização do algoritmo de otimização, aumentando o número de amostras para 1.000 por iteração. Esse aumento foi possível devido ao custo computacional irrisório (segundos) para estimar a função objetivo avaliada (IEC/VPL).

5.2.6 Seleção de amostras para validação da FO prevista pelo FOFE no simulador

Durante cada iteração do processo de otimização, 20 amostras foram selecionadas para validar os valores da FO estimados pelo FOFE no simulador numérico. No cenário em que foram geradas 100 amostras, optou-se por selecionar as 20 amostras com os maiores valores de indicador econômico em cada iteração. No entanto, no cenário em que foram geradas 1.000 amostras, foi implementada uma estratégia de seleção mais rigorosa para evitar viés na alimentação do banco de dados utilizado para o retreinamento iterativo da rede neural. Nesse contexto, as 200 melhores estratégias, avaliadas com base no valor do indicador econômico, foram inicialmente filtradas. Neste contexto, inicialmente filtrou-se as 200 melhores estratégias baseadas no valor do indicador econômico. Destas, 20 amostras foram escolhidas de forma uniformemente espaçadas para assegurar uma representação equitativa da distribuição de valores. A Figura 16 ilustra esta metodologia, destacando com um retângulo tracejado vermelho a faixa de amostras que seriam selecionadas sem este critério de seleção específico, e com um retângulo verde, a faixa efetivamente explorada.

5.2.7 Critério de parada da otimização

O critério de parada estabelecido na Parte 2 da metodologia segue exatamente o mesmo aplicado na Parte 1. Essa abordagem garante uma comparação equitativa e consistente entre as metodologias estudadas.

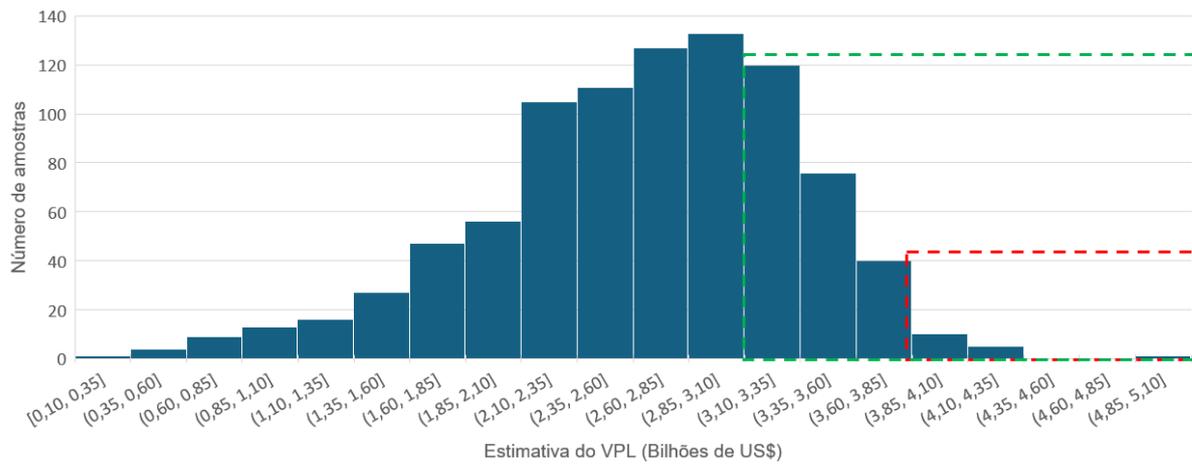


Figura 16: Histograma da distribuição das estimativas do VPL das simulações para o cenário que gerou 1.000 amostras a cada iteração da otimização indicando a estratégia adotada a fim de evitar vies na seleção das amostras.

5.2.8 Fases da otimização

Assim como na Parte 1, a otimização na Parte 2 é realizada através de um processo hierárquico. Na primeira fase, ocorre a otimização do número e do posicionamento dos poços (P2-F1). Em seguida, com base na quantidade e localização dos poços pré-selecionados, é realizado o ajuste fino tanto na posição quanto na completção desses poços (P2-F2). Por fim, a otimização do cronograma de abertura dos poços é realizada (P2-F3). O processo de otimização é considerado concluído quando a otimização de cada uma das três fases é finalizada.

6 RESULTADOS E DISCUSSÕES

Neste capítulo são apresentados e discutidos os resultados deste trabalho, divididos em duas partes. A Parte 1 apresenta os resultados da otimização utilizando apenas modelos de simulação. O objetivo foi obter uma estratégia de produção ótima aplicando as melhores práticas de simulação para reduzir o número de simulações necessárias para completar esse estudo. Essa abordagem foi adotada para garantir uma comparação justa com os resultados obtidos posteriormente usando o FOFE. A Parte 2 trata da mesma problemática abordada na Parte 1, porém com um foco na aceleração do processo de otimização com o uso de FOFE e de simulação. Nesta fase, as partes mais custosas do processo foram substituídas por um estimador analítico baseado em técnicas de aprendizado de máquina. Essa substituição teve como objetivo principal a redução de tempo e custo computacional necessários para a otimização, mantendo a eficiência e a eficácia ao longo do processo.

6.1 Parte 1 - Otimização com RSM e IDLHC

A primeira parte apresenta a otimização utilizando os RSM em conjunto com o IDLHC, mostrando o mapa de qualidade e as regiões onde foram definidas as posições candidatas na etapa 5.1.5. Em seguida, é discutida a evolução da otimização hierárquica ao longo das três fases propostas, finalizando com a análise do tempo necessário para otimizar a estratégia de produção. A Figura 17 apresenta três mapas de qualidade com a distribuição das posições candidatas dos poços.

6.1.1 Otimização da posição e número dos poços

A Figura 18 apresenta o gráfico de produção de líquidos dos quatro poços produtores já perfurados. Ao analisar o gráfico, foi possível identificar que a capacidade da plataforma opera sobre seu limite máximo após o tempo de referência até um período de aproximadamente 4.100 dias, conforme indicado pela linha tracejada vertical vermelha. Após a definição das posições candidatas, foi aplicado o algoritmo IDLHC para encontrar a configuração ideal do número e posição dos poços. A restrição na capacidade de processamento de líquidos mostrou ser um fator impactante no número de poços da estratégia.

A Figura 19 ilustra o IEC de cada simulação em função do número total de poços (produtores e injetores), destacando as duas estratégias que alcançaram o maior IEC. Os

resultados indicam uma tendência de melhoria nos valores de IEC com a redução do número de poços, indicado pela flecha verde. As estratégias mais eficazes variam entre oito e nove poços incluindo tanto poços já existentes quanto produtores e injetores adicionados às estratégias, conforme indicado pelas flechas vermelhas.

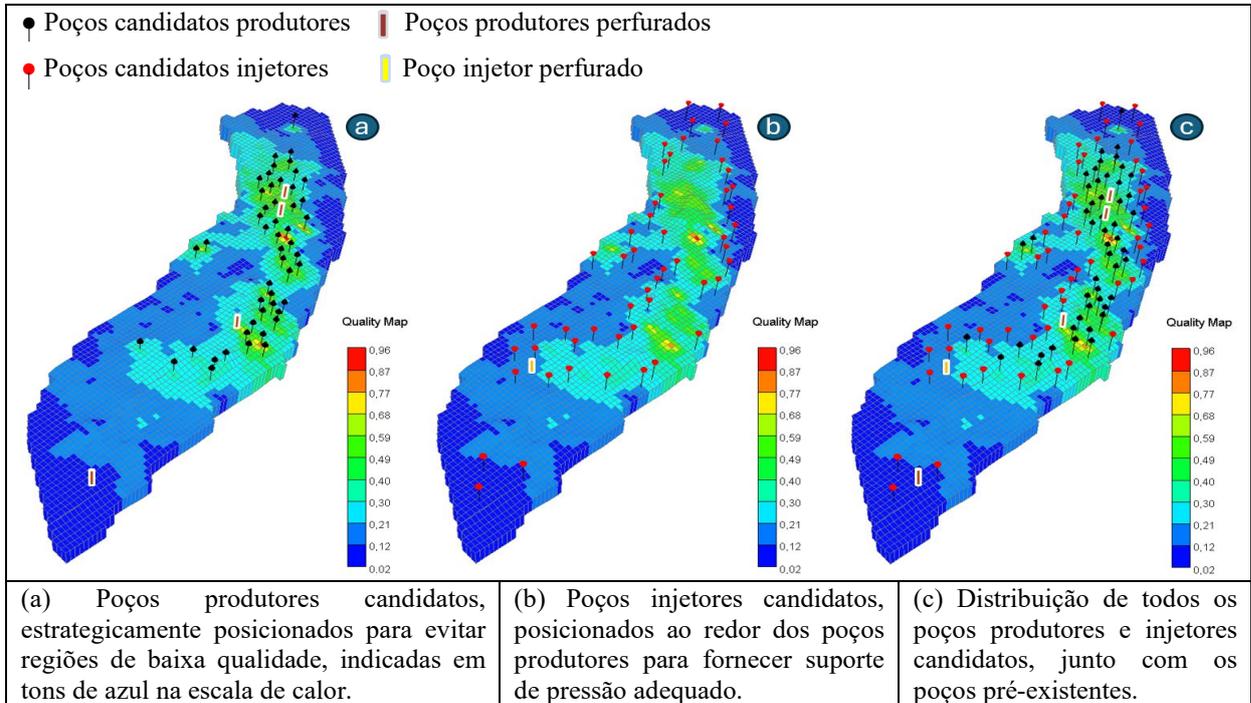


Figura 17: Definição das posições candidatas no reservatório.

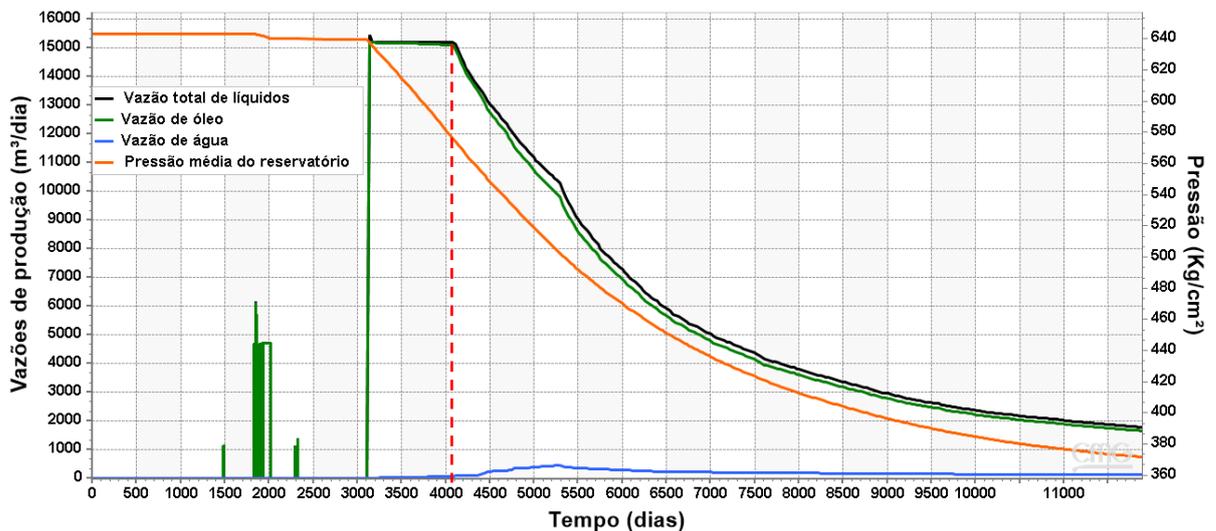


Figura 18: Curvas de vazão de produção total de líquidos, óleo e água dos quatro poços produtores já perfurados no campo, indicando o início do declínio da produção pela linha tracejada.

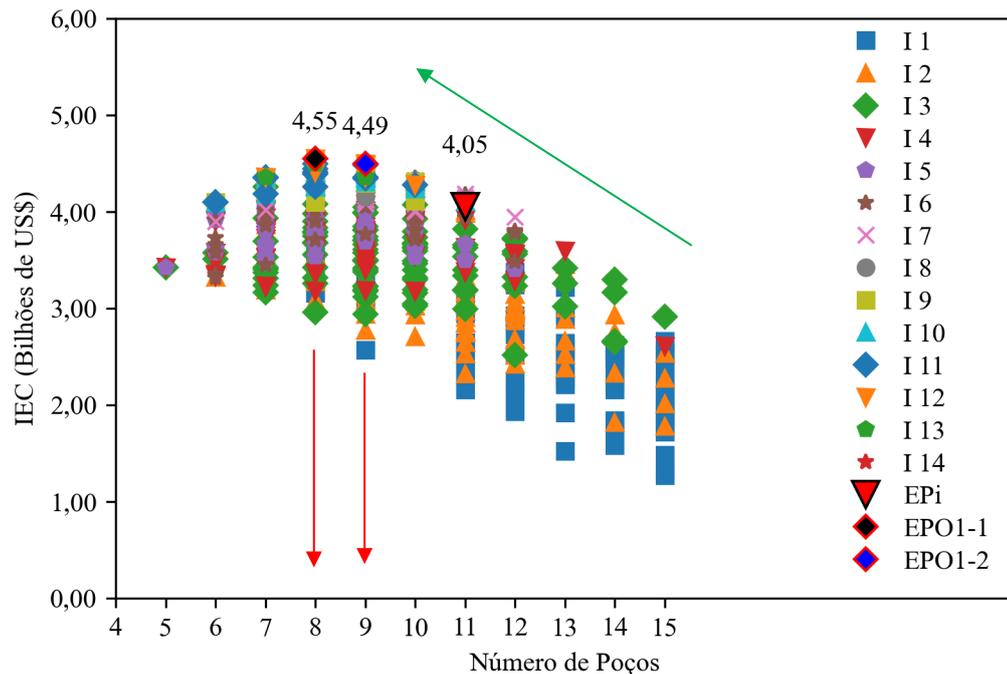


Figura 19: Variação do IEC em função do número de poços ao longo de 14 iterações do processo de otimização. A estratégia de produção inicial (EPI) é representada por um triângulo vermelho invertido, definido como caso base. As duas estratégias com melhores indicadores econômicos (EPO1-1 e EPO1-2), alcançadas na primeira etapa do processo hierárquico, são destacadas por losangos preenchidos em preto e azul contornados em vermelho, respectivamente.

As curvas de vazão de produção, apresentadas na Figura 20, mostram que as estratégias com oito e nove poços justificam a tendência observada no gráfico anterior. Na estratégia com nove poços (cinco produtores e quatro injetores), a plataforma trabalha com a sua capacidade máxima durante todo o período de extrapolação (11.888 dias), conforme destacado pela linha contínua preta no gráfico. Já na estratégia com oito poços (cinco produtores e três injetores), a capacidade é quase totalmente utilizada durante todo o tempo simulado (linha tracejada preta).

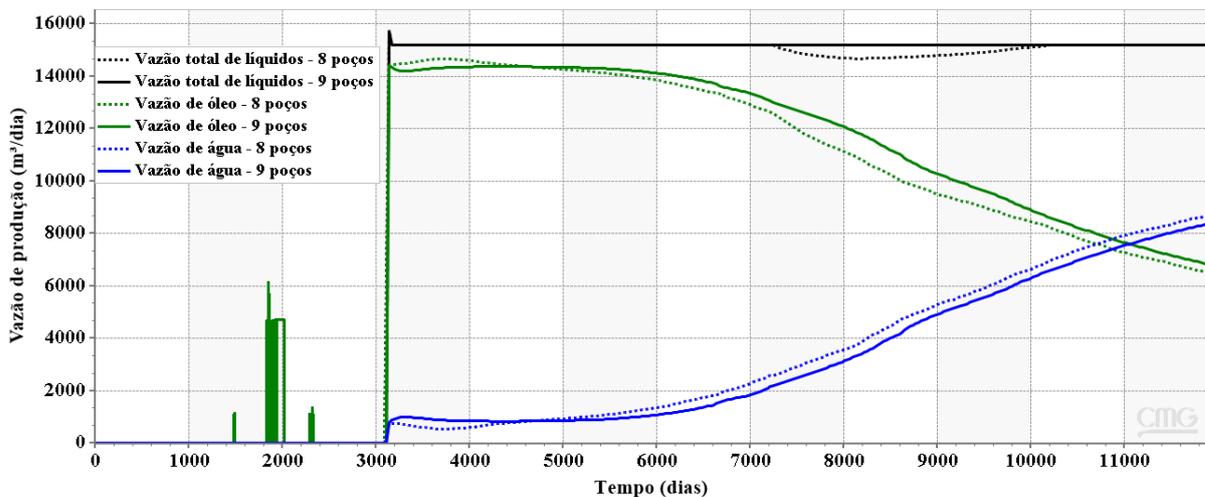


Figura 20: Curvas de vazão de produção de líquidos, óleo e água das duas estratégias que retornaram o melhor IEC da otimização do número e posição dos poços na P1-F1.

Na Tabela 11 é possível identificar que embora a estratégia com nove poços apresente melhores indicadores de Np Wp e FRo, o IEC ficou 1,3% menor (equivalente a 60 milhões de dólares) que na estratégia com oito poços.

Tabela 11: Diferença do IEC e indicadores de produção (Np, Wp e FRo) entre a estratégia com nove e oito poços da P1-F1.

Estratégia	IEC (bilhões de US\$)	Np (m ³)	Wp (m ³)	FRo (%)
8 poços	4,55	1,01E+08	3,23E+07	14,8
9 poços	4,49	1,04E+08	3,01E+07	15,3
Δ (%)	-1,3	3,1	-6,6	3,1

O caso adotado como base para avaliar a evolução da otimização utilizando os RSM e IDLHC é a EPI, que foi desenvolvida por Mello *et al.* (2024) utilizando conhecimentos de Engenharia de Reservatórios e alcançou um IEC de 4,05 bilhões de dólares. A comparação com a EPI permitiu avaliar o ganho da estratégia otimizada na Parte 1 da metodologia, em termos de valor da função objetivo. A Figura 21 apresenta uma comparação do posicionamento dos poços entre a EPI e a EP otimizada nesta fase (EPO1-P1).

