



Copyright 2004, Instituto Brasileiro de Petróleo e Gás - IBP

Este Trabalho Técnico Científico foi preparado para apresentação no 3º Congresso Brasileiro de P&D em Petróleo e Gás, a ser realizado no período de 2 a 5 de outubro de 2005, em Salvador. Este Trabalho Técnico Científico foi selecionado e/ou revisado pela Comissão Científica, para apresentação no Evento. O conteúdo do Trabalho, como apresentado, não foi revisado pelo IBP. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões do Instituto Brasileiro de Petróleo e Gás, Sócios e Representantes. É de conhecimento e aprovação do(s) autor(es) que este Trabalho será publicado nos Anais do 3º Congresso Brasileiro de P&D em Petróleo e Gás

Melhoria da convergência de Algoritmo Genético utilizado em alocação de poços

Luis Fernando Pozas¹, Agostinho Plucenio²

¹ UFSC, DAS – Trindade - Florianópolis, SC, Brasil, lfpozas@das.ufsc.br

² UFSC, DAS – Trindade, Florianópolis, SC, Brasil, plucenio@das.ufsc.br

Resumo – No estudo de otimização da alocação de poços, busca-se através da manipulação das variáveis de decisão - a localização, a taxa de injeção e a taxa de produção dos poços - aumentar o total de hidrocarbonetos recuperado ou ainda, do ponto de vista econômico, aumentar o Valor Presente Líquido (VPL), em inglês (NPV). Os Algoritmos Genéticos são uma das ferramentas que têm sido utilizadas para alcançar este objetivo. Apresenta-se neste trabalho a contribuição de um algoritmo genético modificado que agrega algoritmos de convergência e técnicas de identificação com o objetivo de acelerar a convergência e melhorar a qualidade da solução final.

Palavras-Chave: Alocação de Poços; Algoritmo Genético; Politopo;

Abstract – In the well placement optimization problem the goal is to raise NPV by manipulation of decision variables, such as well location, injection and production rates. Since this is a hard combinatorial problem, Genetic Algorithm has been used as one way of solving it. The authors developed in the present work a hybrid genetic algorithm, which combines polytope algorithm and identification techniques. Our goal is to speed up convergence rate and improve the quality of final solution.

Keywords: Well Placement; Genetic Algorithm; Polytope

1. Introdução

Muitas vezes no exercício cotidiano da Engenharia de Petróleo, as decisões tomadas não obedecem a um critério de otimização. O fato de não se procurar a melhor solução para o problema, ou pelo menos uma solução próxima do ótimo, pode fazer com que pontos importantes de projetos passem despercebidos ao grupo multidisciplinar de profissionais encarregados de resolver o problema. Esta negligência pode resultar em grandes perdas econômicas e técnicas. A principal tarefa da engenharia de reservatórios é desenvolver uma estratégia para alcançar a produção da maior quantidade possível de hidrocarbonetos dentro dos limites físicos e econômicos existentes. As duas questões que determinam o nível de produção são: o sistema de exploração a ser instalado *in situ* e o reservatório geológico existente.

A previsão do comportamento dos reservatórios pode ser realizada através de simulação numérica. O **modelo de simulação** se constitui na principal ferramenta para tomada de decisão. Através deles, uma função-objetivo (f_{obj}), que quantifica matematicamente o objetivo global do projeto, deve ser maximizada através da variação dos parâmetros de entrada do modelo.

No caso específico da otimização da alocação de poços, as variáveis de decisão geralmente são as localizações dos poços. Algumas vezes considera-se também a taxa de injeção e/ou de produção dos poços. A saída, ou função objetivo, é geralmente medida em termos do total de hidrocarbonetos recuperado ou, em termos econômicos, o Valor Presente Líquido (VPL), em inglês (NPV). O VPL é calculado basicamente como a diferença entre o valor presente e o custo de desenvolvimento, $VPL = VP - D$, aonde VP é o valor presente e D é o custo de desenvolvimento.

