

Aplicações de Programação Genética em Reservatórios de Petróleo

Jeff A. Maynard, Alejandra Camacho, Rafael L. Oliveira, Alvaro Talavera, Marco Aurélio Pacheco

Laboratório de Inteligência Computacional Aplicada
Pontifícia Universidade Católica do Rio de Janeiro
Rio de Janeiro, Brasil

jeffmg@ele.puc-rio.br, alecs14@ele.puc-rio.br, rafaello@ele.puc-rio.br, alvaro@ele.puc-rio.br, marco@ele.puc-rio.br

Resumo—O presente artigo tem como objetivo motivar a aplicação de programação genética multi-gene na indústria de petróleo, especificamente na área de caracterização e engenharia de reservatórios. O primeiro problema exposto aborda a estimação do perfil de poço sônico em poços petrolíferos, utilizando alguns dados de perfis geológicos de poço. O segundo problema trata sobre a quantificação de incerteza, num modelo de simulação de reservatórios, por meio da aproximação da função distribuição de probabilidade da produção acumulada de óleo, gás e água. Uma vantagem na utilização da programação genética é a possibilidade de obter uma equação matemática explícita para cada problema tratado, sendo simples a avaliação desta em softwares para a caracterização de reservatórios e para a utilização de proxies (ou modelos substitutos) de reservatórios.

Palavras Chave—*programação genética; reservatórios de petróleo; estimação de perfis; quantificação de incerteza; distribuição de probabilidade.*

I. INTRODUÇÃO

Técnicas de inteligência computacional têm sido utilizadas com muita frequência na indústria de petróleo. Simuladores para a caracterização de reservatórios (como o Petrel da Schlumberger e OpendTect) já possuem desde alguns anos ferramentas de redes neurais (RN) e aplicações com algoritmos genéticos. Trabalhos como [1], [2] e [3] formulam propostas de soluções em diferentes áreas da engenharia de petróleo, por meio de algoritmos genéticos, redes neurais e lógica fuzzy. Por outro lado, programação genética (PG), a qual está incluída na família de técnicas de computação evolucionária, vem se destacando cada vez mais em aplicações onde se tenha pouca informação e seja preciso uma expressão matemática explícita para descrever um problema [4], [5].

Este artigo tenta evidenciar aplicações de PG na indústria de petróleo em duas áreas de muito interesse. A primeira aplicação é a estimação de perfis de poços por meio de PG multi-gene (PGmg), para a caracterização de reservatórios. Trabalhos nesse sentido são encontrados em [6] onde é utilizado PGmg na previsão da permeabilidade. Neste artigo estamos interessados na previsão do perfil de poço sônico (DT) por meio de outros perfis de poços, onde o resultado será comparado com uma rede neural artificial. Uma das vantagens de utilizar PG em relação a outras técnicas como, redes neurais [7],[8] e lógica fuzzy [9], é que a expressão matemática explícita resultante pode ser

incorporada em qualquer software de caracterização de reservatórios (como o Petrel).

A segunda aplicação visa o tratamento de incerteza em modelos de reservatórios de petróleo, isto devido aos modelos de simulação de reservatórios estarem sujeitos à incerteza presente em uma grande variedade de seus parâmetros de entrada. Esta incerteza é o resultado da heterogeneidade das formações geológicas, erros nas medições dos dados e na modelagem petrofísica e estrutural. Uma análise precisa da incerteza requer normalmente uma quantidade elevada de simulações, o qual é usualmente inviável se adicionarmos o tempo consumido ao simular modelos de grande escala. Por outro lado, uma análise correta da propagação da incerteza através do modelo, aumenta a qualidade e robustez das decisões tomadas para o gerenciamento do campo. Por isso, é preciso dar um enfoque eficiente na quantificação de incerteza. A necessidade de determinar uma resposta robusta frente à estocasticidade presente nestes modelos fizeram com que o estudo do tratamento de incertezas ganhasse bastante destaque na literatura atual de petróleo, ver [10], [11] e [12].

Na engenharia de reservatórios é preciso quantificar a incerteza em função de todas as variáveis do reservatório [13]. Desafortunadamente, a relação entre as variáveis do modelo do reservatório (tais como a permeabilidade ou a porosidade) e os dados (tais como as medições do corte de água ou pressão) é com frequência altamente não linear, e nestes casos pode ser muito complicado caracterizar a função distribuição de probabilidade (pdf) a posteriori. O método mais conhecido para resolver este problema é o Monte Carlo (MC), porém a grande desvantagem desta técnica encontra-se em sua lenta convergência, ou seja, geralmente, este método tem um custo computacional alto. Por outro lado, PG é uma técnica que vem sendo utilizada com sucesso na estimação da saída de modelos físicos. Para o problema de aplicação considerado neste artigo, é encontrada uma expressão analítica mediante PGmg que depende de alguns operadores definidos previamente (soma, subtração, potência, raiz, etc.), para assim gerar um proxy, ou seja, um modelo que substituiu o simulador que gera as curvas de produção acumulada de óleo, água ou gás de um reservatório de petróleo. Esta função matemática resultante, é também uma variável aleatória e por tanto, pode-se quantificar sua incerteza estimando sua média, variância, percentis ou sua distribuição de probabilidade completa.