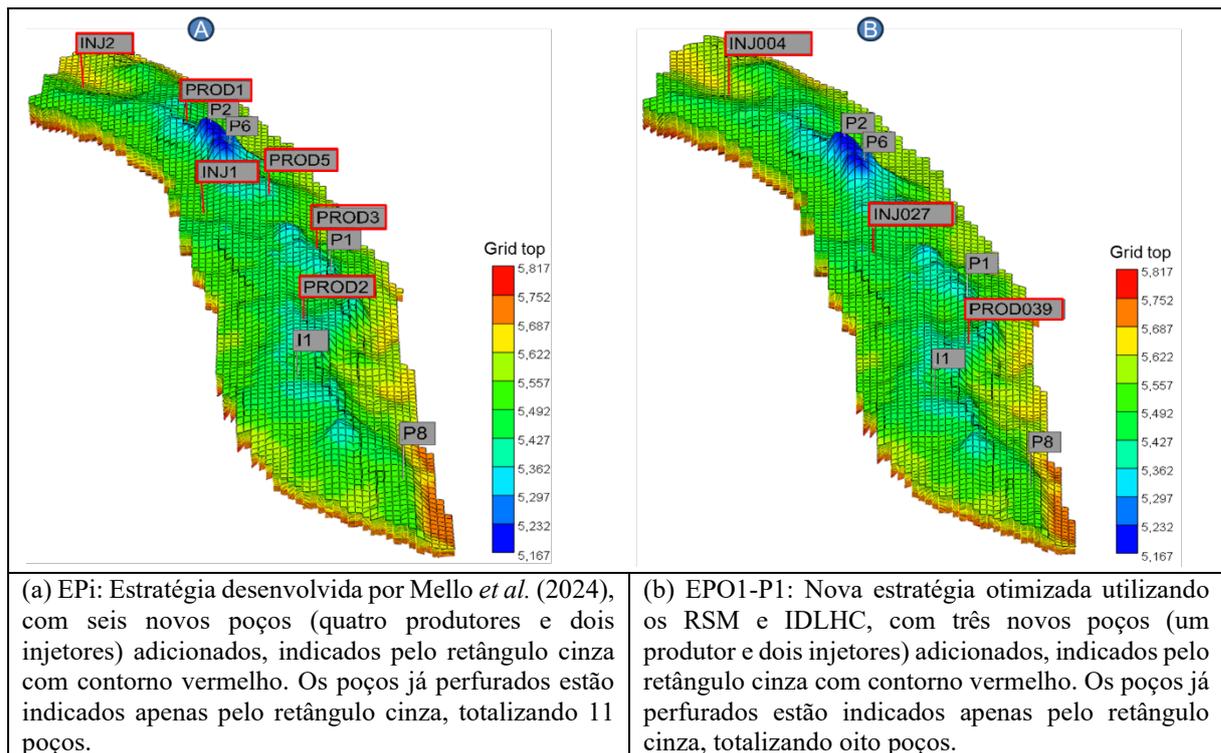


Figura 21: Comparação entre a distribuição dos poços na EPI e na EPO1.

Ao longo das iterações, o processo de otimização utilizando RSM e IDLHC mostrou uma evolução significativa na busca pela melhor estratégia. Cada iteração envolveu, inicialmente, a simulação de 100 estratégias com configurações de poços distintas, buscando maximizar o valor do IEC ao longo de sucessivas iterações.

Na primeira iteração, uma das estratégias testadas já apresentou um IEC superior ao da EPi. No entanto, na segunda iteração, o maior IEC apresentou um leve declínio, ficando um pouco abaixo do valor alcançado pela EPi, por ser um método populacional, não há garantia de que a otimização da FO seja contínua ao longo das iterações devido a exploração contínua do espaço de busca. Isso ocorre porque esses algoritmos buscam evitar ficar presos em ótimos locais e aumentar a chance de encontrar o ótimo global (Mardle e Pascoe, 1999). Através da Figura 22, é possível observar o comportamento evolutivo da FO ao longo das iterações, típico de um processo de otimização. Na 6ª iteração, foi alcançada a melhor estratégia, que resultou em um IEC de 4,55 bilhões de dólares. Este valor representa um ganho de aproximadamente 12%, equivalente a 500 milhões de dólares (Equação 7), em relação à EPi (4,05 bilhões de dólares).

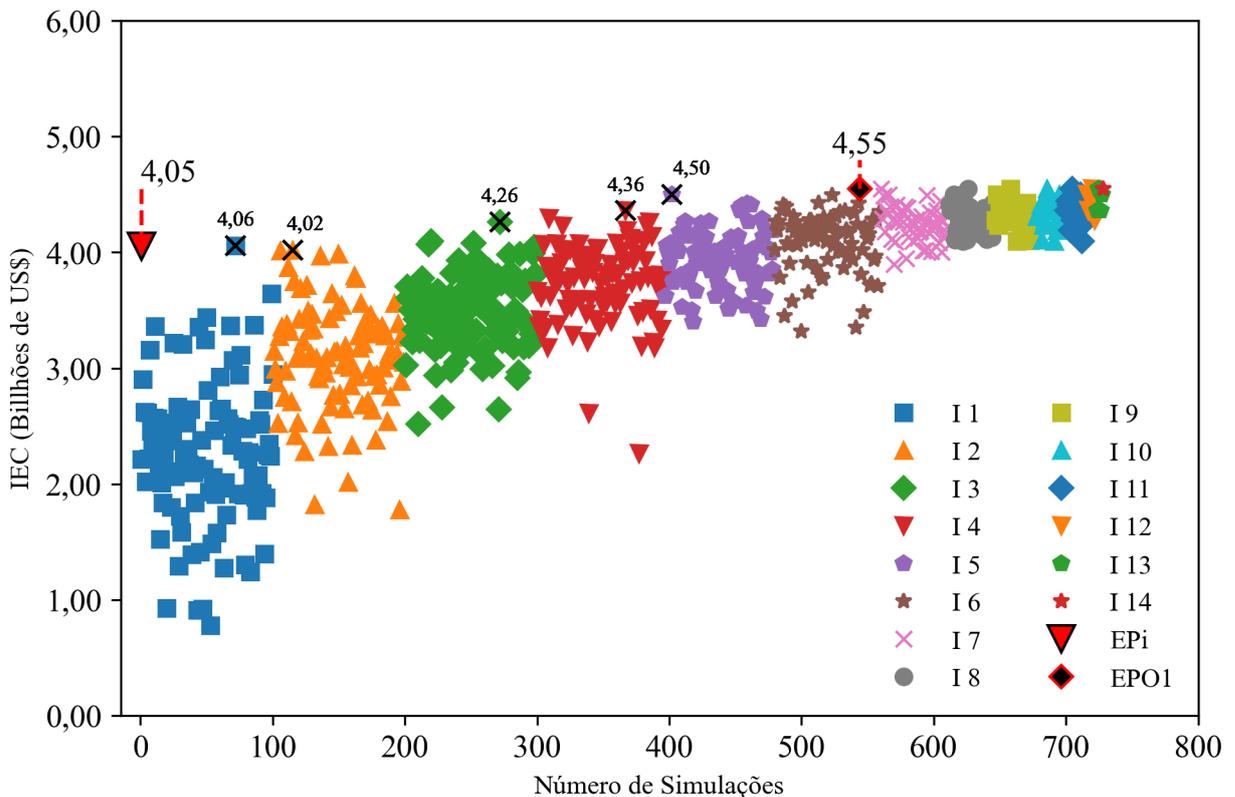


Figura 22: Evolução do IEC ao longo das simulações nas diferentes iterações, com a melhor estratégia (EPO1) identificada na 6ª iteração.

É importante destacar que a Parte 1 da metodologia buscou maximizar a FO da forma mais eficiente possível, aplicando conhecimentos de engenharia de reservatórios e otimização. Nesse sentido, além dos dois critérios de parada que fazem parte da configuração do método de otimização (número mínimo de níveis e número máximo de iterações), um terceiro critério foi adicionado e implementado para reduzir o número máximo de simulações necessárias para obter a EP otimizada. Esse critério envolve a quantificação do ganho (incremento) da FO

escolhida em um intervalo de três iterações. Ou seja, caso não ocorra uma melhora da FO ao longo de três iterações subsequentes, o método é finalizado pelo terceiro critério.

Esses critérios de parada são essenciais para garantir a eficiência do processo de otimização, evitando simulações desnecessárias e economizando tempo e recursos computacionais. A Figura 23 ilustra um caso prático aplicado na P1-F1 para validar a eficiência do critério de parada estabelecido, onde a otimização foi executada somente com os dois critérios originais até que eles fossem violados e interrompessem automaticamente o processo.

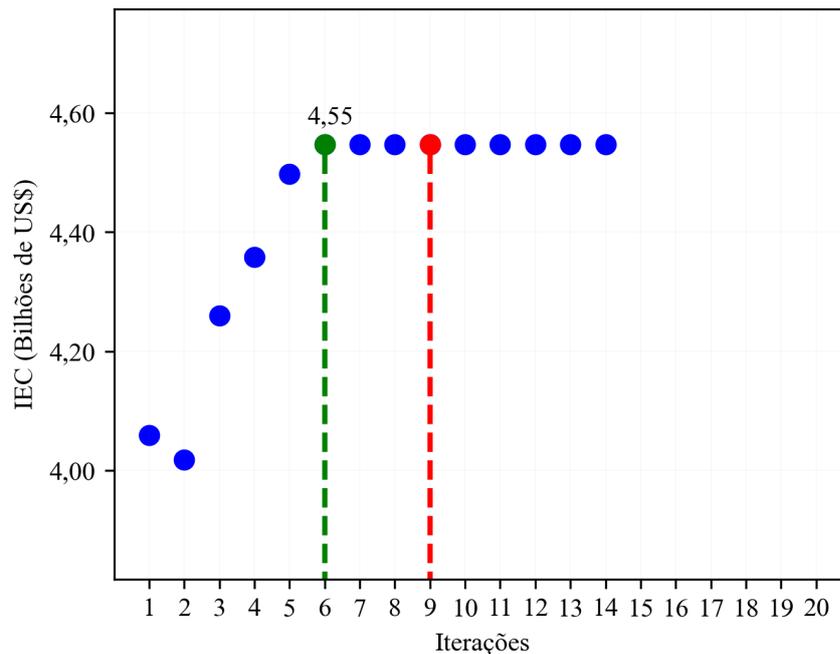


Figura 23: Critério de parada indicando o momento de interrupção do processo.

Conforme observado na Figura 23, o processo rodou até a 14ª iteração. No entanto, com o terceiro critério proposto, o processo pararia na nona iteração (destacada pela linha vertical vermelha tracejada), que é o momento em que foi constatado não haver ganho significativo na FO em três iterações consecutivas, após encontrar a melhor resposta na sexta iteração (indicada pela linha vertical verde tracejada), confirmando a eficiência do critério adicional definido.

Nessa primeira fase do processo hierárquico, considerando que a otimização foi interrompida na nona iteração, foi necessário executar 677 simulações para obter a EPO1. O incremento total do IEC é dado pela diferença entre o IEC da EPO1 e o IEC da EPi, conforme destacado na Equação 7, cada simulações contribuiu para incrementar o IEC em aproximadamente 740 mil dólares (Equação 8).

$$IT_{IEC (P1-F1)} = 4,55 \times 10^9 \text{ US\$} - 4,05 \times 10^9 \text{ US\$} = 500 \times 10^6 \text{ US\$} \quad \text{Equação 7}$$

$$IPS = \frac{IT_{IEC}}{N^{\circ} \text{ de simulações}} = \frac{500 \times 10^6 \text{ US\$}}{677} \cong 740 \times 10^3 \text{ US\$} \quad \text{Equação 8}$$

onde:

IT_{IEC} é o incremento total do Indicador Econômico de Campo na etapa do processo hierárquico;
 IPS é o Incremento por simulação.

6.1.2 Ajuste fino da posição dos poços

Nessa etapa de ajuste fino, há uma redução significativa no espaço de busca em comparação com a fase anterior, onde foram definidos o número e as posições ideais dos poços. Com o objetivo de maximizar a FO, foi realizado um ajuste fino tanto nas posições quanto na completção dos três novos poços adicionados à estratégia. O poço injetor INJ004 foi deslocado quatro blocos e os poços INJ027 e PROD039 foram deslocados dois blocos, conforme indicado pela seta na Figura 24. Esse deslocamento foi realizado considerando apenas o movimento horizontal.

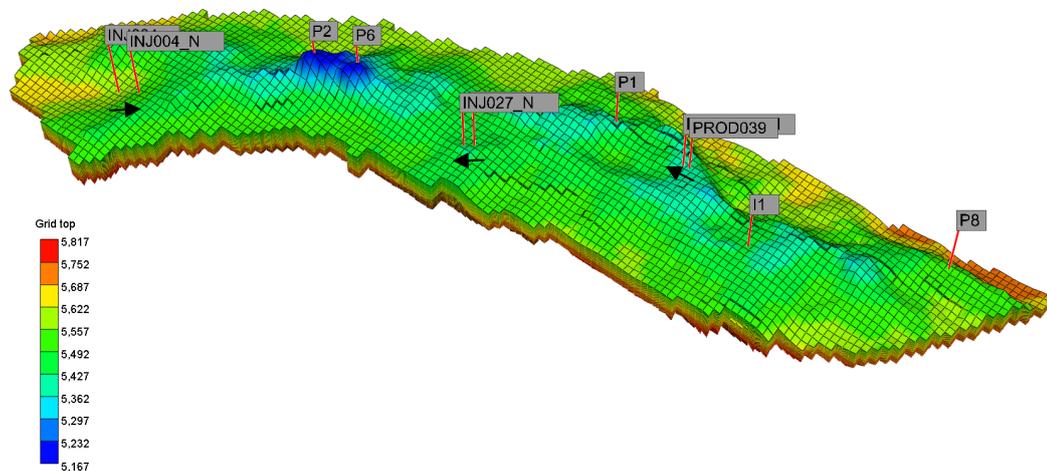


Figura 24: Indicação do deslocamento dos poços após o ajuste fino.

Além disso, foi realizado um ajuste fino na orientação vertical, conforme ilustrado na Figura 25, indicando o antes (a) e o depois (b). A otimização após esse ajuste fino (EPO2) resultou em um gerenciamento aprimorado da completção dos poços, deslocando-os para regiões superiores do reservatório, afastando-os da região do contato água/óleo, o que contribuiu para a redução na produção de água.

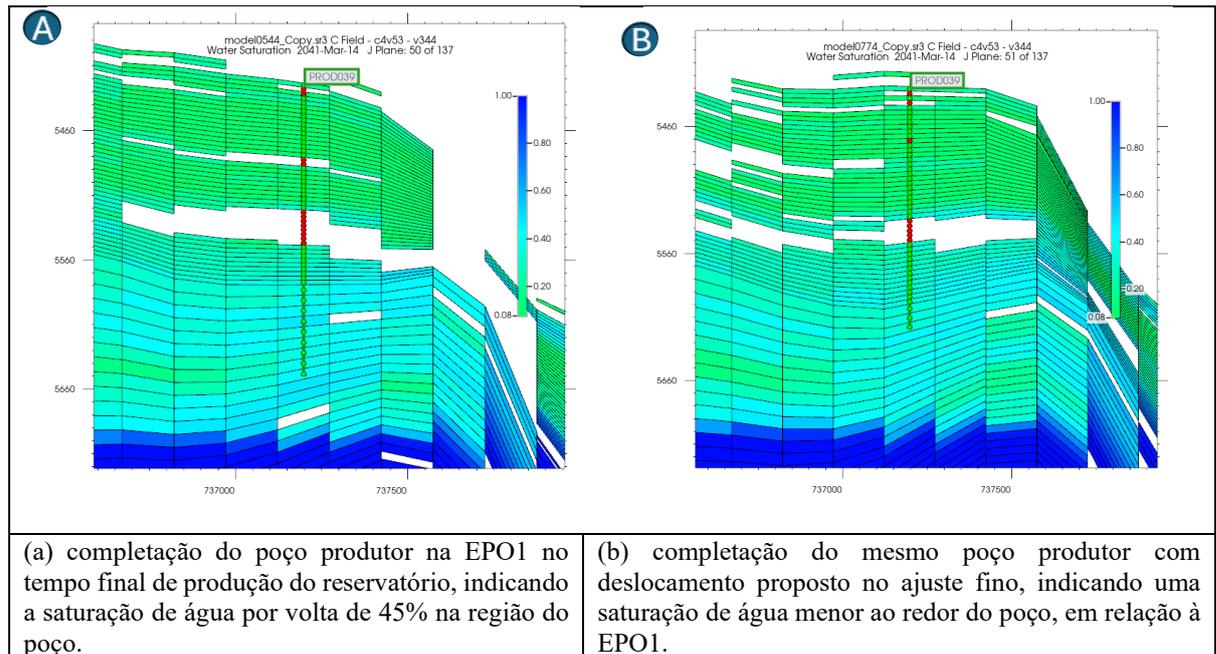


Figura 25: Comparação da saturação de água na região do poço antes e após o ajuste fino na posição dos poços adicionados à estratégia.

Na otimização da EPO2, essa alteração também contribuiu para a manutenção da pressão média do reservatório e para o aumento na produção acumulada de óleo, conforme ilustrado na Figura 26.

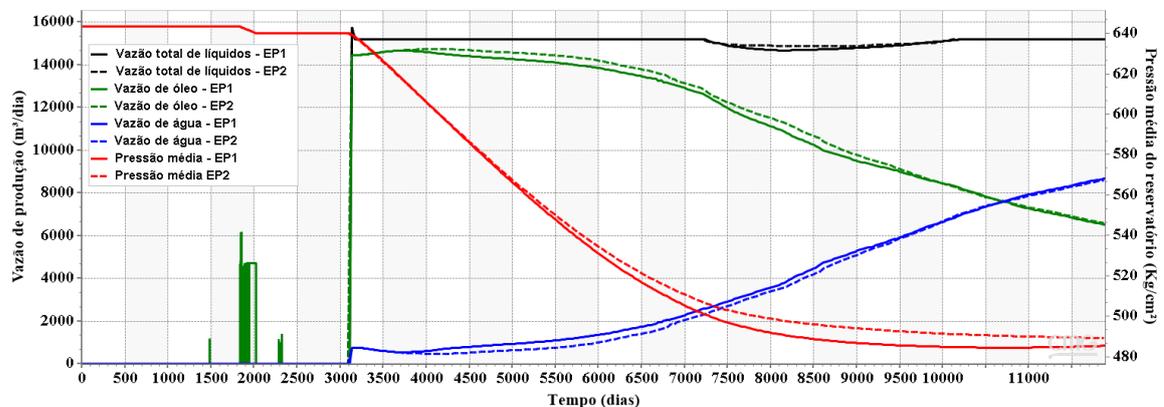


Figura 26: Comparação entre a EPO1 (linha contínua) e EPO2 (linha tracejada) indicando que o ajuste fino atuou de forma positiva no gerenciamento da produção de líquidos, promovendo uma redução na produção de água e maximização na produção de óleo.

A Figura 27 demonstra que a estratégia ideal ocorreu na nona iteração, alcançando um IEC de 4,64 bilhões de dólares. Essa fase da otimização proporcionou um aumento de 14,5% em relação ao caso base, representando aproximadamente 590 milhões de dólares adicionais. Em comparação com a fase anterior, a evolução do IEC foi menos acentuada em termos percentuais, promovendo um ganho de 2%, o que representa cerca de 90 milhões de dólares, conforme destacado na Equação 9. Apesar do aumento abaixo do esperado no IEC e do número adicional de simulações, o objetivo era mostrar o ganho real ao fazer o ajuste fino.