O Custo de Desenvolvimento D é calculado como:

$$D = (a + r)n_p + b + \sum_{j=1}^{n_p} |p_j - P_{PL}| \cdot c, \quad (1)$$

onde

- a - Custo médio de perfuração para cada poço mais o custo da “Árvore de Natal”;
- r - Custo do riser que corresponde ao tamanho da linha d'água multiplicado pelo custo de linha por quilômetro;
- b - Custo de transferência mais o custo de plataforma e planta;
- p_j - Posição do poço j ;
- P_{PL} - Posição da plataforma;
- c - Custo de linha por quilômetro;
- n_p - Número de poços.

Normalmente, estes valores são conhecidos a priori e independem do tempo para cada alternativa selecionada. O valor presente VP representa um valor atribuído no presente momento para objetos ou eventos que não são realizados neste momento. Isto implica na existência de um fator de ajuste que depende da distância temporal existente entre o “hoje” e o dia da precificação do evento a realizar-se. Assim, o valor presente do reservatório é composto pela diferença entre o valor presente da receita VP_R menos o valor presente do custo de operação VP_{Cop} , na qual é aplicada a alíquota de imposto I :

$$VP = (VP_R - VP_{Cop})(1 - I) \quad (2)$$

A possibilidade de se obter diferentes perfis de produção para diferentes configurações pela simulação destes modelos, sugere o desenvolvimento de um método iterativo para encontrar uma configuração que forneça a melhor estratégia, isto é, a alternativa ótima do ponto de vista econômico. Define-se uma alternativa de desenvolvimento como o número e localização de poços produtores e injetores em um campo petrolífero. Portanto, a otimização da alternativa é a otimização do número, da localização e do tipo dos poços. Alguns trabalhos na literatura propõem a resolução do problema de otimização através de algoritmos genéticos (AG). Cita-se nas referências alguns trabalhos sobre a otimização de alocação de poços, que demonstram a validade e a atualidade do assunto (Baris; Wang; Túpac, Vellasco e Pacheco; Junior).

Apesar dos ganhos obtidos com a utilização dos simuladores, cabe ressaltar que estes modelos são bastante complexos e suas simulações demandam um grande esforço computacional o que resulta em um tempo de simulação também grande. Neste sentido, em alguns trabalhos, como por exemplo em Baris (2000), são desenvolvidos AG's híbridos com o intuito de acelerar a convergência do AG e minimizar o acionamento do simulador de reservatório. Neste último em particular, o AG conta com a implementação adicional de um método de extrapolação (Politopo) e um método de interpolação (Kriegging).

Pretende-se neste artigo contribuir com o trabalho de Baris (2000,2001), aumentando a velocidade e precisão de convergência do AG, através da implementação adicional de um algoritmo de identificação. Foca-se a melhoria do desenvolvimento do AG e por isso, para estudo de caso, utilizou-se funções numéricas ao invés dos simuladores propriamente ditos.

A hibridização do AG com a implementação adicional destes algoritmos ajuda a superar uma de suas dificuldades fundamentais, a saber, encontrar facilmente regiões promissoras no espaço de busca, mas ser incapaz de refinar a busca dentro das regiões encontradas.

1.1. Algoritmo Genético – Codificação do Cromossoma

A função numérica utilizada no presente trabalho como modelo do NPV do simulador é:

$$\frac{(1 - 0.1x_1) \sin(0.1x_1) (1 - 0.1x_2) \sin(0.1x_2)}{0.1(0.1(x_1 - 10)^2 + 1) (0.1(0.1x_2 - 10)^2 + 1)} \quad (3)$$

onde x_1 e x_2 são as coordenadas da localização do poço no campo.