O restante deste artigo é organizado como segue. Na segunda seção é apresentada a definição e as propriedades do algoritmo de PGmg, o qual será utilizado na resolução dos problemas considerados nas seguintes seções. Na terceira seção é formulado o problema da estimação de perfis de poços a partir de outros perfis, em reservatórios de petróleo, mediante PGmg. Na quarta seção é descrito o problema da quantificação de incerteza por meio da aproximação da distribuição de probabilidade, num modelo de simulação de reservatórios de petróleo utilizando PGmg. Na quinta seção, são apresentados os resultados dos problemas formulados nas seções anteriores, além de apresentar uma breve descrição dos campos de petróleo que foram utilizados. Finalmente, na sexta seção são enunciadas as conclusões deste artigo.

II. PROGRAMAÇÃO GENÉTICA MULTI-GENE

O algoritmo de programação genética é uma coleção de técnicas de computação evolucionária, que cria programas de computador para solucionar problemas automaticamente sem que o usuário tenha a necessidade de conhecer ou especificar a forma da solução [14]. Análogo à evolução dos seres vivos, estes programas são submetidos a um processo de evolução com etapas de avaliação, seleção, cruzamentos e mutações. Os indivíduos são estruturas de dados em forma de árvore, compostas por nós e folhas. Os nós representam funções aritméticas básicas (soma, subtração) ou operadores mais complexos (sen, cos, exp) e as folhas são os terminais que podem ser variáveis ou constantes. O algoritmo de PG cria uma população inicial, conformada por estes indivíduos. A cada indivíduo é atribuído um valor numérico chamado aptidão, que indica o quanto a solução representada por este indivíduo é melhor em relação às outras soluções da população. A população de soluções candidatas é modificada iterativamente, cada iteração envolve a aplicação de operadores genéticos (seleção, crossover e mutação) na expectativa de gerar novos e melhores candidatos [15]. Programação genética multi-gene é uma variação de programação genética clássica, na qual, cada indivíduo multi-gene está conformado por um ou mais genes e cada gene é uma árvore da programação genética tradicional. O modelo é uma combinação linear ponderada de cada gene e os pesos dos genes, neste artigo, foram otimizados mediante uma regressão de mínimos quadrados ordinária. Deste modo, o resultante modelo pseudo-linear pode capturar o comportamento não linear [16]. Matematicamente o modelo de regressão multi-gene pode ser escrito como, $\hat{y} = d_0 + d_1 \times \text{árvore}_1 + \dots + d_M \times \text{árvore}_M$, onde cada árvore é uma função de zero ou mais das N variáveis de entrada, x_1, \dots, x_N , d_0 é o viés, os termos d_1, \dots, d_M são os pesos dos genes e M é o número de genes, ou seja, de árvores que compõem o indivíduo [16]. Em Fig. 1 pode-se observar um modelo multi-gene com saída y e variáveis de entrada x_1, x_2 e x_3 .

Finalmente, o tamanho da população, o número de gerações, a taxa de cruzamento, a profundidade máxima da árvore e demais parâmetros são testados para encontrar o modelo que tenha o erro menor. Como é usual nas técnicas de inteligência computacional, estas provas são feitas em conjuntos de dados de treinamento e validação fornecidos pelo usuário.

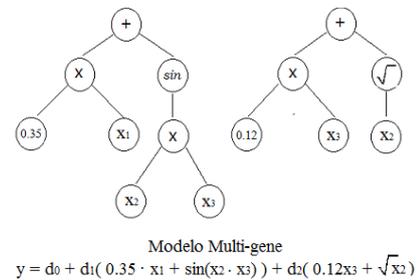


Fig. 1. Modelo PG multi-gene.

A solução obtida é avaliada posteriormente num conjunto de teste. Neste artigo foi utilizado o RMSE (a raiz da média dos erros quadráticos) como função de avaliação.