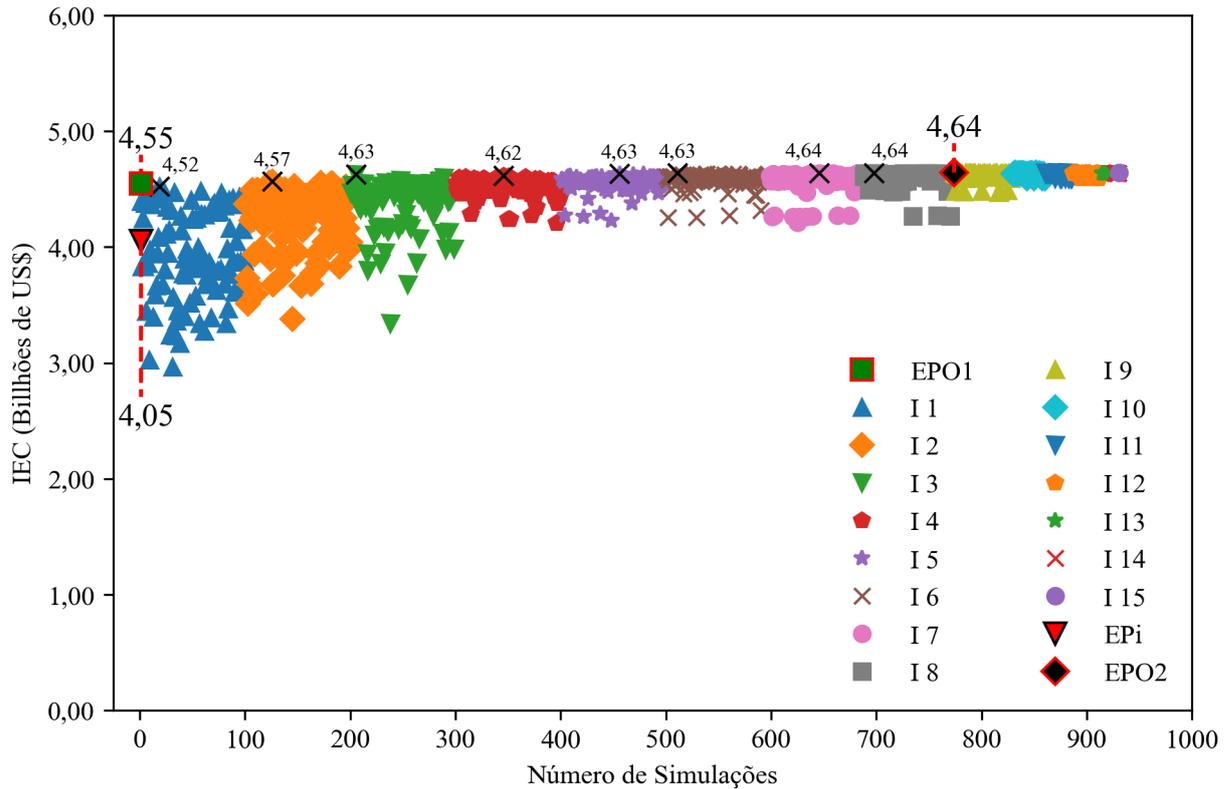


Figura 27: Evolução do IEC ao longo das simulações nas diferentes iterações, com a melhor estratégia identificada na nona iteração.

A Figura 28 destaca a iteração que resultou no melhor IEC. Com o critério original, a simulação rodou até a 15ª iteração. No entanto, com o terceiro critério adicionado, o processo pararia na décima segunda iteração, apresentando uma redução significativa no que tange ao número de simulações.

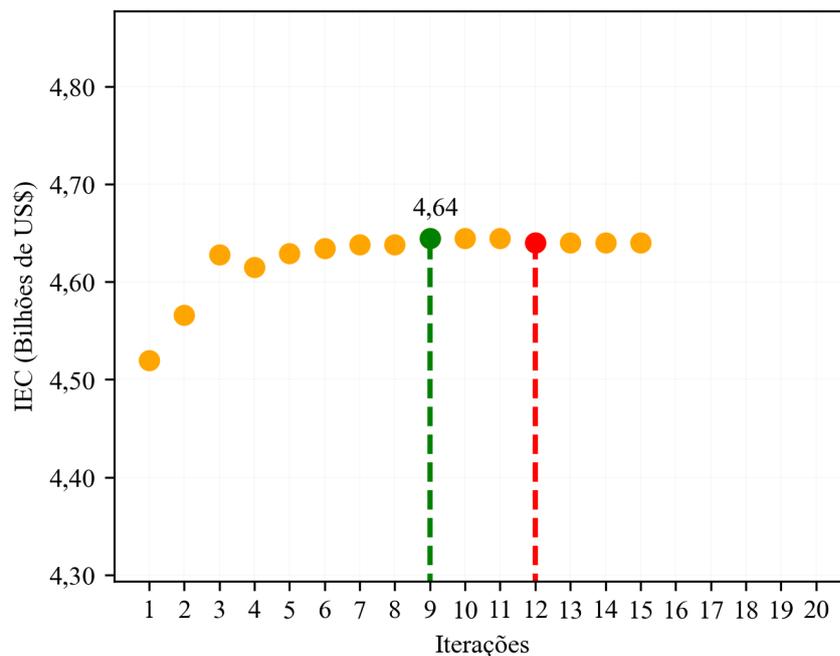


Figura 28: Evolução do IEC ao longo das iterações indicando o momento em que o processo deve ser interrompido para evitar realizar simulações desnecessárias.

Ao final da P1-F2, foi necessário executar 914 simulações para obter a EPO2. Utilizando a Equação 9, calculamos o incremento total do IEC nesta fase. A partir da Equação 10, é possível concluir que cada simulação contribuiu para incrementar o IEC em aproximadamente 98 mil dólares, um valor que equivale a um número 7,5 vezes menor do que na P1-F1.

$$IT_{IEC (P1-F2)} = 4,64 \times 10^9 \text{ US\$} - 4,55 \times 10^9 \text{ US\$} = 90 \times 10^6 \text{ US\$} \quad \text{Equação 9}$$

$$IPS = \frac{IT_{IEC}}{N^{\circ} \text{ de simulações}} = \frac{90 \times 10^6 \text{ US\$}}{914} \cong 98 \times 10^3 \text{ US\$} \quad \text{Equação 10}$$

6.1.3 Otimização do cronograma de abertura dos poços

Essa última fase da otimização hierárquica estabeleceu o cronograma de abertura para os três poços adicionados à estratégia. Para os poços injetores foram atribuídos 31 níveis, representando os possíveis intervalos de abertura, enquanto o poço produtor foi limitado a apenas três níveis, conforme especificado na Tabela 7. É importante destacar que, sempre que possível, é essencial colocar o poço produtor em operação o mais cedo possível, uma vez que este contribui diretamente para a geração de receita do projeto. Além disso, nesta etapa do processo hierárquico, a função objetivo foi modificada de IEC para VPL, refletindo a inclusão do cronograma de abertura dos poços.

A Figura 29 ilustra a evolução da otimização do cronograma de abertura dos poços ao longo das simulações, comparando os resultados entre a EPI, EPO1 e EPO2. Observa-se que, embora a variabilidade das respostas da FO seja menor nesta fase, o ganho obtido em relação à EPO2 foi maior do que o ganho da EPO2 em relação à EPO1. Isso demonstra a importância de um cronograma bem gerido para maximizar o VPL, principalmente quando a capacidade da plataforma é limitada, o que reflete a necessidade de um melhor gerenciamento na produção dos fluidos. Neste caso, a melhor estratégia alcançou um VPL de 4,77 bilhões de dólares.

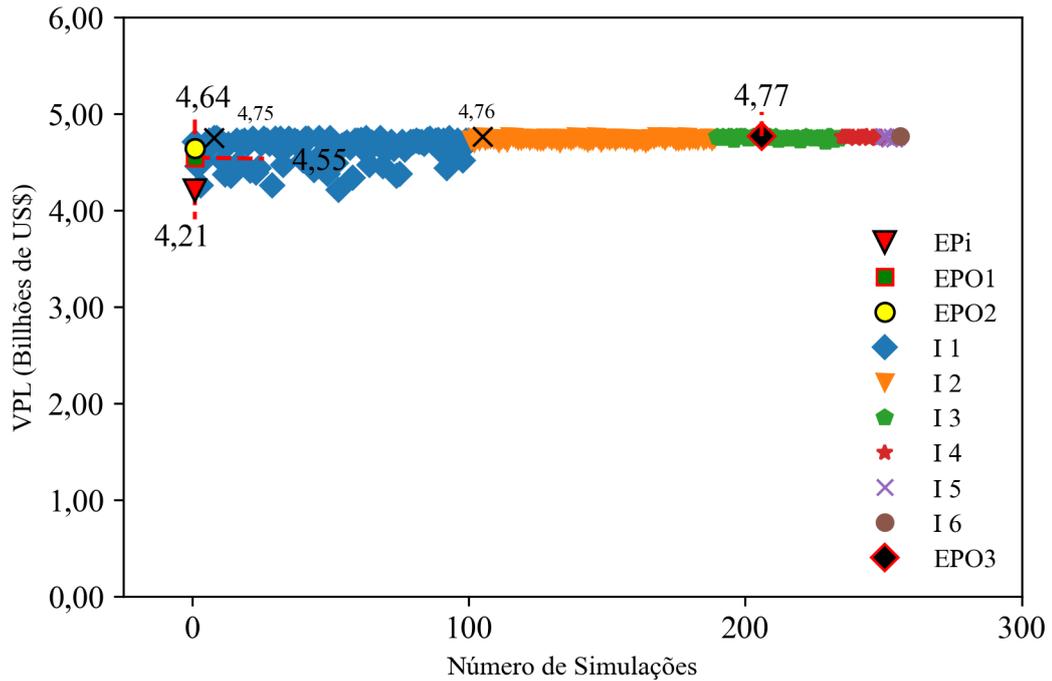


Figura 29: Evolução da otimização do cronograma de abertura dos poços, comparando os resultados com a EPI, EPO1 e EPO2.

A última etapa da otimização hierárquica destaca através do número baixo de iterações, na Figura 30, que a otimização alcançou uma zona de convergência. A melhor estratégia foi obtida na terceira iteração, indicada pela linha tracejada verde, culminando na interrupção do processo na sexta iteração, tanto pelo critério de parada indicando não haver ganho em 3 iterações consecutivas, quanto alcançar o número mínimo de níveis.

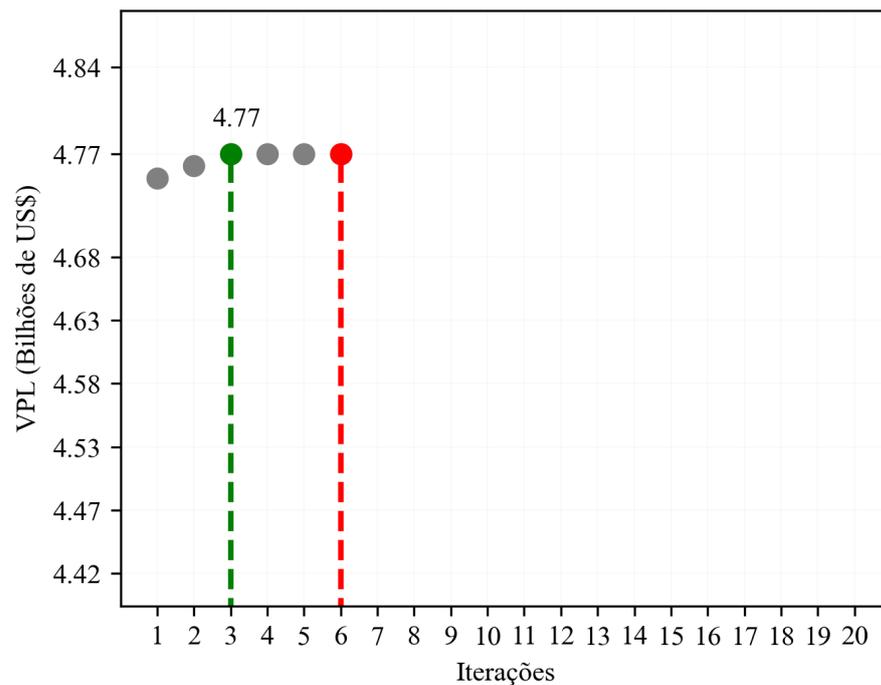


Figura 30: Evolução do VPL ao longo das iterações do processo de otimização, destacando o momento de convergência e o fim da otimização nas iterações iniciais.

Seguindo o critério de parada proposto neste trabalho, podemos concluir que nesta etapa (P1-F3) foi necessário executar 256 simulações para obter a EPO3. Utilizando a Equação 11, calculamos o incremento total do VPL nesta fase. A partir da Equação 12, é possível concluir que cada simulação contribuiu para incrementar o VPL em aproximadamente 508 mil dólares, um valor que demonstra um ganho expressivo quando comparado à etapa anterior (P1-F2), cerca de cinco vezes maior.

$$IT_{VPL(P1-F3)} = 4,77 \times 10^9 \text{ US\$} - 4,64 \times 10^9 \text{ US\$} = 130 \times 10^6 \text{ US\$} \quad \text{Equação 11}$$

$$IPS = \frac{IT_{VPL}}{N^{\circ} \text{ de simulações}} = \frac{130 \times 10^6 \text{ US\$}}{256} \cong 508 \times 10^3 \text{ US\$} \quad \text{Equação 12}$$

onde:

IT_{VPL} é o incremento total do VPL na etapa do processo hierárquico;

IPS é o Incremento por simulação.

A Tabela 12 fornece uma visão abrangente dos resultados da otimização, apresentando VPL, N_p , W_p e FR_o . Antes da implementação do cronograma de abertura de poços, o indicador utilizado durante a otimização é o IEC, conforme discutido na fundamentação teórica (Item 2.7.2). No entanto, para garantir uma comparação justa da diferença no retorno econômico (ΔVPL) apresentado, foi calculado o VPL também para a EP_i .

Tabela 12: Evolução das funções de objetivo ao longo das fases hierárquicas do processo de otimização comparadas com o caso base (EP_i).

Caso	IEC / VPL (10^9 US\$)		N_p (10^8 m ³)	W_p (10^7 m ³)	FR_o (%)	ΔVPL (%)
EP_i	4,05	4,21	0,950	3,43	13,96	-
EPO1	4,55		1,01	3,23	14,84	12,3
EPO2	4,64		1,03	3,08	15,10	14,5
EPO3	4,77		1,02	2,77	14,99	13,30

A partir de uma análise na Tabela 12, observa-se que o processo de otimização apresentou uma melhoria significativa no indicador econômico avaliado, passando de 4,21 (EP_i) para 4,77 bilhões de dólares (EPO3). Para alcançar a estratégia mais eficaz, foram necessárias 1.847 simulações. Além disso, olhando para as EPO2 e EPO3, a implementação do cronograma de abertura de poços causou uma redução na produção acumulada de óleo, um resultado esperado devido ao adiamento da produção. Uma avaliação sob a mesma perspectiva para os poços injetores indica um gerenciamento otimizado (redução) da produção de água, resultando em um aumento significativo do retorno econômico durante a P1-F3. Idealmente, a

otimização hierárquica deveria incluir uma etapa de revalidação, especialmente ao lidar com mudanças em variáveis que possuem grande impacto nos resultados. No entanto, neste estudo, o foco principal foi na aplicabilidade do FOFE para otimização, o que justificou a ausência dessa revalidação. Recomenda-se que estudos futuros incluam essa etapa para garantir estratégias de otimização ainda mais refinadas.

6.2 Parte 2 - Otimização com FOFE, RSM e IDLHC

A segunda parte aborda a otimização aplicando o FOFE, na qual um algoritmo de aprendizado de máquina foi selecionado com base em métricas definidas para avaliação, indicando seu potencial de aplicação. Essa parte da metodologia aplicou a mesma parametrização das variáveis otimizadas utilizada na Parte 1. Assim como na primeira parte, é apresentada a evolução da otimização ao longo do processo hierárquico, finalizando com a análise do tempo consumido nesta parte da metodologia e uma avaliação final ao utilizar a técnica proposta, que é o principal objetivo deste trabalho. Todos os resultados apresentados foram validados por meio de um simulador numérico comercial, reforçando a confiabilidade das respostas obtidas.

6.2.1 Seleção do algoritmo de aprendizado de máquina

A primeira tarefa da metodologia na Parte 2 foi selecionar a técnica de aprendizado de máquina a ser empregada. Utilizando o mesmo conjunto de dados inicial, foi avaliado o desempenho de três técnicas: Floresta Aleatória, GBR e RNA. O objetivo foi identificar a técnica com o melhor poder preditivo dado um conjunto inicial de dados. A RNA apresentou os melhores ajustes ao avaliar em conjunto as métricas R^2 e SMAPE para treinamento (60% das amostras), validação (20% das amostras) e teste (20% das amostras), conforme destacado na Figura 31. Quanto maior o valor de R^2 , mais o modelo é capaz de explicar a variação dos dados em torno de sua média; quanto menor o valor do SMAPE, mais precisa é a previsão em termos de erro percentual.