Na presente aplicação, cada cromossoma da população representa as coordenadas x_1 e x_2 da função avaliada. As coordenadas estão limitadas em um intervalo entre 0 e 511, e são codificadas através da transformação da coordenada da base 10 para seu equivalente na base 2 em uma palavra binária com nove algoritmos. A f_{obj} utilizada é a própria função que descreve a NPV do campo. Assim, por exemplo, o indivíduo (010000010|000111011) representa a coordenada $(x_1, x_2) = (130, 59)$, e sua avaliação da f_{obj} (no caso da função 3) vale -0.0239 . Em outras palavras, um poço localizado em $(x_1, x_2) = (130, 59)$ teria um NPV de -0.0239 .

Tabela 1. Exemplo de codificação do cromossoma

(x_1, x_2)	Indivíduo	Avaliação
(130, 59)	1000001000111011	-0.0239

1.2. Politopo

O algoritmo Politopo (AP), também conhecido como método simplex, foi utilizado como um método auxiliar para refinar a procura do AG. Este método, introduzido por Nelder e Mead (1965), não requer o cálculo de derivadas. Um politopo é uma figura geométrica na dimensão n constituída de segmentos de linha interconectando seus $n+1$ vértices. Em duas dimensões um politopo é um triângulo, em três dimensões um tetraedro e assim por diante. A cada passo do algoritmo, um novo vértice é gerado. Se o novo vértice é melhor do que pelo menos um dos vértices existentes, ele substitui o pior vértice. Este algoritmo apresenta como desvantagem demandar muitas avaliações da f_{obj} , o que corresponderia a acionar o simulador muitas vezes. No entanto, as soluções que são obtidas através dele são muito úteis no sentido de gerar indivíduos mais aptos para a nova população do AG.

1.3. Algoritmo de Identificação

A identificação é uma técnica muito utilizada na área de controle. Basicamente, a idéia é gerar uma função cuja performance se aproxime suficientemente bem da função original em torno dos pontos escolhidos. No presente trabalho, procurou-se gerar uma função quadrática através tratamento matemático detalhado a seguir.

Sejam sete os pontos escolhidos para fazer a identificação (x_{11}, x_{21}) a (x_{17}, x_{27}) , e f_{av_1} a f_{av_7} suas respectivas avaliações da f_{obj} . A partir dos sete pontos, e de suas avaliações, são formadas as matrizes:

$$X = \begin{bmatrix} x_{11} & x_{21} & x_{11}^2 & x_{21}^2 & 1 \\ x_{12} & x_{22} & x_{12}^2 & x_{22}^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{17} & x_{27} & x_{17}^2 & x_{27}^2 & 1 \end{bmatrix} \quad e \quad Y = \begin{bmatrix} f_{av_1} \\ f_{av_2} \\ \vdots \\ f_{av_7} \end{bmatrix} \quad (4)$$

Busca-se então gerar uma superfície convexa que passe por estes pontos. Para tanto, faz-se a identificação dos parâmetros da seguinte função:

$$k_1 \cdot x_1 + k_2 \cdot x_2 + k_3 \cdot x_1^2 + k_4 \cdot x_2^2 + k_5 \cdot 1 = F_{av} \quad (5)$$

Utilizando as informações contidas nos pontos selecionados faz-se $X \cdot K = Y$, aonde $K = [k_1 \ k_2 \dots \ k_5]^T$. Portanto, para se obter K faz-se $K = \text{pinv}(X) \cdot Y$. Para realizar a inversão de matrizes que não sejam quadradas, utiliza-se a operação pseudo-inversa, representada como pinv . Como os parâmetros estimados k são em número de cinco, optou-se por selecionar para identificação um conjunto de sete pontos de maneira a garantir um grau de liberdade razoável.

A motivação de se utilizar a identificação aqui é que espera-se que uma vez que o AG já tenha atingido um certo grau de convergência, a superfície obtida com a identificação se aproxime bastante da função original em torno do ponto de máximo. Assim, através do ponto de máximo obtido com a superfície refinar-se-ia a convergência do AG.