III. ESTIMAÇÃO DE PERFIS DE POÇO

A exploração de petróleo depende da correta caracterização dos materiais que compõem o reservatório. Para isso são utilizadas medições diretas, obtidas mediante um deslocamento contínuo de uma sonda dentro dos poços [13]. O gráfico destas amostras contra a profundidade é denominado perfil de poço. Existem distintos tipos de perfis de poço e estes recebem o nome pela informação que trazem ou pelo método de aquisição. Os principais são: raios gama (GR), porosidade neutrônica (NPHI), resistividade (ILD), densidade (RHOB) e sônico (DT). Um exemplo, meramente ilustrativo, do comportamento destes cinco perfis é apresentado em Fig. 2. A informação dos perfis é fundamental para o desenvolvimento de modelos geológicos. Posteriormente os dados localizados unicamente nos poços são extrapolados para o resto do reservatório utilizando métodos geoestatísticos e técnicas que utilizam dados sísmicos ou uma combinação de ambos.

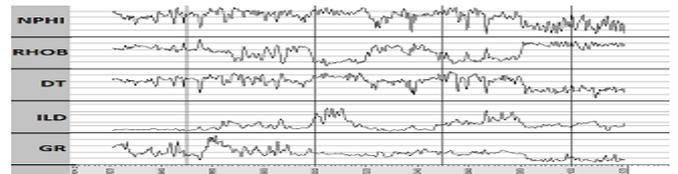


Fig. 2. Perfis de poço.

A. Formulação do problema

O perfil de poço sônico mede a diferença nos tempos de trânsito de uma onda mecânica através das rochas, o qual é utilizado para estimativas da porosidade, correlação poço a poço, detecção de fraturas e apoio à sísmica para a elaboração do sismograma sintético [17], uma ferramenta para a estimação de propriedades geológicas [18]. Este perfil é utilizado principalmente em poços pioneiros, que recebem maiores investimentos para aquisição de informação [19]. Por conseguinte, o perfil sônico não está disponível numa porcentagem considerável dos poços. Com o propósito de utilizar a maior quantidade de informação para a caracterização de reservatórios, deve-se estimar este perfil com a informação de outros perfis disponíveis. Portanto, PGmg pode ser utilizada para estimar a relação, não necessariamente linear, que existe entre o perfil de poço sônico e outros perfis de poço. Estas estimativas inicialmente foram desenvolvidas mediante equações empíricas de relações observadas, depois por

estimacões lineais [20] e atualmente com métodos multiatributos, como redes neurais [21], lógica fuzzy [22] e PG.

Neste trabalho objetiva-se estimar o perfil de poço sônico, nos poços onde não se tem esta informação, utilizando perfurações existentes que apresentem os cinco perfis de poço antes descritos (GR, ILD, RHOB, NPHI e DT). Por meio do algoritmo de PGmg vai se construir uma expressão explícita que represente a relação entre os perfis de poço considerados. O desempenho desta função é avaliado num conjunto de poços não contemplados para o processo de aprendizado, simulando assim, uma estimativa em poços onde o perfil sônico nunca foi registrado. Os resultados gerados mediante PGmg serão validados com estimativas obtidas com RN, por meio do erro MAPE.

IV. PROPAGAÇÃO DA INCERTEZA MEDIANTE A ESTIMAÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADE DA PRODUÇÃO ACUMULADA

A. Formulação do problema

Considere um modelo de simulação de reservatórios de petróleo (f), em função de M parâmetros incertos independentes, $Z = (Z_1, \dots, Z_M)$, por exemplo, permeabilidade, porosidade, compressibilidade da rocha, profundidade dos contatos dos fluidos etc.. Deseja-se estimar alguns estatísticos, ou mais rigorosamente, a distribuição de probabilidade das saídas do modelo de simulação: produção acumulada de óleo, gás e água. Estas análises estatísticas e probabilísticas das respostas fornecem conhecimento da incerteza e da sua propagação através do modelo, assim, é acrescentada a confiança nos resultados da simulação do modelo. O algoritmo de PGmg será utilizado para representar as saídas do modelo de simulação por meio de uma função estocástica. Para encontrar esta expressão explícita é necessário construir um conjunto de treino, validação e teste. Espera-se que, mediante o treinamento e a validação o algoritmo de PGmg aprenda e consiga representar a relação existente entre o conjunto de dados de entrada, $Z \in \mathbb{R}^M$ e a saída do modelo $Y = f(Z) \in \mathbb{R}$, em Fig. 3 pode-se observar este processo. Devido ao modelo apresentar aleatoriedade nas variáveis de entrada, é preciso gerar uma matriz de dados para os conjuntos de treino, validação e teste, seguindo a densidade de probabilidade dos parâmetros de entrada.