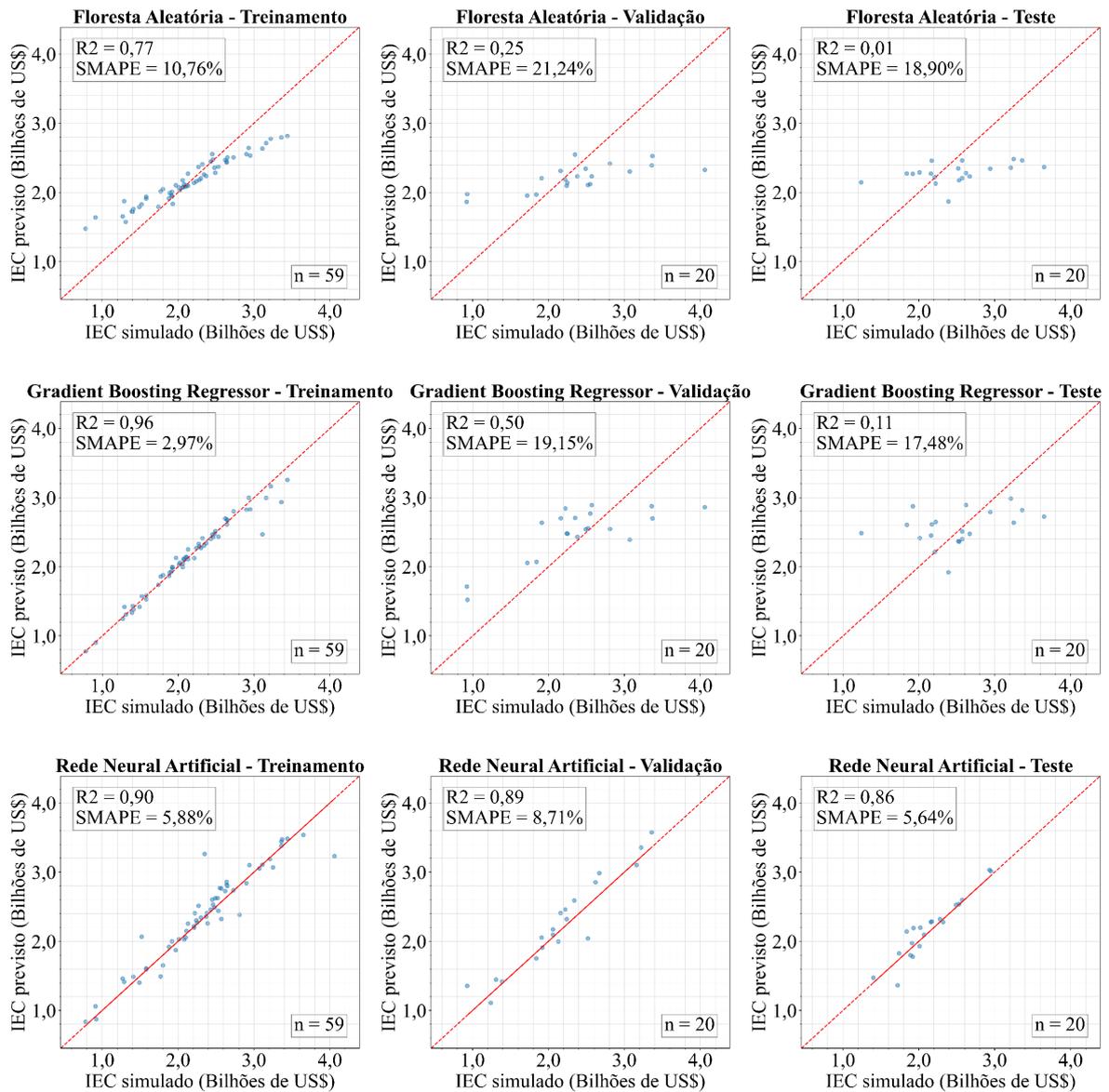


Figura 31: Comparação entre as métricas R^2 e SMAPE do treinamento, validação e teste nos três métodos de aprendizado de máquina avaliados, para identificar o que melhor captura o comportamento dos dados e realiza estimativas mais precisas.

Após definir a rede neural como técnica para estimar o valor da função objetivo, ela foi utilizada para substituir parte do processo de otimização. Durante cada iteração, os valores de todo o conjunto de amostras geradas pelo IDLHC foram estimados pelo FOFE. Diferente da metodologia tradicional aplicada na Parte 1, onde o conjunto inteiro de amostras geradas era simulado, apenas 20 amostras de cada iteração foram simuladas.

As amostras simuladas eram então adicionadas iterativamente ao conjunto inicial de treinamento, a fim de retreinar a rede neural e melhorar seu poder de estimativa. Esse procedimento foi realizado nas três fases do processo hierárquico.

A Rede Neural Artificial pode ter demonstrado melhor desempenho devido à sua capacidade de capturar relações complexas e não-lineares nos dados, o que é uma vantagem significativa em cenários onde essas características predominam (Goodfellow *et al.*, 2016). Além disso, as RNA ajustam dinamicamente os pesos entre as camadas durante o treinamento, permitindo um ajuste fino que supera as abordagens baseadas em árvores, como a utilizada pelo Random Forest e o Gradient Boosting Regressor (Haykin, 2009; Breiman, 2001). A estrutura em camadas das RNA também oferece uma maior capacidade de modelagem, capturando nuances que os outros métodos podem não identificar (LeCun; Bengio; Hinton, 2015).

6.2.2 Otimização da posição e número dos poços FOFE – 1000 amostras

Como o custo computacional da estimativa da FO é irrisório, nessa primeira avaliação foram geradas 1.000 amostras em cada iteração. Dessas, foram filtradas as 200 melhores, baseadas na FO avaliada, e selecionadas 20 igualmente espaçadas para estimar o valor da FO com o FOFE. Se as 20 melhores amostras baseadas na FO fossem selecionadas, os valores resultantes seriam muito próximos, reduzindo a representatividade da variabilidade. A Figura 32 apresenta essas amostras, comparando a evolução do IEC ao longo das iterações.

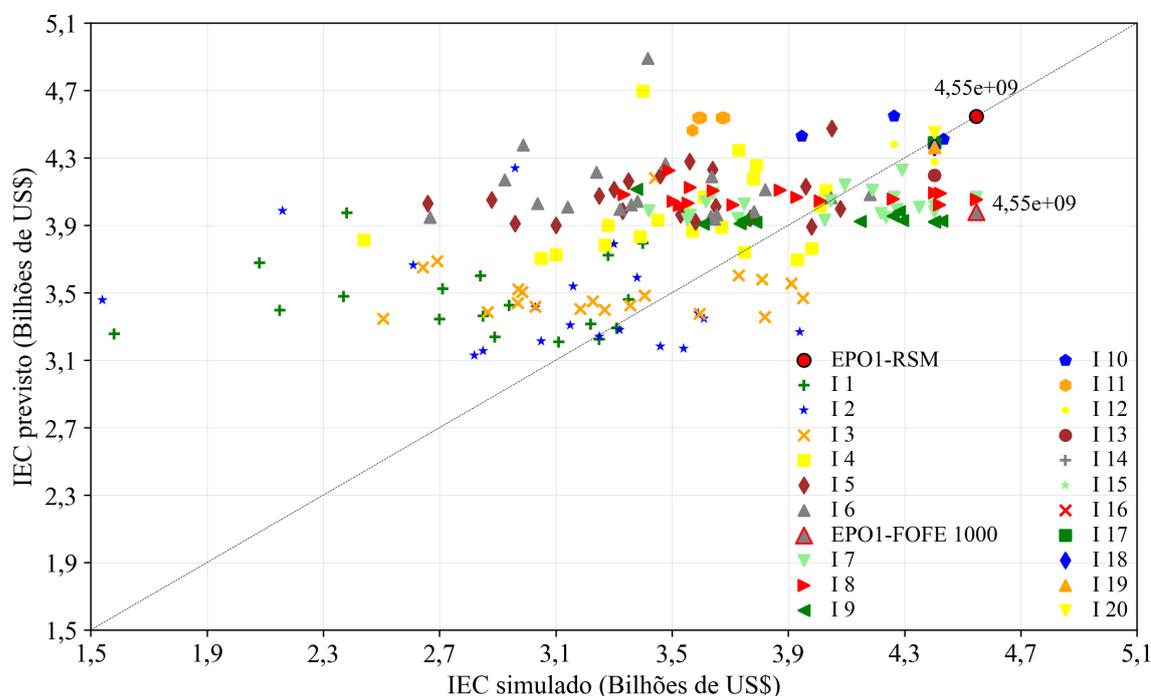


Figura 32: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do número e posição dos poços na P2-F1, com base na otimização inicial utilizando os RSM desenvolvida na P1-F1 (destacada com um círculo vermelho sobre a linha identidade).

Na Figura 32, é possível observar a dispersão dos pontos em relação à linha de referência (a linha onde o valor estimado é igual ao valor simulado). Inicialmente, os pontos estão mais

dispersos da linha de referência, indicando um menor poder de reprodução das estimativas obtidas pelo simulador.

À medida que o processo de otimização avança, as redes neurais são retreinadas de forma dinâmica, incorporando novos dados a cada iteração. As 20 amostras selecionadas para validação no simulador são incluídas no banco de dados inicial de treinamento, proporcionando, dessa forma, um refinamento contínuo do modelo de estimativa.

No decorrer das iterações, o algoritmo de otimização IDLHC vai refinando o espaço de busca, facilitando encontrar as melhores soluções. Embora na sexta iteração o poder de estimativa do FOFE ainda não esteja acurado, é nesta iteração que o processo atinge a melhor estratégia, destacada por um triângulo cinza com contorno vermelho. Esta iteração apresenta a mesma configuração de poços da EP otimizada na P1-F1, que resulta no maior valor de IEC estimado pelo FOFE.

A Figura 33 destaca, com uma linha vertical verde tracejada, que a melhor estratégia foi alcançada na iteração seis, retornando um IEC de 4,55 bilhões de dólares. Para identificar o momento ideal para interromper o processo de otimização e torná-lo mais eficiente, aplicou-se o mesmo critério de parada utilizado na Parte 1 da metodologia. Após a sexta iteração, não houve melhora na função objetivo avaliada, indicando que o processo poderia ser interrompido na nona iteração, conforme destacado pela linha vertical vermelha tracejada.

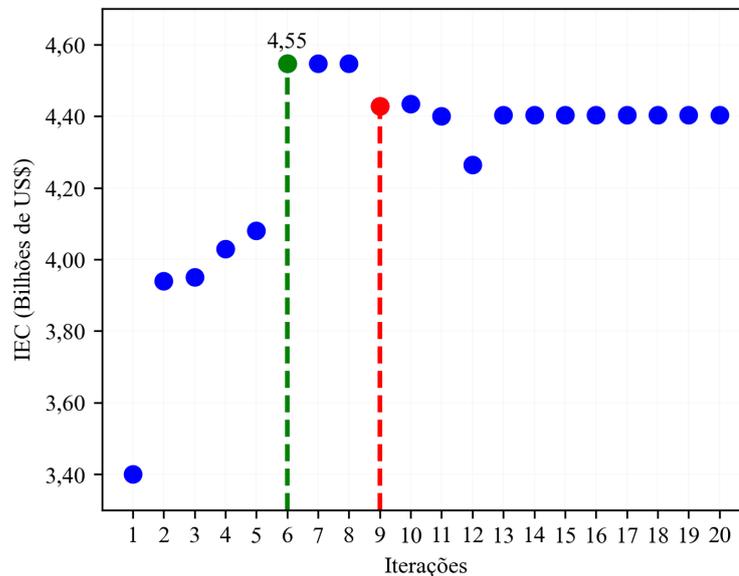


Figura 33: Critério de parada indicando o momento de parada do processo de otimização do número e posição dos poços (F2-P1).

A contabilização do número de amostras utilizadas nesta etapa consiste nas 100 amostras utilizadas para treinar a rede neural, mais as amostras validadas no simulador numérico. Como a indicação de que o processo deve ser interrompido ocorre na nona iteração,

consideraram-se mais 180 simulações (20 amostras validadas por iteração). Isso totaliza 280 amostras utilizadas durante esta fase de otimização.

6.2.3 Ajuste fino na posição dos poços FOFE – 1000 amostras

Nesta fase do processo de otimização, foi utilizada uma RNA treinada especificamente com a parametrização definida para o ajuste fino da posição dos poços. De forma similar à Parte 1, mas com um número de amostras igual a 1.000 para estimar os valores de IEC, o mesmo critério de seleção das melhores estratégias foi seguido para realizar a validação com os dados estimados pelo simulador.

Durante esta fase, a análise dos pontos na Figura 34 revela que, ao longo das iterações, as estimativas tornam-se mais precisas e próximas da linha de referência. Isso reflete a eficácia do treinamento contínuo da RNA, que se adapta e evolui à medida que novas amostras são incorporadas ao processo de retreinamento.

É importante ressaltar que, apesar dessa maior precisão, o maior valor previsto pelo FOFE nesta fase ainda é inferior ao valor obtido com os RSM. Este melhor resultado é atingido na décima iteração, conforme destacado na Figura 34. A redução observada é de aproximadamente 0,9%, o que corresponde a uma diferença de cerca de 40 milhões de dólares. O valor de referência obtido com os RSM está circulado em vermelho, enquanto o melhor com a FOFE está destacado com um pentágono azul com contorno vermelho.

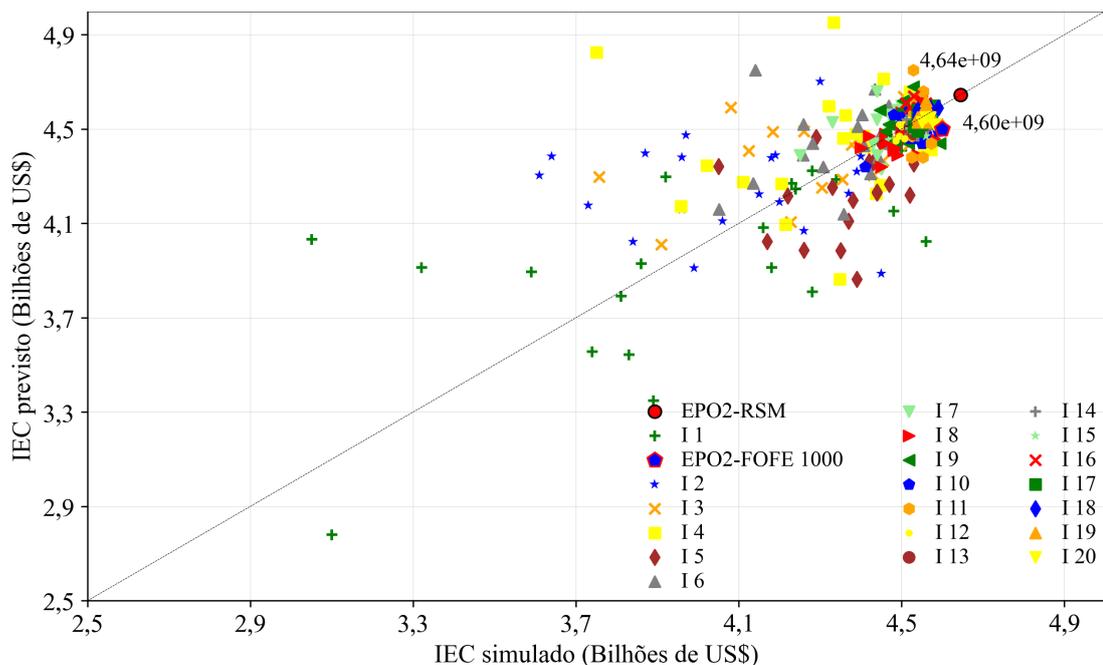


Figura 34: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução do ajuste fino na posição e completção dos poços na P2-F2, utilizando como referência o resultado obtido na otimização utilizando os RSM na P1-F2 (destacado com um círculo vermelho sobre a linha identidade).

Apesar da dispersão dos pontos, principalmente nas iterações iniciais, o processo de otimização vai refinando o espaço de busca e convergindo para configurações de poços cada vez mais eficientes. A melhor estratégia, identificada na décima iteração, demonstra como o processo é capaz de encontrar soluções que, ainda que sejam um pouco inferiores às encontradas utilizando RSM, representam uma otimização significativa do ajuste fino na posição e completção dos poços, aumentando em torno de 53 milhões de dólares, o valor da FO em comparação com o IEC alcançado na EPO1 da P2-F1.

Nesta segunda fase, a melhor estratégia foi identificada na iteração 10, com um IEC de 4,60 milhões de dólares, seguindo o critério de parada estabelecido (parada na 13ª iteração), conforme destacado na Figura 35. Assim, foram contabilizadas as 100 simulações iniciais utilizadas para o treinamento da rede neural, mais 260 simulações realizadas ao longo do processo, totalizando 360 simulações.

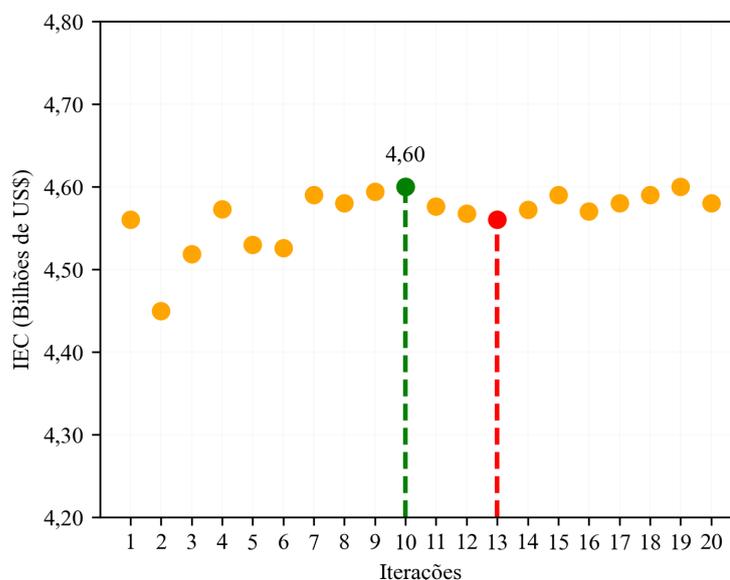


Figura 35: Evolução do valor da função objetivo ao longo da otimização, indicando o momento de interrupção do ajuste fino na posição e completção dos poços após encontrar a melhor resposta e não obter melhora nas três iterações subsequentes.

6.2.4 Otimização do cronograma de abertura dos poços FOFE – 1000 amostras

Nesta etapa do processo de otimização, as amostras geradas e os critérios de seleção para a validação no simulador permaneceram os mesmos. A Figura 36 apresenta uma comparação dos valores estimados pela FOFE e pelo simulador ao longo das etapas P2-F3 e P1-F3, respectivamente.

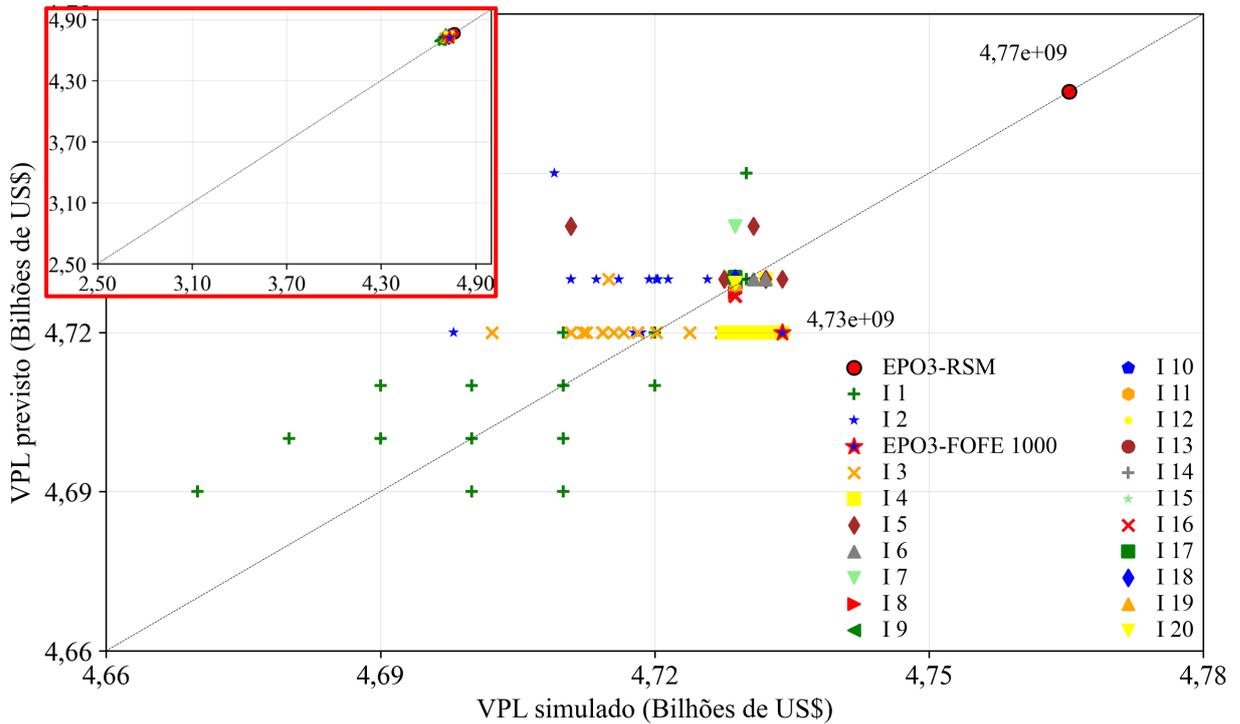


Figura 36: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do cronograma de abertura dos poços na P2-F3, tomando com base o resultado da otimização conduzida na P1-F3 (destacado por um círculo vermelho posicionado sobre a linha identidade).