1.4. Híbridação do AG

Neste trabalho, para garantir que a solução obtida pelo AG já tenha atingido um certo nível de convergência, obtém-se quarenta novas gerações de indivíduos através da utilização do AG puro. Logo em seguida, é acionado o AP,

onde um número pré-definido de politopos é construído (N_{POLY}). Para cada politopo criado, se a solução obtida for melhor do que a solução do indivíduo com pior avaliação da f_{obj} , ela o substitui, caso contrário, a solução do politopo é descartada. Conclui-se, portanto, que há desperdício de um certo número de avaliações da f_{obj} , número que depende do ajuste do número de expansões e contrações utilizadas.

Após a aplicação de politopo, o algoritmo de identificação (AI) é acionado. O algoritmo busca, dentre os indivíduos que pertencem à última população gerada (já incluídos os gerados por politopo), os que representem os sete pontos com maior avaliação da f_{obj} . Os pontos escolhidos têm ainda que atender à restrição que suas distâncias ao ponto de maior avaliação da f_{obj} não exceda uma dada constante (para garantir que estejam próximos do ponto de máximo), e também de que a distância entre eles esteja dentro de uma faixa aceitável (para garantir que os pontos sejam representativos). Caso o AI não consiga selecionar mais do que 4 pontos, o algoritmo termina. Caso seja selecionado um número entre 4 e 6 pontos, geram-se novos politopos para completar os sete pontos para identificação. Se ainda assim os pontos não atenderem aos requisitos, o AI é suspenso. Assim, por exemplo, se não mais do que cinco pontos forem selecionados pelo critério de avaliação e de distância, o AP é acionado duas vezes para obter os dois pontos que restam.

2. Resultados Experimentais

Em um primeiro momento, buscou-se determinar o melhor ajuste dos parâmetros do AG puro, quando a f_{obj} avaliada foi a Função 3, cuja visualização tridimensional está mostrada na Figura 1. Assim, foram obtidas dez simulações para diferentes ajustes de *taxa de crossover* (r) e *taxa de mutação* (m), obtendo-se os resultados demonstrados na Tabela 2 para 50 gerações e população de 40 indivíduos, e para uma população de 50 indivíduos. Outros ajustes de tamanho de população foram avaliados, mas não se mostraram interessantes.

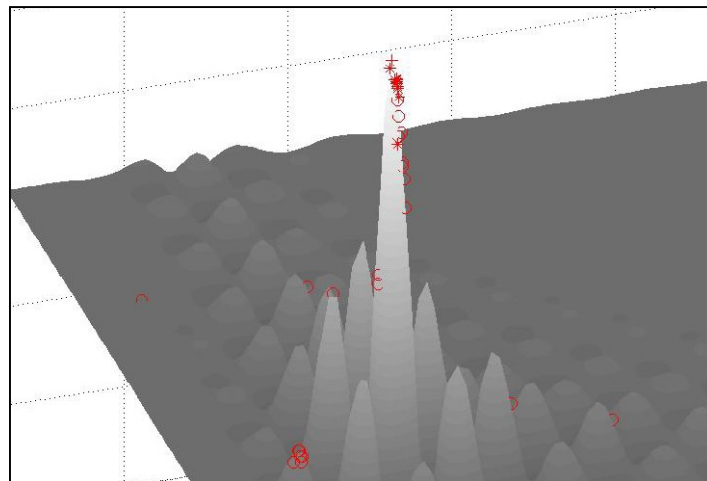


Figura 1. Visualização tridimensional da Função 3.

A partir destes dados, decidiu-se ajustar o AG com uma *taxa de crossover* de 50%, *taxa de mutação* de 20%, e população com 50 indivíduos.

Tabelas 2 – Avaliação máxima atingida para diferentes ajustes dos parâmetros do AG.