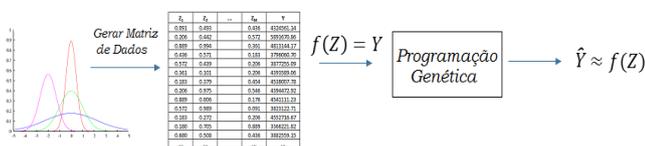


Fig. 3. Programação genética na aproximação da saída do modelo.

A resposta de PGmg (\hat{Y}) é uma função que será considerada um modelo substituto da saída real do modelo (Y), esta função é estocástica e depende da densidade das variáveis de entrada.

B. Estimacão da distribuição de probabilidade da saída do modelo

Uma vez obtida a aproximação da saída do modelo (\hat{Y}), é possível calcular sua distribuição de probabilidade empírica a fim de quantificar sua incerteza. Por definição, a função distribuição de probabilidade de Y , é a função $F(y)$, tal que, $F(y)$

$= P(Y \leq y)$ ou $F(y) = P(f(Z) \leq y)$. Esta função pode ser estimada a partir de um conjunto de N amostras,

$$F(y) \approx \frac{1}{N} \sum_{k=1}^N \mathbf{1}_{(-\infty, y]}(Y_{(k)})$$

Onde $Y_{(k)} = \hat{Y}(\beta_k)$ e os β_k são N variáveis aleatórias independentes geradas conforme a sua densidade de probabilidade, (ver [23]). O α -percentil y_α de Y pode ser aproximado mediante o conjunto ordenado $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(k)}$,

$$y_\alpha \approx Y_{(\lceil \alpha N \rceil)}$$

Finalmente, uma das maiores vantagens do algoritmo de PGmg é obter uma expressão explícita que pode ser manipulada e utilizada como um proxy sob incerteza para estimar a produção acumulada de óleo, gás e água. Uma vez construído o modelo substituto, é possível fazer análises estatísticas e probabilísticas da resposta do simulador, sem a necessidade de simular novamente o campo de petróleo. Desta maneira o custo computacional de realizar estas análises é diminuído consideravelmente. Outra vantagem de PGmg é, caso se tenha uma ideia do comportamento da resposta do modelo, poder utilizar operadores específicos (exp, log, tan) para orientar a construção das soluções candidatas.

V. CASOS DE ESTUDO

Nesta seção é mostrada a capacidade do algoritmo de PGmg para encontrar uma expressão que represente a relação existente entre diferentes dados envolvidos em modelos de reservatórios de petróleo. Primeiramente é apresentado um caso de estudo onde objetiva-se estimar um perfil de poço a partir de outros perfis de poços de um reservatório, utilizando PGmg. Posteriormente, no segundo caso de estudo, estima-se a distribuição de probabilidade da produção acumulada de óleo, água e gás, utilizando a expressão obtida mediante PGmg como um substituto sob incerteza da resposta do simulador.

A. Campo Namorado

O campo Namorado é um reservatório de petróleo localizado a 80 km da costa, ao sudeste do Brasil, no estado do Rio de Janeiro. Ele está situado em zonas de pouca profundidade, sob lâmina d'água entre 110 e 250m, delimitado por fechamentos estratigráficos e estruturais a uma área aproximadamente de 21 km² [24]. É composto principalmente por arenitos e folhelhos e apresenta uma grande complexidade geológica [17]. É uma parte fundamental da bacia sedimentar de Campos, a mais importante já explorada do país [25]. Foi selecionado para este estudo por ser um reservatório extensamente explorado, com um total de 56 perfurações. Somente 18 poços apresentam o perfil sônico, dos quais 15 também têm os outros quatro perfis de poço anteriormente mencionados. Portanto, foram utilizados estes 15 poços completos, para a criação de um modelo preditivo. Os dados foram divididos aleatoriamente da seguinte maneira, onze poços para o conjunto de treinamento e validação e quatro poços para o conjunto de teste. A distribuição espacial destes conjuntos pode ser apreciada em Fig. 4, onde os elementos circulares representam os poços de teste e os quadrados os poços de treino. Deseja-se construir uma função global que represente a relação entre os perfis considerados, e desta

maneira, realizar uma estimativa do perfil sônico, nos 41 poços restantes.

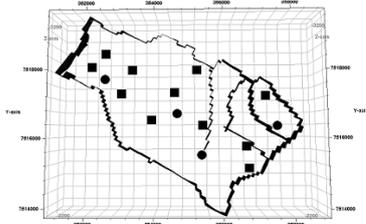


Fig. 4. Distribuição dos poços utilizados do reservatório Namorado.