Na Figura 36, é possível observar que o espaço de busca foi reduzido de forma significativa, de modo que houve a necessidade de ampliar a escala do gráfico para identificar com maior clareza o comportamento das amostras. Esse nível de detalhe revela uma significativa redução na dispersão dos pontos, com as estimativas realizadas pelo FOFE cada vez mais próximas dos valores estimados pelo simulador.

Entretanto, mesmo com essa melhora na precisão, o valor de referência (EPO3-RSM), representado pelo círculo vermelho sobre a linha de referência, não foi alcançado. A diferença observada é de aproximadamente 0,8%, o que corresponde a uma redução de cerca de 36 milhões de dólares.

A melhor estratégia nesta fase final foi identificada na segunda iteração e está destacada com uma estrela azul com contorno vermelho. A estratégia, apesar de não alcançar o valor de referência dos RSM, demonstra a eficácia do processo de otimização e o potencial de melhoria contínua com o uso do modelo FOFE. Esta fase evidencia a capacidade do modelo em se ajustar e melhorar progressivamente, mesmo com um espaço de busca reduzido.

A última fase da otimização hierárquica apresenta um cenário diferente, onde o impacto das variáveis está reduzido, resultando em uma baixa variabilidade da FO ao longo do processo. A Figura 37 mostra o resultado do VPL máximo ao longo de cada iteração.

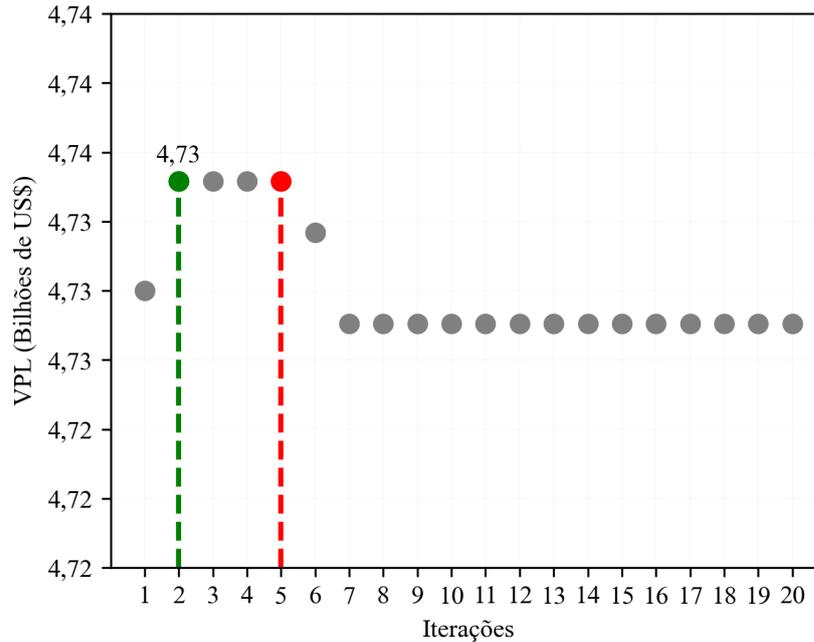


Figura 37: Evolução do valor da função objetivo ao longo da otimização, indicando o momento de interrupção do processo após violar o critério de parada na otimização do cronograma de abertura dos poços.

Na Figura 37 observa-se que a melhor estratégia foi alcançada na iteração 10, representada pela linha tracejada verde, com um valor de 4,73 bilhões de dólares. Nota-se que houve uma redução na ordem de grandeza do VPL (eixo das ordenadas), em comparação com os gráficos apresentados na P2-F1 e P2-F2, para demonstrar a variação do VPL máximo ao longo de cada iteração.

A fim de garantir uma comparação justa entre as Partes 1 e 2 da metodologia geral e avaliar o impacto do número de amostras geradas em cada iteração durante a otimização, foi conduzida uma nova otimização. Nesta, gerou-se 100 amostras por iteração, um número equivalente ao aplicado no estudo com RSM e IDLHC, conforme descrito a seguir.

6.2.5 Otimização da posição e número dos poços FOFE – 100 amostras

Na primeira fase hierárquica do processo de otimização, houve uma modificação na configuração da melhor estratégia obtida. Neste caso, foram selecionados quatro novos poços para a melhor estratégia, além dos cinco já perfurados, dos quais três são injetores e um produtor. Esta configuração está representada na Figura 38 através de um quadrado amarelo com contorno vermelho.

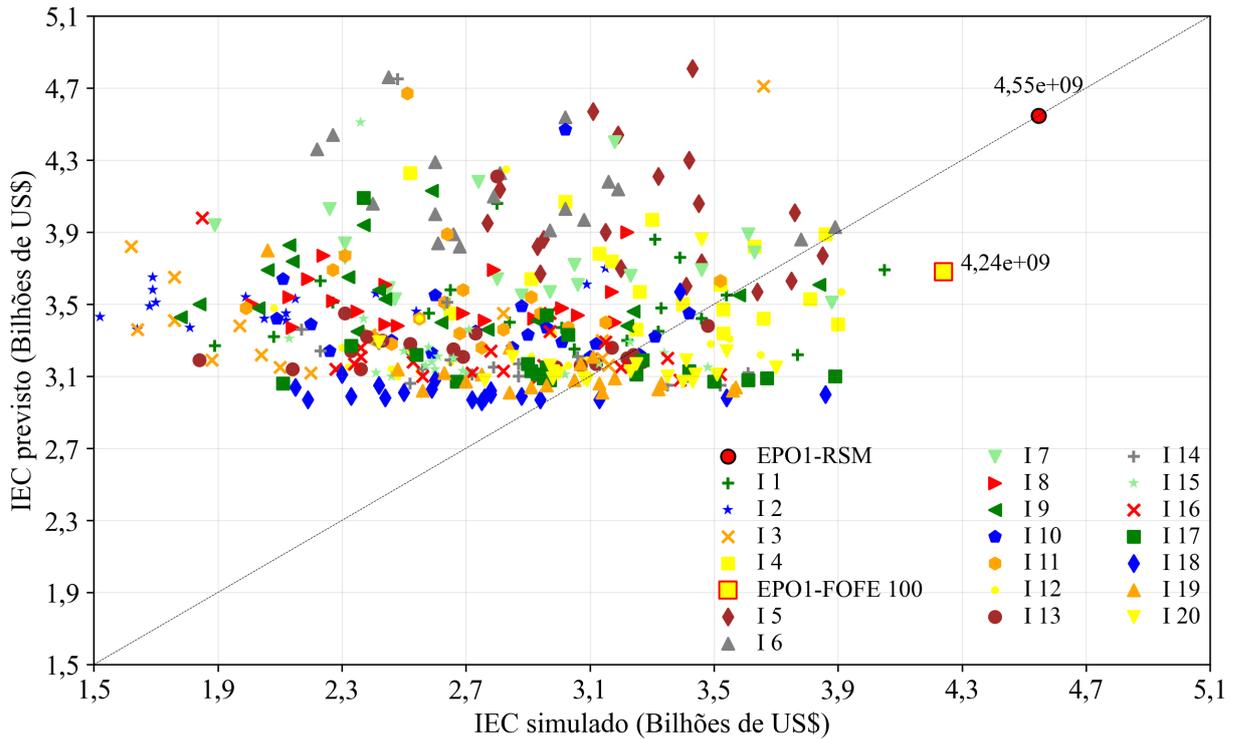


Figura 38: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização do número e posição dos poços na P2-F1, com base na otimização utilizando os RSM desenvolvida na P1-F1 (destacada com um círculo vermelho sobre a linha identidade).

Nota-se na Figura 38 que houve uma redução de aproximadamente 7% no valor do IEC da melhor estratégia utilizando FOFE em comparação com a otimização utilizando os RSM e IDLHC (EPO1). A melhor estratégia com FOFE atingiu um valor de 4,24 bilhões de dólares na quarta iteração, após um total de 240 simulações (100 de treinamento e 140 de validação da estimativa da RNA, considerando três iterações referentes à aplicação do critério de parada). Em contrapartida, a estratégia EPO1 (círculo vermelho com contorno preto) obteve um valor de 4,55 bilhões de dólares.

6.2.6 Ajuste fino na posição dos poços FOFE – 100 amostras

A segunda fase do processo hierárquico conseguiu gerenciar adequadamente a estratégia com um poço adicional. A otimização promoveu um ganho de 0,1% no valor do IEC, em comparação com a EPO2-RSM superando também a EPO2-FOFE 1000. A Figura 39 destaca a estratégia EPO2 através de um círculo vermelho com contorno preto, e um valor de IEC de 4,64 bilhões de dólares.

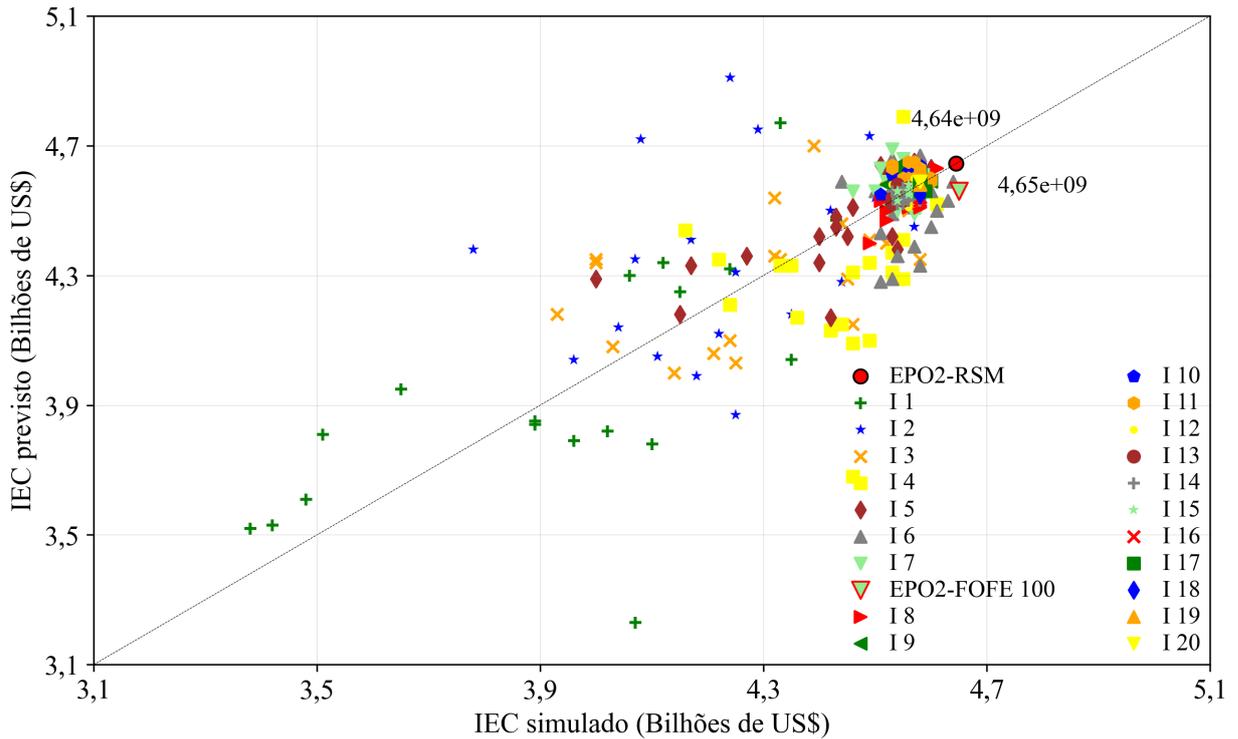


Figura 39: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução do ajuste fino na posição e completção dos poços na P2-F2, utilizando como referência o resultado obtido na otimização aplicando os RSM na P1-F2 (destacado com um círculo vermelho sobre a linha identidade).

Nesta fase, a melhor estratégia, representada por um triângulo verde invertido com contorno vermelho, atingiu um valor de 4,65 bilhões de dólares na sétima iteração. Para alcançar esse resultado, foram realizadas 300 simulações, das quais 100 foram utilizadas para o treinamento da rede neural e 200 para a validação das estimativas realizadas pelo FOFE, considerando três iterações subsequentes referentes à aplicação do critério de parada.

Através da Figura 39, é possível também observar uma tendência de alinhamento das estimativas do FOFE com os valores simulados à medida que avança o número de iterações. Essa convergência sugere que a melhor estratégia se aproxima cada vez mais da linha de referência, indicando um aumento na precisão da estimativa do IEC.

6.2.7 Otimização do cronograma de abertura dos poços FOFE – 100 amostras

A otimização do cronograma de abertura dos poços nesta fase também apresentou baixa variabilidade nas respostas. A última fase do processo hierárquico obteve resultados muito próximos aos alcançados pela otimização que utilizou RSM e IDLHC (EPO3), representada por um círculo vermelho com contorno preto e um VPL de 4,77 bilhões de dólares.

A melhor estratégia foi alcançada na quinta iteração, encontrando um valor de 4,81 bilhões de dólares, representada na Figura 40 por um losango marrom com contorno vermelho.

Esta fase apresentou um ganho de aproximadamente 1% no VPL em comparação com a estratégia EPO3. A Figura 40 destaca a proximidade dos valores previstos pelo modelo FOFE em relação aos valores simulados, indicando boa precisão nas estimativas.

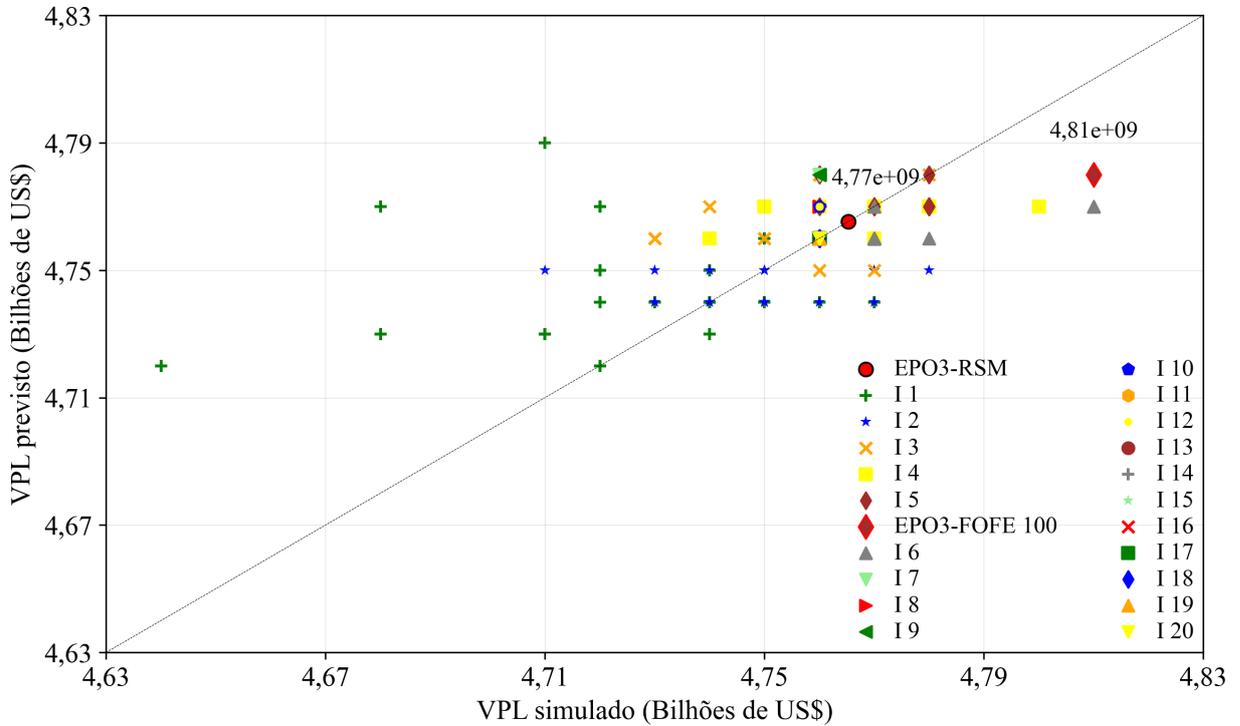


Figura 40: Comparação entre os valores previstos pelo FOFE (eixo y) vs. os valores simulados (eixo x), indicando a evolução da otimização no cronograma de abertura dos poços na P2-F3, tomando com base o resultado da otimização conduzida na P1-F3 n(destacado por um círculo vermelho posicionado sobre a linha identidade).

6.3 Tempo computacional

Após concluir a otimização hierárquica utilizando as duas abordagens da metodologia, Partes 1 e 2, realizou-se uma avaliação do desempenho do FOFE. Esta comparação teve como objetivo demonstrar a eficiência do FOFE em acelerar o processo de obtenção de uma estratégia de produção otimizada, reduzindo significativamente o número de simulações necessárias.

Na Parte 2, a contabilização das simulações envolveu tanto as simulações realizadas ao longo do processo quanto o tempo computacional, que foi convertido em número equivalente de simulações, necessário para o treinamento da rede neural em cada fase do processo hierárquico. Esse cálculo permitiu uma comparação justa com a otimização conduzida na Parte 1 da metodologia, conforme apresentado na Tabela 13.

Tabela 13: Comparação do tempo computacional em relação às execuções de simulação (realizadas no simulador mais as equivalentes), destacando o aumento de velocidade obtido com o uso do FOFE.