A relação entre os perfis de poço GR, ILD, NPHI, RHOB e o perfil de poço sônico, foi estudada por meio de uma análise de correlação. Foram selecionados os perfis de poço GR (54.4%), NPHI (81.0%) e RHOB (69.7%), pois apresentaram uma correlação significativa com o perfil de poço sônico. Portanto, estes são os que contém informação mais acertada sobre o comportamento do perfil a estimar. O perfil ILD apresentou uma correlação de 5.7%, por este motivo não foi considerado neste trabalho. A avaliação dos resultados será feita utilizando o erro MAPE, que é uma medida de erro percentual.

A seleção dos parâmetros de PGmg foi realizada mediante múltiplos experimentos realizados sequencialmente, otimizando um a um o valor de cada parâmetro. Os valores são mostrados em TABELA 1. Similarmente, foram selecionados os operadores que compõem a função de saída, estes são: soma, subtração, produto, potência quadrada, seno, cosseno, mínimo e máximo.

TABELA 1. PARÂMETROS DE PG

Parâmetros de PGmg	Valor
Tamanho População	200
Número de Gerações	100
Tamanho do Torneio	4
Profundidade máxima de árvore	7
Número máximo de genes	6
Taxa de cruzamento	0.65
Taxa de mutação	0.3
Taxa de Elitismo	0.05
Intervalo de variação dos nós constantes	[-10,10]

Além desta avaliação, foi utilizada a técnica de RN com o único propósito de validar os resultados obtidos com PGmg. Empregou-se a RN multilayer perceptron (MLP) com erros otimizados mediante backpropagation (BP), implementada no software de interpretação e análise sísmica Petrel. Esta ferramenta faz uma seleção interna do número de camadas escondidas, número de neurônios e funções de ativação, permitindo ao usuário somente determinar a quantidade de iterações. Neste caso foram utilizadas 2000 iterações, este valor foi escolhido a partir de um conjunto de experimentos, realizados para evitar um treinamento excessivo. O conjunto de aprendizado foi o mesmo utilizado na PGmg. Em TABELA 2 podem ser observados os erros MAPE para cada método. Estes erros correspondem à média de dez rodadas independentes para cada um dos métodos.

Ambos os algoritmos conseguem com um erro aceitável, estimar para cada um dos poços de teste o perfil sônico. Destacando que PGmg retorna uma expressão explícita que pode ser utilizada para análises posteriores. Um exemplo disto é a possibilidade de exportar o resultado para softwares de interpretação sísmica.

TABELA 2. COMPARAÇÃO DE RESULTADOS DE PG E RN

Poços de Teste	Erro MAPE nos poços de teste	
	Programação Genética	Redes Neurais
Poço 1	4.53%	4.66%
Poço 2	4.81%	4.90%
Poço 3	3.99%	4.02%
Poço 4	4.27%	4.41%
Média	4.40%	4.50%

Em Fig. 5 são apresentados os resultados dos algoritmos PGmg e RN para analisar visualmente as diferenças e determinar a proximidade entre as estimativas e os valores reais.

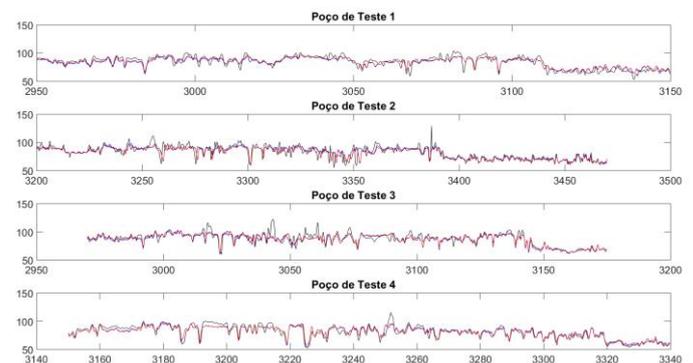


Fig. 5. Resultados da estimativa do perfil sônico por poço com PGmg e RN.

Alguns pontos extremos nos poços de teste não conseguiram ser atingidos por nenhum dos métodos, no entanto o comportamento do perfil é capturado tanto por PGmg como por RN. A função de saída construída por PGmg foi utilizada posteriormente para o cálculo dos perfis faltantes.

B. Campo PUNQ-S3

A fim de medir a incerteza propagada nos modelos de simulação de reservatórios de petróleo, foi considerado o campo PUNQ-S3. O modelo determinístico do reservatório foi tomado a partir de um estudo de engenharia industrial, baseado num campo real que foi operado por Elf Exploration Production e desenvolvido no projeto PUNQ (Production forecasting with UNcertainty Quantification). A descrição completa pode ser encontrada em [26]. O modelo numérico contém 19x28x5 blocos, com 1761 ativos. O reservatório produz mediante 10 poços. Além disso, é limitado por uma falha ao leste e sul e tem comunicação com um forte aquífero ao oeste e norte, por isto não precisa de poços injetores.

O modelo de simulação considerado neste artigo, tem três parâmetros incertos independentes: porosidade, permeabilidade horizontal e vertical. Estes afetam a resposta e portanto tem

sentido propagar a incerteza presente neles. A porosidade mede a capacidade de uma rocha de poder armazenar fluidos em seus espaços interiores, chamados poros. Os poros podem ser preenchidos por gases, águas ou óleo. Assim, a porosidade é definida como sendo a relação entre o volume de vazios de uma rocha e o volume total da mesma [27]. Esta característica geológica toma valores no intervalo [0,1]. Por outro lado, a permeabilidade de um meio poroso é a medida de sua capacidade de se deixar atravessar por fluidos. É uma medida da condutividade de fluidos de um material. A permeabilidade não possui um valor uniforme em todo o campo, na maioria das rochas porosas, há uma variação da permeabilidade com respeito à posição considerada. Na prática observa-se que a permeabilidade segue uma distribuição log-normal [27]. Portanto, vão ser consideradas estas propriedades como variáveis aleatórias com distribuição de probabilidade log-normal, de média um e desvio padrão zero, truncadas para manter as propriedades geológicas.

O objetivo do estudo é quantificar a incerteza na produção acumulada de óleo, água e gás, e como não existe solução analítica, os resultados vão ser comparados com Monte Carlo de 300,000 simulações. Mediante PGmg tenta-se aproximar cada uma das saídas do modelo de simulação de reservatórios de petróleo, por meio de uma função analítica estocástica. Estas expressões foram utilizadas como modelos substitutos (*proxies*) sob incerteza, para estimar a média, variância e mais rigorosamente, a distribuição de probabilidade empírica. Os parâmetros de PGmg foram selecionados depois de uma grande quantidade de experimentos realizados, de maneira sequencial, otimizando um a um o valor de cada parâmetro. Em TABELA 3, podem ser observados os valores escolhidos. Similarmente, foram otimizadas as funções que compõem as soluções candidatas do algoritmo de PGmg, estas são: soma, subtração, produto, potência quadrada, mínimo e máximo.

TABELA 3. PARÂMETROS DE PG

Parâmetros	Óleo	Gás	Água
Tamanho População	30	30	30
Número de Gerações	850	850	850
Tamanho do Torneio	4	4	4
Profundidade máxima de árvore	3	3	3
Número máximo de genes	20	20	20
Taxa de cruzamento	0.65	0.65	0.65
Taxa de mutação	0.3	0.3	0.3
Taxa de Elitismo	0.05	0.05	0.05
Intervalo de variação dos nós constantes	[-10,10]	[10,10]	[10,10]
Amostras no Treinamento	40	60	80
Amostras na Validação	20	30	50
Simulações totais (Treino + Validação)	60	90	130

A quantidade de simulações requeridas (treino e validação) para cada uma das funções obtidas foi determinada por meio de uma sequência de experimentos, reduzindo progressivamente o número de amostras, sem uma perda significativa no coeficiente de determinação no teste. Deve-se destacar a eficiência em termos de custo computacional, medido no número de chamadas ao simulador, se comparado com a técnica de Monte Carlo tradicional. Em TABELA 4 são apresentadas as médias e variâncias, aproximadas pelas soluções obtidas por meio de PGmg. Pode-se observar a proximidade dos valores alcançados

com os valores obtidos pelo método de simulação de Monte Carlo. As maiores variações são apresentadas na variância, enquanto a média apresenta-se estável.

TABELA 4. MÉDIA E VARIÂNCIA

Saída	Técnica	Média	Variância
Óleo	Monte Carlo	4,060,102.18	1,791,266,013,125.12
	PGmg	4,101,232.57	1,829,258,367,026.30
Gás	Monte Carlo	363,106,047.35	11,153,185,807,589,300.00
	PGmg	361,563,693.27	12,413,092,352,971,800.00
Água	Monte Carlo	5,019,276.05	32,386,882,380,460.30
	PGmg	4,922,285.04	32,715,248,581,023.80

Em Fig. 6, Fig. 7 e Fig. 8 são mostradas estimativas das funções distribuição de probabilidade das saídas com PGmg e Monte Carlo.

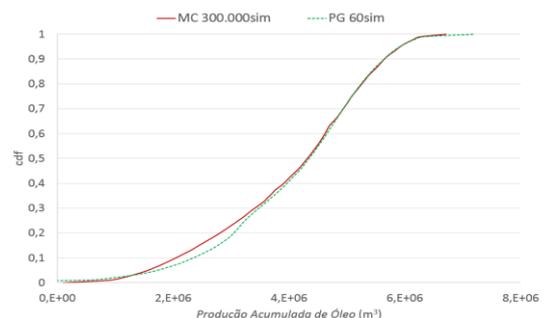


Fig. 6. Distribuição empírica da produção acumulada de óleo.