Número de simulações executadas no simulador numérico e cálculo das simulações equivalentes ao tempo de treinamento da RNA			
		RSM e FOFE-100	RSM e FOFE-1000
Fase 1: Número e posições	Iteração final	7	9
	Tempo de treinamento (min)	62,67	75,25
	Número de simulações equivalentes	5	6
	Número de simulações executadas	240	280
Fase 2: Ajuste fino	Iteração final	10	13
	Tempo de treinamento (min)	80,48	93,36
	Número de simulações equivalentes	7	8
	Número de simulações executadas	300	360
Fase 3: Cronograma de abertura	Iteração final	8	5
	Tempo de treinamento (min)	69,36	46,16
	Número de simulações equivalentes	6	4
	Número de simulações executadas	260	200

1 simulação de reservatório = 13 min

Objetivo final - Número total de simulações em cada parte do processo de otimização da produção				
	Fase 1 Número e posições dos poços	Fase 2 Ajuste fino	Fase 3 Cronograma de abertura	Total de Simulações
RSM	677	914	256	1.847
RSM e FOFE 100	245	307	266	818
RSM e FOFE 1000	286	368	204	858

Aceleração total do processo:	FOFE 1000 amostras	54%	FOFE 100 amostras	56%
--------------------------------------	--------------------	------------	-------------------	------------

Comparando os resultados, o FOFE mostrou-se eficaz em reduzir significativamente o número de simulações necessárias para alcançar uma estratégia de produção otimizada. Na aplicação do FOFE-1000, que utilizou 1000 amostras por iteração, a redução foi de 1.003 simulações quando comparado ao processo tradicional (que utilizou apenas RSM); enquanto no processo com FOFE-100, que utilizou 100 amostras por iteração, a redução foi de 1.044 simulações, representando uma redução de 54% e 56%, respectivamente. Esses resultados destacam o potencial do uso integrado do FOFE com o simulador para acelerar o processo de otimização, substituindo parte do processo tradicional e melhorando a eficiência computacional em estudos de produção de petróleo.

Após a fase final da otimização, observou-se que os resultados obtidos com o uso combinado de RSM e FOFE ficaram muito próximos dos valores alcançados com a otimização, utilizando apenas os RSM. Conforme ilustrado na Tabela 14, na abordagem que gerou 1.000 amostras por iteração, a performance da FOFE ficou apenas 0,7% abaixo do valor obtido com

os RSM. Por outro lado, na abordagem que gerou 100 amostras por iteração, o FOFE conseguiu superar os resultados dos RSM por aproximadamente 0,9%. Além disso, essa variação de desempenho está dentro da margem de erro esperada para processos de otimização.

Tabela 14: Análise da qualidade do resultado comparando a otimização utilizando apenas RSM com a otimização utilizando RSM e FOFE gerando 1000 amostras (destacado na cor verde) e gerando 100 amostras (destacado na cor azul) nas três etapas do processo hierárquico.

Fase da Otimização Hierárquica	Indicador econômico RSM (bilhões de US\$)	Indicador econômico RSM e FOFE - 100 amostras (bilhões de US\$)	Qualidade do resultado	Indicador econômico RSM e FOFE - 1000 amostras (bilhões de US\$)	Qualidade do resultado
Número e posição (IEC)	4,55	4,24	-6,8%	4,55	0,0%
Ajuste fino (IEC)	4,64	4,65	0,1%	4,60	-1,0%
Cronograma de abertura (VPL)	4,77	4,81	0,9%	4,73	-0,7%

É relevante notar que o número de amostras geradas pelo FOFE, seja com 100 ou 1.000 amostras por iteração, não apresentou uma influência significativa nos resultados finais da otimização. Ambas as configurações demonstraram eficácia na redução do número total de simulações necessárias para alcançar uma estratégia ótima.

Essa proximidade nos resultados evidencia a robustez e a precisão do FOFE, consolidando como uma ferramenta eficaz na otimização de estratégias de produção, independente das diferentes configurações de amostras. A metodologia utilizando o FOFE mostrou-se capaz de acelerar significativamente o processo de otimização, reduzindo o número total de simulações necessárias e preservando a qualidade das soluções obtidas. O tempo economizado pode ser utilizado em outros estudos de simulação.

7 CONCLUSÕES

Este trabalho descreve uma metodologia utilizada para acelerar o processo na otimização de estratégias de produção de campos de petróleo em reservatórios carbonáticos com heterogeneidades complexas pois a representação adequada dessas complexidades nos modelos de simulação resulta em um custo computacional elevado. Para isso, foi adotada a aplicação de técnicas de aprendizado de máquina para substituir partes computacionalmente custosas do processo.

O trabalho foi desenvolvido em duas partes. Na Parte 1, utilizou-se os modelos de simulação tradicionais (RSM) para otimizar uma estratégia de produção, aplicando as melhores práticas para obter a configuração operacional que retornasse a melhor função objetivo avaliada. Na Parte 2, o processo foi desenvolvido seguindo a mesma abordagem do problema definido na Parte 1, aplicando uma rede neural para estimar a função objetivo avaliada de estratégias de produção, reduzindo a necessidade de realizar muitas simulações para obter uma estratégia otimizada.

O processo de otimização utilizando RSM em conjunto com o algoritmo IDLHC provou ser uma técnica muito eficaz na otimização da estratégia de produção do Campo C. Na fase inicial do processo de otimização hierárquica, o algoritmo identificou que um número reduzido de poços era ideal para a estratégia de produção, considerando a capacidade de produção limitada da plataforma. A fase de ajuste fino na localização e completação dos poços, juntamente com a otimização do cronograma de abertura dos poços, gerenciou a produção de forma eficaz, promovendo uma melhoria significativa no valor da função objetivo em comparação com a estratégia base (EPi).

A substituição das fases mais custosas computacionalmente ao longo da otimização pelo FOFE mostrou-se eficaz na abordagem do problema proposto, principalmente na fase inicial de otimização, onde o espaço de busca é mais amplo. No entanto, à medida que a otimização se tornou mais refinada e a necessidade de precisão aumentou, a representatividade do FOFE em relação ao simulador se mostrou dentro da margem de erro esperada para processos de otimização. Mesmo assim, a diferença entre os valores da função objetivo obtidos pelo RSM e pelo FOFE ao final do processo foi mínima. Além disso, a redução no número de simulações necessárias para alcançar uma estratégia otimizada foi extremamente significativa com o uso de RSM e FOFE.

O uso do treinamento dinâmico, através da aplicação com RMS e FOFE, mostrou ser importante para a obtenção de resultados consistentes. Observou-se uma melhoria contínua nas previsões à medida que novos dados eram incorporados ao treinamento a cada iteração.

No caso específico deste estudo, a variação do número de amostras geradas na otimização com RMS e FOFE (100 ou 1000) teve pouca influência na otimização. Os resultados alcançados em ambos os cenários foram muito próximos, indicando que o método é robusto em diferentes configurações de amostras.

A principal contribuição deste trabalho é a significativa redução na dependência do simulador numérico em processos de seleção de estratégia de produção, mantendo a qualidade dos resultados. Mesmo com uma quantidade reduzida de simulações, os resultados da otimização utilizando RSM foram alcançados, demonstrando a eficácia do FOFE em proporcionar uma otimização rápida e eficiente sem comprometer a precisão. Isso evidencia a capacidade do método em acelerar o processo de otimização, reduzindo o tempo computacional e a necessidade de simulações numéricas. Essa metodologia pode ter um impacto significativo na indústria de óleo e gás, onde o tempo e a precisão dos resultados são essenciais. A redução do tempo computacional acelera as tomadas de decisões importantes no projeto, liberando recursos para outras aplicações e estudos.

Recomendações para trabalhos futuros

O trabalho desenvolvido apresenta uma metodologia eficiente para otimização de estratégias de produção em reservatórios carbonáticos utilizando técnicas de inteligência artificial. Entretanto, há oportunidades para aprimorar e expandir a aplicação da metodologia desenvolvida. A seguir, são apresentadas algumas recomendações para trabalhos futuros, com objetivo de aumentar a eficiência, a robustez e a aplicabilidade da metodologia:

- Selecionar dez amostras para validação no simulador numérico, em vez de 20, em cada etapa do processo iterativo. Isso pode reduzir ainda mais o número de simulações necessárias e tornar o processo mais eficiente;
- Durante o processo iterativo de estimativa com o FOFE, realizar validações no simulador numérico em intervalos mais espaçados (por exemplo, a cada três ou cinco iterações, em vez de a cada iteração). Isso pode reduzir significativamente o número de simulações necessárias, mantendo a eficiência computacional e a precisão dos resultados;
- Treinar a rede neural com um conjunto de 50 amostras, em vez de 100, e avaliar se o desempenho (relacionado a economia de tempo e recurso computacional) é mantido;

- Testar a metodologia com abordagens que incluam outras variáveis do G1, ou que combine G1 e G2 simultaneamente, ou ainda que usem parametrizações diferentes;
- Desenvolver abordagens que considerem incertezas geológicas, além das operacionais, de forma mais abrangente. Isso permitirá uma avaliação mais completa e robusta das estratégias de produção, aumentando a confiabilidade das decisões tomadas;
- Aplicar a metodologia desenvolvida em diferentes tipos de reservatórios, com características geológicas distintas, com objetivo de validar a robustez do método e a flexibilidade em diferentes tipos de reservatórios;
- Repetir os experimentos para minimizar o efeito de aleatoriedade do método IDLHC – populacional.

REFERÊNCIAS

AGADA, S. *et al.* Deciphering the Fundamental Controls of Flow in Carbonates Using Numerical Well-Testing, Production Optimisation, and 3D High-resolution Outcrop Analogues for Fractured Carbonate Reservoirs. In: **EAGE Annual Conference & Exhibition incorporating SPE Europec**, Londres, 2013. DOI 10.2118/164805-MS. Disponível em: <https://doi.org/10.2118/164805-MS>. Acesso em: 19 set. 2022.

AGAPITOS, A.; BRABAZON, A.; O'NEILL, M. Regularised gradient boosting for financial time-series modelling. **Computational Management Science**, v. 14, n. 3, p. 367-391, jul. 2017. ISSN 1619-6988. DOI 10.1007/s10287-017-0280-y. Disponível em: <https://doi.org/10.1007/s10287-017-0280-y>. Acesso em: 19 set. 2023.

AJISAFE, F. O. *et al.* Impact of Well Spacing and Interference on Production Performance in Unconventional Reservoirs, Permian Basin. In: **SPE/AAPG/SEG Unconventional Resources Technology Conference**, Austin, Texas, EUA, julho de 2017. DOI 10.15530/URTEC-2017-2690466. Disponível em: <https://doi.org/10.15530/URTEC-2017-2690466>. Acesso em: 20 fev. 2024.

AL JAWHARI, M. O. *et al.* Integration of a Production Optimization System with Intelligent Well Surveillance for an Effective Reservoir Management in Abu Dhabi Field. In: **ADIPEC**, Abu Dhabi, Emirados Árabes Unidos, outubro de 2023. DOI 10.2118/216895-MS. Disponível em: <https://doi.org/10.2118/216895-MS>. Acesso em: 3 dez. 2023.

ANDREI, N. Scaled conjugate gradient algorithms for unconstrained optimization. **Computational Optimization and Applications**, [S.l.], v. 38, n. 3, p. 401-416, dez. 2007. DOI 10.1007/s10589-007-9055-7. Disponível em: <https://doi.org/10.1007/s10589-007-9055-7>. Acesso em: 5 nov. 2022.

ASADIAN-PAKFAR, M.; SEDAE, B.; NAKHAEI, A. A time-bounded approach for reducing runtime of well placement optimization using PSO and PSOGA algorithms. **Geoenery Science and Engineering**, [S.l.], v. 231, p. 212391, dez. 2023. ISSN 2949-8910. DOI 10.1016/j.geoen.2023.212391. Disponível em: <https://doi.org/10.1016/j.geoen.2023.212391>. Acesso em: 22 fev. 2024.

ASHENA, R.; THONHAUSER, G. Application of Artificial Neural Networks in Geoscience and Petroleum Industry. In: **Artificial Intelligent Approaches in Petroleum Geosciences**.

Cham: Springer International Publishing, 2024. p. 115-154. DOI 10.1007/978-3-031-52715-9_4. Disponível em: https://doi.org/10.1007/978-3-031-52715-9_4. Acesso em: 26 de ago. 2024.

AVANSI, G. D.; SCHIOZER, D. J. UNISIM-I: Synthetic Model for Reservoir Development and Management Applications. **International Journal of Modeling and Simulation for the Petroleum Industry**, v. 9, n. 1, p. 21–30, 2015. Disponível em: https://doi.org/10.1007/978-3-031-52715-9_4. Acesso em: 26 de ago. 2024.

AVANSI, G. D.; HOHENDORFF FILHO, J. C. V.; SCHIOZER, D. J. Simulation Models and Fast Objective Function Estimators Classification for Petroleum Reservoir Studies. **UNISIM ON-LINE**, v. 15, n. 3, 138th ed., jul. 2020. Disponível em: <https://www.unisim.cepetro.unicamp.br/index.php/publicacoes/unisim-on-line>. Acesso em: 15 set. 2022.

AZAMIPOUR, V. *et al.* An Injection Rate Optimization in a Water Flooding Case Study with an Adaptive Simulated Annealing Techniques. **Journal of Petroleum Science and Technology**, v. 7, n. 4, p. 13-31, 2017. DOI 10.22078/jpst.2017.1974.1350. Disponível em: https://jpst.ripi.ir/article_813.html. Acesso em: 24 de fev. 2024.

BABU, D. K.; ODEH, A. S. Productivity of a Horizontal Well. **SPE Reservoir Engineering**, v. 4, n. 4, p. 417-421, nov. 1989. ISSN 0885-9248. DOI 10.2118/18298-PA. Disponível em: <https://doi.org/10.2118/18298-PA>. Acesso em: 11 jun. 2023.

BARKER, J. W.; CUYPERS, M.; HOLDEN, L. Quantifying Uncertainty in Production Forecasts: Another Look at the PUNQ-S3 Problem. **SPE Journal**, v. 6, n. 4, p. 433-441, dez. 2001. DOI 10.2118/74707-PA. Disponível em: <https://doi.org/10.2118/74707-PA>. Acesso em: 29 nov. 2023.

BECKNER, B. L.; SONG, X. Field Development Planning Using Simulated Annealing - Optimal Economic Well Scheduling and Placement. In: SPE Annual Technical Conference and Exhibition, Dallas, Texas, outubro de 1995. DOI 10.2118/30650-MS. Disponível em: <https://doi.org/10.2118/30650-MS>. Acesso em: 08 mar. 2024.

BERTINI JUNIOR, J. R. *et al.* A comparison of machine learning surrogate models for net present value prediction from well placement binary data. **Journal of Petroleum Science and Engineering**, [S.l.], v. 208, p. 109208, jan. 2022. ISSN 0920-4105. DOI

<https://doi.org/10.1016/j.petrol.2021.109208>. Disponível em:
<https://www.sciencedirect.com/science/article/pii/S0920410521008615>. Acesso em: 17 jan. 2024.

BERTINI JUNIOR, J. R. *et al.* A comparison of machine learning algorithms as surrogate model for net present value prediction from wells arrangement data. In: **2019 International Joint Conference on Neural Networks (IJCNN)**, Budapeste, Hungria, 2019. p. 1-8. DOI 10.1109/IJCNN.2019.8851708. Disponível em: <https://doi.org/10.1109/IJCNN.2019.8851708>. Acesso em: 07 out. 2023.

BOTECHIA, V. E. **Análise de Desempenho de Poços na Seleção da Estratégia de Produção de Petróleo sob Incertezas**. 2012. 147 p. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) - Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2012. Disponível em: <https://www.unisim.cepetro.unicamp.br/publicacoes/Vinicius%20Eduardo%20Botechia.pdf>. Acesso em: 13 ago. 2022.

BREIMAN, L. **Population theory for boosting ensembles**. The Annals of Statistics, v. 32, n. 1, p. 1-11, fev. 2004. DOI 10.1214/aos/1079120126. Disponível em: <https://doi.org/10.1214/aos/1079120126>. Acesso em: 14 jan. 2024.

BREIMAN, L. **Random Forests**. Machine Learning, v. 45, n. 1, p. 5-32, out. 2001. ISSN 1573-0565. DOI 10.1023/A:1010933404324. Disponível em: <https://doi.org/10.1023/A:1010933404324>. Acesso em: 21 nov. 2023.

BRUYELLE, J.; GUÉRILLOT, D. Well Placement Optimization with an Artificial Intelligence Method Applied to Brugge Field. In: **SPE Gas & Oil Technology Showcase and Conference**, Dubai, Emirados Árabes Unidos, outubro de 2019. DOI 10.2118/198656-MS. Disponível em: <https://doi.org/10.2118/198656-MS>. Acesso em: 16 set. 2024.

CAVALCANTE FILHO, J. S. de A. **Metodologia de Geração de Mapas de Qualidade com Aplicação na Seleção e Otimização de Estratégias de Produção**. 2005. 106 p. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) - Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2005. Disponível em: <https://repositorio.unicamp.br/Busca/Download?codigoArquivo=473672>. Acesso em: 15 jan. 2023.

CIRAC, G. *et al.* Cross-Domain Feature learning and data augmentation for few-shot proxy development in oil industry. **Applied Soft Computing**, v. 149, p. 110972, dez. 2023. ISSN 1568-4946. DOI 10.1016/j.asoc.2023.110972. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494623009900>. Acesso em: 25 de fev. 2024.

CIRAC, G. *et al.* Deep hierarchical distillation proxy-oil modeling for heterogeneous carbonate reservoirs. **Engineering Applications of Artificial Intelligence**, v. 126, p. 107076, nov. 2023. ISSN 0952-1976. DOI 10.1016/j.engappai.2023.107076. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0952197623012605>. Acesso em: 14 de set. 2023.

COMPUTER MODELLING GROUP LTD. **CMOST-AI User Guide**. 10 ed. Calgary, Canadá: Computer Modelling Group Ltd, 2021.

CRUZ, P. S. **Reservoir management decision-making in the presence of geological uncertainty**. 2000. 217 p. Tese (Doutorado em Ciências Aplicadas) – Stanford University, Califórnia, Estados Unidos, 2000. Disponível em: <https://www.proquest.com/dissertations-theses/reservoir-management-decision-making-presence/docview/304627834/se-2?accountid=8113>. Acesso em: 20 set. 2022.

DANES, L. *et al.* Presenting a Concept for Accelerating Robust Oil Production Strategy Optimizations: Evolutionary Representativeness. **Petroleum Exploration and Development**, 2024. No prelo.

DENG, H. *et al.* Ensemble learning for the early prediction of neonatal jaundice with genetic features. **BMC Medical Informatics and Decision Making**, v. 21, n. 1, p. 338, dez. 2021. ISSN 1472-6947. DOI 10.1186/s12911-021-01701-9. Disponível em: <https://doi.org/10.1186/s12911-021-01701-9>. Acesso em: 05 de nov. 2023.

EL-MIHOUB, T, A. *et al.* Hybrid Genetic Algorithms: A Review. **Engineering Letters**, v. 13, n. 2, p. 124-137, 2006. Disponível em: https://www.engineeringletters.com/issues_v13/issue_2/EL_13_2_11.pdf. Acesso em: 28 de out. 2023.