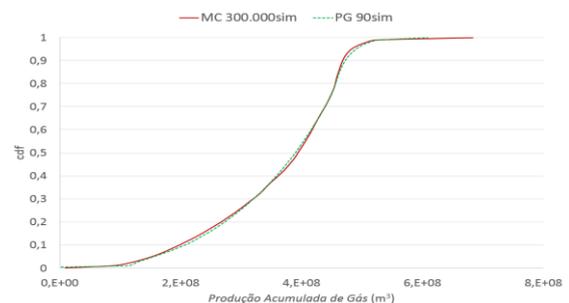


Fig. 7. Distribuição empírica da produção acumulada de gás.

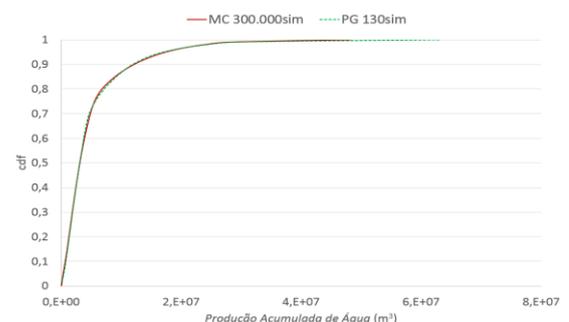


Fig. 8. Distribuição empírica da produção acumulada de água.

Pode-se observar nos três gráficos, que PGmg consegue aproximar com exatidão as distribuições de probabilidade. Os percentis estimados para cada produção acumulada estão próximos aos valores gerados pelo Monte Carlo, inclusive os altos que têm maior dificuldade para ser aproximados. Além

disso, foi calculado o coeficiente de determinação, comumente denotado por R^2 , com o fim de analisar o ajuste da estimação obtida por PGmg. Os valores são, para a produção acumulada de óleo $R^2=0.936$, para água $R^2=0.973$ e finalmente para gás $R^2=0.965$. Com isto é confirmada a eficiência do ajuste mostrado nos gráficos. A função de saída estimada consegue representar a relação entre os dados de entrada e a resposta dada pelo modelo, tendo em consideração a aleatoriedade dos parâmetros de entrada.

VI. CONCLUSÃO

Neste artigo, foram apresentadas duas aplicações de programação genética multi-gene, a primeira para estimar um perfil de poço por meio de outros perfis no reservatório Namorado. A segunda aplicação foi para aproximar a distribuição de probabilidade da produção acumulada num modelo de simulação do reservatório PUNQ-S3, considerando a incerteza presente em alguns parâmetros de entrada. PGmg foi utilizada para obter nas duas aplicações, uma expressão para estimar as saídas desejadas. Este algoritmo multi-gene conseguiu, em ambos os casos, aprender a partir dos dados fornecidos para o treinamento, a relação não linear existente entre as entradas e a saída do problema considerado, além de construir uma função analítica para representá-la. Para a estimação do perfil sônico, os resultados obtidos mediante PGmg foram satisfatórios, apresentando um erro MAPE médio menor do que 4.5%. A validação dos resultados foi realizada por meio de comparação com redes neurais e observou-se que PGmg consegue, em todos os poços de teste, um erro menor do que os apresentados pelas redes neurais. Destaca-se que, a propriedade estimada e as variáveis que se relacionam a ela, estão presentes unicamente nos poços. Em um caso onde as variáveis de entrada estiverem disponíveis, em todo o reservatório, será possível realizar a estimação completa no campo, e ter assim um mapa 3D da propriedade. Na aproximação da distribuição empírica da produção acumulada de óleo, água e gás, o algoritmo de PGmg alcançou valores próximos aos resultados obtidos por Monte Carlo. Os percentis altos foram estimados com um nível de exatidão eficiente e com um custo computacional baixo. Além disso, uma das maiores vantagens do algoritmo de PGmg é, obter uma expressão explícita que pode ser usada como um proxy sob incerteza da produção acumulada de óleo, gás e água. Após a construção do modelo substituto, é possível quantificar a incerteza nestas saídas do modelo, mediante o cálculo de estatísticas e análises probabilísticas, sem a necessidade de simular novamente o modelo do campo de petróleo, e assim, diminuir consideravelmente o custo computacional destes estudos.