EMERICK, A. A. *et al.* Well Placement Optimization Using a Genetic Algorithm With Nonlinear Constraints. In: **SPE Reservoir Simulation Symposium**, The Woodlands, Texas,

fevereiro de 2009. DOI 10.2118/118808-MS. Disponível em: <https://doi.org/10.2118/118808-MS>. Acesso em: 20 set. 2024.

ERTEKIN, T.; ABOU-KASSEM, J. H.; KING, G. R. **Basic Applied Reservoir Simulation**. Society of Petroleum Engineers, [S.l.], 2001. ISBN 978-1-55563-089-8. DOI 10.2118/9781555630898. Disponível em: <https://doi.org/10.2118/9781555630898>. Acesso em: 12 fev. 2023.

ERTEKIN, T.; SUN, Q. Artificial Intelligence Applications in Reservoir Engineering: A Status Check. **Energies**, Basel, v. 12, n. 15, p. 2897, 2019. ISSN 1996-1073. DOI 10.3390/en12152897. Disponível em: <https://www.proquest.com/scholarly-journals/artificial-intelligence-applications-reservoir/docview/2316982751/se-2?accountid=8113>. Acesso em: 12 dez. 2023.

EVENSEN, O.; WOMACK, D.; LIN, S. **How AI can pump new life into oilfields**. Armonk, NY: IBM Institute for Business Value, 7 fev. 2020. Disponível em: <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/oil-gas-production-optimization>. Acesso em: 28 abr. 2023.

FANCHI, J. R. **Principles of Applied Reservoir Simulation**. 4. ed. Gulf Professional Publishing, 2018. 364 p.

FARAMARZI, P.; AMINI, K. A scaled three-term conjugate gradient method for large-scale unconstrained optimization problem. **Calcolo**, v. 56, n. 4, p. 35, set. 2019. ISSN 1126-5434. DOI 10.1007/s10092-019-0333-4. Disponível em: <https://doi.org/10.1007/s10092-019-0333-4>. Acesso em: 17 set. 2024.

FEDER, J. Considering Time and Space in Drilling and Completion Can Reduce Well Interference. **Journal of Petroleum Technology**, v. 71, n. 9, p. 78-80, set. 2019. ISSN 0149-2136. DOI 10.2118/0919-0078-JPT. Disponível em: <https://doi.org/10.2118/0919-0078-JPT>. Acesso em: 20 set. 2022.

FERREIRA, C. *et al.* "Use of Emulator Methodology for Uncertainty Reduction Quantification." Paper presented at the **SPE Latin America and Caribbean Petroleum Engineering Conference**, Maracaibo, Venezuela, May 2014. DOI 10.2118/169405-MS. Disponível em: <https://doi.org/10.2118/169405-MS>. Acesso em: 25 set. 2024

FLETCHER, R.; REEVES, C. M. Function minimization by conjugate gradients. **The Computer Journal**, v. 7, n. 2, p. 149-154, 1964. ISSN 0010-4620. DOI 10.1093/comjnl/7.2.149. Disponível em: <https://doi.org/10.1093/comjnl/7.2.149>. Acesso em: 11 dez. 2023.

FRIEDMAN, J. H. Greedy Function Approximation: A Gradient Boosting Machine. **The Annals of Statistics**, v. 29, n. 5, p. 1189-1232, out. 2001. ISSN 0090-5364, 2168-8966. Disponível em: <http://www.jstor.org/stable/2699986>. Acesso em: 19 abr. 2024.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). **The Annals of Statistics**, v. 28, n. 2, p. 337-407, abr. 2000. DOI 10.1214/aos/1016218223. Disponível em: <https://doi.org/10.1214/aos/1016218223>. Acesso em: 20 abr. 2024.

GUO, D. *et al.* Numerical Analysis of Transient Pressure Behaviors with Shale Gas MFHWs Interference. **Energies**, v. 12, n. 2, p. 262, 2019. DOI 10.3390/en12020262. Disponível em: <https://doi.org/10.3390/en12020262>. Acesso em: Acesso em: 01 set. 2024.

GASPAR, A. T. F. S.; BARRETO, C. E. A. G.; SCHIOZER, D. J. **Journal of Petroleum Science and Engineering**, v. 146, p. 473-488, out. 2016. ISSN 0920-4105. DOI <https://doi.org/10.1016/j.petrol.2016.05.042>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410516302169>. Acesso em: 12 jul. 2022.

GASPAR, A. T. *et al.* Application of Assisted Optimization to Aid Oil Exploitation Strategy Selection for Offshore Fields. In: **SPE Latin America and Caribbean Petroleum Engineering Conference**, Maracaibo, Venezuela, maio de 2014. DOI 10.2118/169464-MS. Disponível em: <https://doi.org/10.2118/169464-MS>. Acesso em: 02 jul. 2022.

GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow: Conceitos, Ferramentas e Técnicas Para a Construção de Sistemas Inteligentes**. Alta Books, 2019. 576 p.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge: MIT Press, 2016. 785 p.

GUGL, R. *et al.* Evaluation of Gas-Based EOR Methods in Gas-Invaded Zones of Fractured Carbonate Reservoir, **Energies**, 15, no. 13: 4921, 2022. DOI 10.3390/en15134921 Disponível em: <https://doi.org/10.3390/en15134921>. Acesso em: 02 set. 2023.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Nova York, NY: Springer New York, 2009. p. 587-604. ISBN 978-0-387-84858-7. Disponível em: https://doi.org/10.1007/978-0-387-84858-7_15. Acesso em: 18 ago. 2024.

HAYKIN, S. **Neural Networks and Learning Machines**. 3. ed. Upper Saddle River, New Jersey: Pearson Education, Inc., 2009. 906 p.

HEMMATI-SARAPARDEH, A. *et al.* **Applications of Artificial Intelligence Techniques in the Petroleum Industry**. Gulf Professional Publishing, 2020. 312 p.

HOHENDORFF FILHO, J. C.; MASCHIO, C.; SCHIOZER, D. J. Production strategy optimization based on iterative discrete Latin hypercube. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, v. 38, n. 8, p. 2473-2480, dez. 2016. ISSN 1806-3691. DOI 10.1007/s40430-016-0511-0. Disponível em: <https://doi.org/10.1007/s40430-016-0511-0>. Acesso em: 19 jul. 2022.

HU, D. Refracturing Candidate Selection in Tight Oil Reservoirs Using Hybrid Analysis of Data and Physics Based Models." Paper presented at the **SPE Canadian Energy Technology Conference**, Calgary, Alberta, Canada, March 2022. DOI 10.2118/208883-MS. Disponível em: <https://doi.org/10.2118/208883-MS>. Acesso em: 25 set. 2024.

IMEX MANUAL. Computer Modelling Group Ltd. (CMG). . [S.l: s.n.]. , 2019.

ILHAN, E.; TURALI, M. Y.; KOZAT, S. S. Gradient Boosting With Moving-Average Terms for Nonlinear Sequential Regression. **IEEE Signal Processing Letters**, v. 30, p. 1182-1186, 2023. ISSN 1558-2361. DOI 10.1109/LSP.2023.3309577. Disponível em: <https://doi.org/10.1109/LSP.2023.3309577>. Acesso em: 09 dez. 2023.

JAHN, F.; COOK, M.; GRAHAM, M. **Hydrocarbon Exploration and Production**. 2. ed. Amsterdam: Elsevier Science, 2008. 456 p. ISBN 978-0-444-53236-7.

KANEKO, T. *et al.* Hybrid model of a physics-based model and machine learning for real-time estimation of unmeasurable parts: Mapping from measurable to unmeasurable variables. **Ocean Engineering**, v. 261, p. 112-123, 2022. Disponível em: <https://doi.org/10.1016/j.oceaneng.2022.112123>. Acesso em: 25 set. 2024.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: **Proceedings of ICNN'95 - International Conference on Neural Networks**, Perth, Australia, 27 Nov.-1 Dec. 1995. v. 4, p. 1942-1948. DOI 10.1109/ICNN.1995.488968. Disponível em: <https://doi.org/10.1109/ICNN.1995.488968>. Acesso em: 09 dez. 2023.

KHAN, M. Y. *et al.* Automated Prediction of Good Dictionary Examples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques. **Complexity**, v. 2021, p. 2553199, jan. 2021. ISSN 1076-2787. DOI 10.1155/2021/2553199. Disponível em: <https://doi.org/10.1155/2021/2553199>. Acesso em: 19 set. 2024.

KHARRAT, R.; OTT, H. A Comprehensive Review of Fracture Characterization and Its Impact on Oil Production in Naturally Fractured Reservoirs. **Energies**, v. 16, n. 8, art. 3437, 2023. ISSN 1996-1073. DOI 10.3390/en16083437. Disponível em: <https://doi.org/10.3390/en16083437>. Acesso em: 09 fev. 2024.

KIM, Y. D.; DURLOFSKY, L. J. Neural network surrogate for flow prediction and robust optimization in fractured reservoir systems. **Fuel**, v. 351, p. 128756, nov. 2023. ISSN 0016-2361. DOI 10.1016/j.fuel.2023.128756. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0016236123013698>. Acesso em: 14 set. 2024.

KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by Simulated Annealing. **Science**, v. 220, n. 4598, p. 671-680, 13 mai. 1983. DOI 10.1126/science.220.4598.671. Disponível em: <https://doi.org/10.1126/science.220.4598.671>. Acesso em: 19 set. 2024.

KUANG, L. *et al.* Application and development trend of artificial intelligence in petroleum exploration and development. **Petroleum Exploration and Development**, v. 48, n. 1, p. 1-14, fev. 2021. ISSN 1876-3804. DOI 10.1016/S1876-3804(21)60001-0. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1876380421600010>. Acesso em: 15 ago. 2023.

KWON, S. *et al.* Determination of oil well placement using convolutional neural network coupled with robust optimization under geological uncertainty. **Journal of Petroleum**

Science and Engineering, v. 201, p. 108118, jun. 2021. ISSN 0920-4105. DOI 10.1016/j.petrol.2020.108118. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410520311724>. Acesso em: 20 ago. 2023.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, n. 521, 436–444, 2015. DOI 10.1038/nature14539. Disponível em: <https://doi.org/10.1038/nature14539>. Acesso em: 26 set. 2024. LI, H. *et al.* Machine Learning-Assisted Prediction of Oil Production and CO₂ Storage Effect in CO₂-Water-Alternating-Gas Injection (CO₂-WAG). **Applied Sciences**, v. 12, n. 21, 2022. ISSN 2076-3417. DOI 10.3390/app122110958. Disponível em: <https://doi.org/10.3390/app122110958>. Acesso em: 06 fev. 2023.

LI, W. *et al.* Reservoir production prediction with optimized artificial neural network and time series approaches. **Journal of Petroleum Science and Engineering**, v. 215, p. 110586, Aug. 2022. ISSN 0920-4105. DOI 10.1016/j.petrol.2022.110586. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410522004624>. Acesso em: 19 fev. 2023.

LI, X.; ZHAO, D.; LI, Z. An overview of the reservoir simulation. In: **7th International Conference on Modelling, Identification and Control (ICMIC)**, 2015, [S.l.]. p. 1-4. DOI 10.1109/ICMIC.2015.7409348. Disponível em: <https://doi.org/10.1109/ICMIC.2015.7409348>. Acesso em: 13 fev. 2023.

LOOMBA, A. K.; BOTECHIA, V. E.; SCHIOZER, D. J. A comparative study to accelerate field development plan optimization. **Journal of Petroleum Science and Engineering**, v. 208, p. 109708, jan. 2022. ISSN 0920-4105. DOI 10.1016/j.petrol.2021.109708. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410521013358>. Acesso em: 29 nov. 2023.

LUCIA, F. J. **Limestone Reservoirs**. In: Carbonate Reservoir Characterization: An Integrated Approach. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 181-215. ISBN 978-3-540-72742-2. DOI 10.1007/978-3-540-72742-2_6. Disponível em: https://doi.org/10.1007/978-3-540-72742-2_6. Acesso em: 19 set. 2022.

MARDLE, S.; PASCOE, S. An overview of genetic algorithms for the solution of optimisation problems. **Computers in Higher Education Economics Review**, v. 13, n. 1, p. 16-20, 1999. Disponível em: <https://ideas.repec.org/a/che/chepap/v13y1999i1p16-20.html>. Acesso em: 06 set. 2023.

MARTINI, R F; SCHIOZER, D J; NAKAJIMA, L. Use of quality maps in reservoir management. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, v. 27, p. 463–468, 2005. DOI 10.1590/S1678-5878200500. Disponível em: <https://doi.org/10.1590/S1678-58782005000400016>. Acesso em: 06 set. 2022.

MASCHIO, C.; AVANSI, G. D.; SCHIOZER, D. J. Data Assimilation Using Principal Component Analysis and Artificial Neural Network. **SPE Reservoir Evaluation & Engineering**, v. 26, n. 3, p. 795-812, ago. 2023. ISSN 1094-6470. DOI 10.2118/214688-PA. Disponível em: <https://doi.org/10.2118/214688-PA>. Acesso em: 21 jan. 2024.

MASCHIO, C.; NAKAJIMA, L.; SCHIOZER, D. J. Production Strategy Optimization Using Genetic Algorithm and Quality Map. In: **Europec/EAGE Conference and Exhibition**, Roma, Itália, junho de 2008. DOI 10.2118/113483-MS. Disponível em: <https://doi.org/10.2118/113483-MS>. Acesso em: 19 set. 2022.

MASCHIO, C.; SCHIOZER, D. J. A new methodology for history matching combining iterative discrete Latin Hypercube with multi-start simulated annealing. **Journal of Petroleum Science and Engineering**, v. 169, p. 560-577, out. 2018. ISSN 0920-4105. DOI <https://doi.org/10.1016/j.petrol.2018.06.004>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410518304881>. Acesso em: 21 jun. 2022.

MASCHIO, C.; SCHIOZER, D. J. Probabilistic history matching using discrete Latin Hypercube sampling and nonparametric density estimation. **Journal of Petroleum Science and Engineering**, v. 147, p. 98-115, nov. 2016. ISSN 0920-4105. DOI <https://doi.org/10.1016/j.petrol.2016.05.011>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410516301759>. Acesso em: 21 jun. 2022.

MATTAX, C. C.; DALTON, R. L. Reservoir Simulation (includes associated papers 21606 and 21620). **Journal of Petroleum Technology**, v. 42, n. 6, p. 692-695, jun. 1990. ISSN 0149-2136. DOI 10.2118/20399-PA. Disponível em: <https://doi.org/10.2118/20399-PA>. Acesso em: 21 jul. 2022.

MAYERHOFER, M. A Cost/Benefit Review of Completion Choices in the Williston Basin Using a Hybrid Physics-Based-Modeling/Multivariate-Analysis Approach." **SPE Prod & Oper** 34 (2019): 24–40. DOI 10.2118/187254-PA. Disponível em: [doi: https://doi.org/10.2118/187254-PA](https://doi.org/10.2118/187254-PA). Acesso em: 25 set. 2024.

MAZZULLO, S. J.; RIEKE, H. H.; CHILINGARIAN, G. V. **Carbonate Reservoir Characterization: A Geologic-Engineering Analysis, Part II**. 1. ed. Amsterdam: Elsevier, 1996. 993 p.

MCKAY, M. D.; BECKMAN, R. J.; CONOVER, W. J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. **Technometrics**, v. 42, n. 1, p. 55-61, fev. 2000. ISSN 0040-1706. DOI 10.1080/00401706.2000.10485979. Disponível em: <https://doi.org/10.1080/00401706.2000.10485979>. Acesso em: 21 set. 2024.

MEIRA, L. A. A.; COELHO, G. P.; SILVA, C. G.; ABREU, J. L. A.; SANTOS, A. A. S.; SCHIOZER, D. J. Improving representativeness in a scenario reduction process to aid decision making in petroleum fields. **Journal of Petroleum Science and Engineering**, v. 184, p. 106398, jan. 2020. ISSN 0920-4105. DOI <https://doi.org/10.1016/j.petrol.2019.106398>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410519308198>. Acesso em: 21 set. 2022.

MELLO, S.F. *et al.* A Practical Reservoir Engineering Approach to Build Faster Complex Numerical Models: A Real Field Application. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, 2024. No prelo.

MICHAEL, A. A Hybrid Data-Driven/Physics-Based Modeling Approach for Hydraulic Fracture Initiation and Early-Phase Propagation in Shale. **Geomechanics for Energy and the Environment**, v. 34, p. 100453, 2023. Disponível em: <https://doi.org/10.1016/j.gete.2023.100453>. Acesso em: 25 set. 2024.

MIRZAEI-PAIAMAN, A.; SANTOS, S. M. G.; SCHIOZER, D. J. A review on closed-loop field development and management. **Journal of Petroleum Science and Engineering**, v. 201, p. 108457, jun. 2021. ISSN 0920-4105. DOI <https://doi.org/10.1016/j.petrol.2021.108457>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410521001169>. Acesso em: 03 jul. 2022.

MOGOLLON, J. L. *et al.* Comparative Analysis of Data-Driven, Physics-Based and Hybrid Reservoir Modeling Approaches in Waterflooding. Paper presented at the **SPE Annual Technical Conference and Exhibition**, Houston, Texas, USA, October 2022.

DOI 10.2118/210373-MS. Disponível em: <https://doi.org/10.2118/210373-MS>. Acesso em: 25 set. 2024.

MOHAGHEGH, S. D. Determining the main drivers in hydrocarbon production from shale using advanced data-driven analytics – A case study in Marcellus shale. **Journal of Unconventional Oil and Gas Resources**, v. 15, p. 146-157, set. 2016. ISSN 2213-3976. DOI <https://doi.org/10.1016/j.juogr.2016.07.004>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S221339761630026X>. Acesso em: 25 set. 2022.

MØLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, v. 6, n. 4, p. 525-533, jan. 1993. ISSN 0893-6080. DOI [https://doi.org/10.1016/S0893-6080\(05\)80056-5](https://doi.org/10.1016/S0893-6080(05)80056-5). Disponível em: <https://www.sciencedirect.com/science/article/pii/S0893608005800565>. Acesso em: 21 out. 2022.

MORENO, R. *et al.* Emulation of reservoir production forecast considering variation in petrophysical properties. **Journal of Petroleum Science and Engineering**, v. 165, p. 711-725, 2018. DOI [10.1016/j.petrol.2018.02.056](https://doi.org/10.1016/j.petrol.2018.02.056). Disponível em: <https://doi.org/10.1016/j.petrol.2018.02.056>. Acesso em: 25 set. 2024.

MUSAYEV, K.; SHIN, H.; NGUYEN-LE, V. Optimization of CO₂ injection and brine production well placement using a genetic algorithm and artificial neural network-based proxy model. **International Journal of Greenhouse Gas Control**, v. 127, p. 103915, jul. 2023. ISSN 1750-5836. DOI <https://doi.org/10.1016/j.ijggc.2023.103915>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1750583623000853>. Acesso em: 12 mai. 2024.