REFERENCES

- [1] S. Mohaghegh and S. S. Ameri, Artificial neural network as a valuable tool for petroleum engineers, 1995.
- [2] S. Mohaghegh, Recent developments in application of artificial intelligence in petroleum engineering, 2005, pp. 86-91.
- [3] M. Pacheco and M. M. Vellasco, Intelligent systems in oil field development under uncertainty, vol. 183, Springer, 2009.
- [4] T. Yu, D. Wilkinson and A. Castellini, "Constructing reservoir flow simulator proxies using genetic programming for history matching and production forecast uncertainty analysis," vol. 2, 2008.
- [5] E. Burke, M. Hyde, G. Kendall, G. Ocho and E. Ozcan, "Exploring hyper-heuristic methodologies with genetic programming," pp. 177-201, 2009.
- [6] H. Kaydani, A. Mohebbi and M. Eftekhari, "Permeability estimation in heterogeneous oil reservoirs by multi-gene," *Journal of Petroleum Science and Engineering*, vol. 123, p. 201–206, 2014.
- [7] H. Tang, N. Toomey e S. Meddaugh, "Using an artificial-neural-network method to predict carbonate well log facies successfully," *SPE Journal Paper*, 2011.
- [8] M. Van der Baan e C. Jutten, "Neural networks in geophysical applications," *Geophysics*, vol. 65, pp. 1032-1047, 2000.
- [9] A. A. Taghavi, "Improved permeability estimation through use of fuzzy logic in a carbonate reservoir from southwest, Iran," 2005.
- [10] P. Sarma, L. Durlofsky e K. Aziz, "Efficient closed loop production optimization under uncertainty," 2005.
- [11] P. Sarma, L. Durlofsky, K. Aziz e W. Chen, "Efficient real-time reservoir management using adjoint-based optimal control and model updating," vol. 10, pp. 3-36, 2006.
- [12] L. Heng e P. Sarma, "A comparative study of the probabilistic-collocation and experimental-design methods for petroleum-reservoir uncertainty quantification," vol. 16, pp. 429-439, 2011.
- [13] D. Oliver, A. Reynolds e N. Liu, Inverse theory for petroleum reservoir characterization and history matching, Cambridge University Press, 2008.
- [14] R. Poli, W. Langdon, N. Mcphee and J. Koza, A field guide to genetic programming, Lulu.com, 2008.
- [15] J. R. Koza, Genetic programming: on the programming of computers by means of natural selection, vol. 1, MIT press, 1992.
- [16] D. Searson, D. Leahy and M. Willis, "GPTIPS: an open source genetic programming toolbox for multigene symbolic regression," vol. 1, pp. 77-80, 2010.
- [17] A. C. Vidal, S. S. Sancevero, A. Z. Remacre and C. P. Constanzo, "Modelagem geoestatística 3D da impedância acústica para a caracterização do Campo de Namorado," vol. 25(3), pp. 295-305, 2007.
- [18] J. E. Thomas, Fundamentos de engenharia de petróleo, Interciência, 2001.
- [19] E. S. Chagas, S. L. Russo and V. H. Simon, "Estimativa da perfilagem em poços de petróleo utilizando a estatística multivariada," Simpósio Nacional de Probabilidade em estatística, São Paulo 19, 2010.
- [20] S. Adcock, "In search of the well tie: What if I don't have a sonic log?," vol. 12(12), pp. 1161-1164, 1993.
- [21] M. Silva, R. Santos and S. Fontoura, "Predição de propriedades de perfis com redes neurais: uma aplicação no campo escola de Namorado, Bacia de Campos, Brasil," p. 143, 2001.
- [22] M. Leite, A. Carrasquille and J. D. Silva, "Simulação do perfil sônico a partir dos perfis de raios gama e de resistividade em poços da Bacia de Campos," vol. 26(2), Revista Brasileira de Geofísica, 2008, pp. 141-151.
- [23] J. Ko and J. Garnier, "Multi-element stochastic spectral projection for high quantile estimation," vol. 243, pp. 87-108, 2013.
- [24] S. X. Meneses and T. Adams, "Ocorrência de resistividades anômalas no Campo de Namorado, Bacia de Campos," vol. 4, no. 2, 1990.
- [25] Petrobras, "Petrobras," [Online]. Available: <http://www.petrobras.com.br/pt/nossas-atividades/principais-operacoes/bacias/bacia-de-campos.htm>. [Accessed 05 2015].
- [26] Imperial College London, "PUNQ-S3," [Online]. Available: <http://www3.imperial.ac.uk/earthscienceandengineering/research/perm/punq-s3model>. [Accessed 05 2015].
- [27] A. Rosa, R. De Souza Carvalho and J. Xavier, Engenharia de reservatórios de petróleo, Interciência, 2006.