NAKAJIMA, L.; SCHIOZER, D. J. Horizontal Well Placement Optimization Using Quality Map Definition. In: **Canadian International Petroleum Conference**, Calgary, Alberta, junho de 2003. DOI [10.2118/2003-053](https://doi.org/10.2118/2003-053). Disponível em: <https://doi.org/10.2118/2003-053>. Acesso em: 29 jul. 2022.

NAKAJIMA, L. **Otimização de Desempenho de Poços Horizontais no Desenvolvimento de Campos de Petróleo**. 2003. 128 p. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) - Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2003. Disponível em:

https://www.unisim.cepetro.unicamp.br/publicacoes/tese_lincoln.pdf. Acesso em: 13 ago. 2022.

NASIR, Y.; YU, W.; SEPEHRNOORI, K. Hybrid derivative-free technique and effective machine learning surrogate for nonlinear constrained well placement and production optimization. **Journal of Petroleum Science and Engineering**, v. 186, p. 106726, mar. 2020. DOI 10.1016/j.petrol.2019.106726. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410519311465>. Acesso em: 10 set. 2023.

NELSON, R. A. **Geologic Analysis of Naturally Fractured Reservoirs**. 2. ed. Woburn: Gulf Professional Publishing, 2001. ISBN 978-0-88415-317-7. DOI 10.1016/B978-088415317-7/50000-2. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780884153177500002>. Acesso em: 11 set. 2022.

NIE, R.-S. *et al.* Dual Porosity and Dual Permeability Modeling of Horizontal Well in Naturally Fractured Reservoir. **Transport in Porous Media**, v. 92, n. 1, p. 213-235, mar. 2012. ISSN 1573-1634. DOI 10.1007/s11242-011-9898-3. Disponível em: <https://doi.org/10.1007/s11242-011-9898-3>. Acesso em: 10 set. 2022.

NOGUEIRA, P. B.; SCHIOZER, D. J. An Efficient Methodology of Production Strategy Optimization Based on Genetic Algorithms. In: **Latin American and Caribbean Petroleum Engineering Conference**, 31 maio 2009. DOI 10.2118/122031-MS. Disponível em: <https://doi.org/10.2118/122031-MS>. Acesso em: 20 set. 2022.

NWACHUKWU, A. *et al.* Fast evaluation of well placements in heterogeneous reservoir models using machine learning. **Journal of Petroleum Science and Engineering**, v. 163, p. 463-475, abr. 2018. ISSN 0920-4105. DOI 10.1016/j.petrol.2018.01.019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410518300226>. Acesso em: 20 set. 2024.

ODEH, A. S. Reservoir Simulation ...What Is It. **Journal of Petroleum Technology**, v. 21, n. 11, p. 1383–1388, nov. 1969. ISSN 0149-2136. DOI 10.2118/2790-PA. Disponível em: <https://doi.org/10.2118/2790-PA>. Acesso em: 26 jul. 2022.

PAIVA, H. P. **Simulação da Recuperação de Petróleo em Reservatórios Naturalmente Fraturados**. 2012. 148 p. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) - Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2012. Disponível em: <https://www.unisim.cepetro.unicamp.br/publicacoes/Hernani%20Petroni%20Paiva.pdf>. Acesso em: 13 ago. 2024.

PANDEY, Y. N.; RASTOGI, A.; KAINKARYAM, S.; BHATTACHARYA, S.; SAPUTELLI, L. **Machine Learning in the Oil and Gas Industry: Including Geosciences, Reservoir Engineering, and Production Engineering with Python**. 1. ed. Berkeley, CA: Apress, 2020. ISBN 978-1-4842-6093-7. Disponível em: <https://doi.org/10.1007/978-1-4842-6094-4>. Acesso em: 25 set. 2023.

PANJA, P.; VELASCO, R.; PATHAK, M.; DEO, M. Application of artificial intelligence to forecast hydrocarbon production from shales. **Petroleum**, v. 4, n. 1, p. 75-89, mar. 2018. ISSN 2405-6561. DOI 10.1016/j.petlm.2017.11.003. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2405656117301141>. Acesso em: 20 out. 2023.

PICHON, S.; CAFARDI ORIHUELA, F. G.; LAGARRIGUE, E.; CAVAZZOLI, G. When, where, and how to drill and complete pads of multiple wells? Four-dimensional considerations for field development in the Vaca Muerta Shale. In: **SPE Argentina Exploration and Production of Unconventional Resources Symposium**, 14 ago. 2018. DOI 10.2118/191836-MS. Disponível em: <https://doi.org/10.2118/191836-MS>. Acesso em: 01 set. 2024.

POLAK, E.; RIBIERE, G. Note sur la convergence de méthodes de directions conjuguées. **Revue française d'informatique et de recherche opérationnelle**. Série rouge, Paris: Dunod, v. 3, n. R1, p. 35-43, 1969. Disponível em: http://www.numdam.org/item/M2AN_1969__3_1_35_0/. Acesso em: 11 jul. 2023.

PRESHO, M.; WO, S.; GINTING, V. Calibrated dual porosity, dual permeability modeling of fractured reservoirs. **Journal of Petroleum Science and Engineering**, v. 77, n. 3, p. 326-337, 1 jun. 2011. ISSN 0920-4105. DOI: <https://doi.org/10.1016/j.petrol.2011.04.007>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410511000878>. Acesso em: 07 out. 2023.

RAHMANIFARD, H.; PLAKSINA, T. Application of artificial intelligence techniques in the petroleum industry: a review. **Artificial Intelligence Review**, v. 52, n. 4, p. 2295-2318, 1 dez. 2019. ISSN 1573-7462. DOI: 10.1007/s10462-018-9612-8. Disponível em: <https://doi.org/10.1007/s10462-018-9612-8>. Acesso em: 03 nov. 2023.

RIOS, V. S.; AVANSI, G. D.; SCHIOZER, D. J. Practical workflow to improve numerical performance in time-consuming reservoir simulation models using submodels and shorter period of time. **Journal of Petroleum Science and Engineering**, v. 195, p. 107547, dez. 2020. DOI 10.1016/j.petrol.2020.107547. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410520306185>. Acesso em: 31 out. 2022.

ROSA, A.J.; CARVALHO, R. S.; XAVIER, J. A. D. **Engenharia de Reservatórios de Petróleo**. [S.l.]: Interciência, 2006.

RUDER, S. An overview of gradient descent optimization algorithms. **ArXiv**. CoRR, v. abs/1609.04747, 2016. Disponível em: <http://arxiv.org/abs/1609.04747>. Acesso em: 20 jul. 2023.

RUEDA, C. J. A.; MEJIA S. E. C.; ROEHL, D. Integrated discrete fracture and dual porosity - Dual permeability models for fluid flow in deformable fractured media. **Journal of Petroleum Science and Engineering**, v. 175, p. 644-653, abr. 2019. DOI: 10.1016/j.petrol.2018.12.053. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410518311550>. Acesso em: 13 dez. 2023.

SANTOS, D. R. *et al.* A Machine Learning Approach to Reduce the Number of Simulations for Long-Term Well Control Optimization. In: **SPE Annual Technical Conference and Exhibition**, Virtual, outubro de 2020. DOI 10.2118/201379-MS. Disponível em: <https://doi.org/10.2118/201379-MS>. Acesso em: 3 de mar. 2023.

SANTOS, S. M. G. *et al.* Expected value, downside risk and upside potential as decision criteria in production strategy selection for petroleum field development. **Journal of Petroleum Science and Engineering**, v. 157, p. 81-93, ago. 2017. DOI: 10.1016/j.petrol.2017.07.002. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410517305636>. Acesso em: 20 out. 2022.

SANTOS, S. M. G.; GASPAR, A. T. F. S.; SCHIOZER, D. J. Managing reservoir uncertainty in petroleum field development: Defining a flexible production strategy from a set of rigid candidate strategies. **Journal of Petroleum Science and Engineering**, v. 171, p. 516-528, dez. 2018. DOI: 10.1016/j.petrol.2018.07.048. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410518306235>. Acesso em: 20 out. 2022.

SARMA, P.; DURLOFSKY, L. J.; AZIZ, K.; CHEN, W. H. Efficient real-time reservoir management using adjoint-based optimal control and model updating. **Computational Geosciences**, v. 10, n. 1, p. 3-36, mar. 2006. DOI: 10.1007/s10596-005-9009-z. Disponível em: <https://doi.org/10.1007/s10596-005-9009-z>. Acesso em: 21 out. 2022.

SCHIOZER, D. J. *et al.* Use of representative models in the integration of risk analysis and production strategy definition. **Journal of Petroleum Science and Engineering**, v. 44, n. 1, p. 131-141, 31 out. 2004. Risk Analysis Applied to Petroleum Exploration and Production. DOI: <https://doi.org/10.1016/j.petrol.2004.02.010>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410504000403>. Acesso em: 21 maio 2022.

SCHIOZER, D. J. *et al.* Model-based decision analysis applied to petroleum field development and management. **Oil & Gas Science and Technology – Rev. IFP Energies nouvelles**, v. 74, art. 46, 2019. DOI: <https://doi.org/10.2516/ogst/2019019>. Disponível em: <https://doi.org/10.2516/ogst/2019019>. Acesso em: 21 maio 2022.

SHI, Y.; LI, J.; LI, Z. Gradient Boosting With Piece-Wise Linear Regression Trees. **ArXiv**, v. abs/1802.05640, 2018. Disponível em <https://doi.org/10.48550/arXiv.1802.05640>. Acesso em: 20 set. 2024.

SILVA, L. M. *et al.* Selection of a Dimensionality Reduction Method: An Application to Deal with High-Dimensional Geostatistical Realizations in Oil Reservoirs. **SPE Reservoir Evaluation & Engineering**, v. 26, n. 02, p. 392-410, 2023. DOI 10.2118/212299-PA. Disponível em: <https://doi.org/10.2118/212299-PA>. Acesso em: 20 set. 2024.

SILVA, L. M.; AVANSI, G. D.; SCHIOZER, D. J. Development of proxy models for Petroleum Reservoir Simulation: A Systematic Literature Review and state-of-the-art. **International Journal of Advanced Engineering Research and Science**, v. 7, n. 10, p. 36–62, 2020. DOI 10.22161/ijaers.710.5. Disponível em <https://doi.org/10.22161/ijaers.710.5>. Acesso em: 20 set. 2022.

SIRCAR, A. Application of machine learning and artificial intelligence in oil and gas industry. **Petroleum Research**, v. 6, n. 4, p. 379-391, dez. 2021. ISSN 2096-2495. DOI 10.1016/j.ptlrs.2021.05.009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2096249521000429>. Acesso em: 20 nov. 2022.

SON, C. *et al.* Two-stage sampling scheme for a CNN-based well placement optimization of 3D benchmark reservoirs. **Geoenergy Science and Engineering**, v. 225, p. 211677, jun. 2023. ISSN 2949-8910. DOI 10.1016/j.geoen.2023.211677. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949891023002646>. Acesso em: 25 ago. 2023.

STAFF, G. *et al.* Physics Guided Machine Learning Significantly Improves Outcomes for Data-Based Production Optimization. In: **Abu Dhabi International Petroleum Exhibition & Conference**, 9 nov. 2020. DOI 10.2118/202657-MS. Disponível em: <https://doi.org/10.2118/202657-MS>. Acesso em: 20 nov. 2023.

STAGS, H.M.; HERBECK, E.F. Reservoir Simulation Models: An Engineering Overview. **Journal of Petroleum Technology**, v. 23, n. 12, p. 1428-1436, dez. 1971. DOI 10.2118/3304-PA. Disponível em: <https://doi.org/10.2118/3304-PA>. Acesso em: 15 out. 2022.

SUSLICK, S.B.; SCHIOZER, D.J. Risk analysis applied to petroleum exploration and production: an overview. **Journal of Petroleum Science and Engineering**, v. 44, n. 1, p. 1-9, 31 out. 2004. DOI 10.1016/j.petrol.2004.02.001. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410504000312>. Acesso em: 16 ago. 2022.

TABATABAEI, S. M.; ASADIAN-PAKFAR, M.; SEDAEI, B. Well placement optimization with a novel swarm intelligence optimization algorithm: Sparrow Search Algorithm. **Geoenergy Science and Engineering**, v. 231, p. 212291, 1 dez. 2023. DOI 10.1016/j.geoen.2023.212291. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949891023008783>. Acesso em: 12 nov. 2023.

THAI BA, N. *et al.* Applying the hydrodynamic model to optimize the production for crystalline basement reservoir, X field, Cuu Long Basin, Vietnam. **Journal of Petroleum Exploration and Production Technology**, v. 10, n. 1, p. 31-46, 1 jan. 2020. DOI 10.1007/s13202-019-00755-w. Disponível em: <https://doi.org/10.1007/s13202-019-00755-w>. Acesso em: 30 out. 2023.

THE MATHWORKS INC. *MATLAB version: 9.2.0.518641 (R2017a)*. . Natick, Massachusetts, United States: The MathWorks Inc., 2017a. Disponível em: <https://www.mathworks.com>. Acesso em: 31 jul. 2023.

THE MATHWORKS INC. **trainscg Scaled Conjugate Gradient Backpropagation**. . Natick, Massachusetts, United States: The MathWorks Inc., 2017. Disponível em: https://www.mathworks.com/help/deeplearning/ref/trainscg.html?s_tid=doc_ta. Acesso em: 31 jul. 2023.

TIAB, D.; DONALDSON, E. C. **Petrophysics: theory and practice of measuring reservoir rock and fluid transport properties**. 3. ed. Elsevier/Gulf Professional Pub., 2012. ISBN 9780123838483. Disponível em: <https://cir.nii.ac.jp/crid/1130000794209259136>. Acesso em: 20 fev. 2023.

UNISIM, Universidade Estadual de Campinas (UNICAMP). MERO: Módulo de otimização (OTM). Campinas: [s.n.], 2020.

WANG, D.; TAN, D.; LIU, L. Particle swarm optimization algorithm: an overview. **Soft Computing**, v. 22, n. 2, p. 387-408, 2018. DOI 10.1007/s00500-016-2474-6. Disponível em: <https://doi.org/10.1007/s00500-016-2474-6>. Acesso em: 19 jan. 2024.

WANG, Y.-W.; DAI, Z.-X.; WANG, G.-S.; CHEN, L.; XIA, Y.-Z.; ZHOU, Y.-H. A hybrid physics-informed data-driven neural network for CO₂ storage in depleted shale reservoirs. **Petroleum Science**, n. xxxx, 2023. Disponível em: <https://doi.org/10.1016/j.petsci.2023.08.032>. Acesso em: 25 de set. 2024.

WANG, X.; FENG, Q.; HAYNES, R. D. Optimization of Well Placement and Production for Large-scale Mature Oil Fields. **Journal of Engineering Science and Technology Review** v. 8, n. 5, 2015. Disponível em: <http://jestr.org/downloads/Volume8Issue5/fulltext85192015.pdf>. Acesso em: 25 out. 2022.

WEBBER, K. J.; VAN GEUNS, L. C. Framework for Constructing Clastic Reservoir Simulation Models. **Journal of Petroleum Technology**, v. 42, n. 10, p. 1248-1297, out. 1990. DOI 10.2118/19582-PA. Disponível em: <https://doi.org/10.2118/19582-PA>. Acesso em: 20 ago. 2022.

WILSON, A. Use of Emulator Methodology for Uncertainty-Reduction Quantification. **J Pet Technol** 68 (2016): 63–64. DOI 10.2118/0716-0063-JPT. Disponível em: <https://doi.org/10.2118/0716-0063-JPT>. Acesso em: 25 set. 2024.

XAVIER, A. M. **Análise do Valor da Informação na Avaliação e Desenvolvimento de Campos de Petróleo**. 2004. 159p. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) - Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2004. Disponível em: https://www.unisim.cepetro.unicamp.br/publicacoes/tese_xavier.pdf. Acesso em: 13 ago. 2024.

XIANG, X.; ZHANG, H.; XIA, S. Label Aggregation of Gradient Boosting Decision Trees. In: **Proceedings of the 2020 2nd International Conference on Image Processing and Machine Vision (IPMV '20)**. New York: Association for Computing Machinery, 2020. p. 140–145. DOI 10.1145/3421558.3421581. Disponível em: <https://doi.org/10.1145/3421558.3421581>. Acesso em: 20 jan. 2024.

XU, Z; CHEN, J; LEUNG, J. Y. An improved dual-porosity dual-permeability modeling workflow for representing nonplanar hydraulic fractures. **Gas Science and Engineering**, v. 118, p. 205108, out. 2023. ISSN 2949-9089. DOI: <https://doi.org/10.1016/j.jgsce.2023.205108>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949908923002364>. Acesso em: 02 mar. 2024.

YANG, D.; ZHANG, Q.; GU, Y. Integrated optimization and control of the production-injection operation systems for hydrocarbon reservoirs. **Journal of Petroleum Science and Engineering**, v. 37, n. 1, p. 69-81, fev. 2003. ISSN 0920-4105. DOI: [https://doi.org/10.1016/S0920-4105\(02\)00311-X](https://doi.org/10.1016/S0920-4105(02)00311-X). Disponível em: <https://www.sciencedirect.com/science/article/pii/S092041050200311X>. Acesso em: 19 set. 2023.

YETURU, K. **Machine learning algorithms, applications, and practices in data science**. In: RAO, Arni S.R. Srinivasa; RAO, C.R. (orgs.). *Handbook of Statistics*. v. 43. Elsevier, 2020. p. 81-206. DOI: <https://doi.org/10.1016/bs.host.2020.01.002>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169716120300225>. Acesso em: 20 jan. 2024.

Zabinsky, Z. B. Random search algorithms. **Department of Industrial and Systems Engineering, University of Washington**, USA, 2009. Disponível em: <https://courses.washington.edu/inde510/516/AdapRandomSearch4.05.2009.pdf>. Acesso em: 05 nov. 2023.

ZANDVLIET, M.J. *et al.* Bang-bang control and singular arcs in reservoir flooding. **Journal of Petroleum Science and Engineering**, v. 58, n. 1, p. 186-200, ago. 2007. ISSN 0920-4105. DOI: <https://doi.org/10.1016/j.petrol.2006.12.008>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410507000101>. Acesso em: 05 dez. 2023.

ZHENG, C. Research on Reservoir Numerical Simulation Method and Algorithm Design. In: **2023 4th International Conference for Emerging Technology (INCET)**, 26-28 maio 2023, p. 1-6. DOI 10.1109/INCET57972.2023.10170560. Disponível em: <https://doi.org/10.1109/INCET57972.2023.10170560>. Acesso em: 20 abr. 2023.

ZHOU, Y.; HOOKER, G. Boulevard: Regularized Stochastic Gradient Boosted Trees and Their Limiting Distribution. **ArXiv**, 13 set. 2019. DOI 10.48550/arXiv.1806.09762. Disponível em: <https://arxiv.org/abs/1806.09762>. Acesso em: 19 set. 2024.