Nº da Simulação	Pop. = 40 indivíduos						Pop. = 50 indivíduos					
	$r=0.4$ $m=0.2$	$r=0.5$ $m=0.2$	$r=0.3$ $m=0.2$	$r=0.4$ $m=0.1$	$r=0.5$ $m=0.1$	$r=0.3$ $m=0.1$	$r=0.4$ $m=0.2$	$r=0.5$ $m=0.2$	$r=0.3$ $m=0.2$	$r=0.4$ $m=0.2$	$r=0.5$ $m=0.2$	$r=0.3$ $m=0.2$
1	41.57	15.36	17.79	16.08	40.08	42.34	14.25	14.49	17.64	13.16	17.37	20.46
2	17.45	19.64	17.03	41.89	18.01	17.47	19.62	43.05	42.46	42.34	15.70	11.39
3	17.71	43.28	17.68	17.37	40.07	18.47	14.45	17.13	18.19	41.57	20.44	17.79
4	39.15	42.74	18.28	41.57	14.26	16.34	40.72	40.67	40.67	16.89	43.41	14.89
5	39.55	17.44	40.03	42.46	17.64	42.35	42.35	41.35	18.01	40.58	43.05	14.50
6	42.12	19.51	40.58	15.21	17.37	14.29	41.89	43.02	17.05	40.67	17.03	10.60
7	13.90	42.34	41.35	14.72	17.03	40.67	18.20	41.17	40.60	42.35	17.11	40.67
8	40.03	41.21	18.31	18.22	17.62	40.08	41.74	42.47	41.29	16.35	17.69	20.49
9	40.59	18.01	42.43	10.83	11.09	18.20	40.67	43.29	42.46	40.03	13.60	12.05
10	17.93	17.68	42.34	19.65	41.74	13.96	40.67	43.06	18.10	14.51	12.85	41.34
Média	31.00	27.72	29.58	23.80	23.49	26.42	31.46	36.97	29.65	30.85	21.82	20.42

Uma vez ajustado o AG puro, passa-se à escolha da estratégia para utilização do AP. Em um primeiro momento optou-se por aplicar este algoritmo a cada geração de uma nova população, subsequentemente às operações de *crossover* e *mutação*. Esta estratégia não se mostrou interessante, pois, apesar de acelerar a convergência do AG apresentou o inconveniente de induzir à convergência prematura em torno de um ótimo local, e não do global. Analisou-se então uma pequena modificação desta estratégia onde o AP ainda seria acionado a cada nova geração, mas somente após um certo nível de convergência ter sido alcançado. Esta estratégia também não se mostrou interessante já que não há grandes melhorias na convergência que façam o esforço computacional demandado pelo AP valer a pena. Por fim, optou-se por um esquema em série onde o AG puro é utilizado em uma primeira etapa até um nível de convergência razoável ser atingido. Para garantir a convergência, fez-se uma análise da relação entre o número de gerações e o nível de convergência, como consta na Tabela 3, onde há um indicador da porcentagem de vezes em que o máximo atingido esteve entre 90% e 100% do valor do ótimo global após 20, 25, 30, 35 e 40 gerações. Adotou-se então o número de 40 gerações.

Tabela 3 – Avaliação máxima atingida

Nº da simulação	até 20 ger.	Até 25 ger.	Até 30 ger.	Até 35 ger.	Até 40 ger.
1	43.17	17.71	40.67	41.57	40.67
2	42.46	40.59	17.49	43.41	40.67
3	16.82	20.19	16.36	17.72	43.29
4	04.00	32.39	17.41	41.57	38.07
5	05.84	14.69	38.07	41.34	43.05
6	06.82	18.22	17.64	20.08	43.29
7	43.29	39.81	42.79	17.79	41.57
8	09.45	41.35	19.86	43.29	43.41
9	19.46	15.79	40.66	13.88	17.47
10	42.35	41.57	43.17	17.72	41.35
Pct.(%)	40	40	40	50	80

Em uma segunda etapa, busca-se então 10 pontos novos com o AP, para fazer o primeiro refinamento da busca. Em uma terceira etapa utiliza-se o AI. O AI gera a superfície que passa pelos sete pontos selecionados, e retorna as coordenadas (x_{1max} , x_{2max}) do ponto de máximo desta superfície. Apresenta-se na Figura 2 um exemplo de superfície gerada.

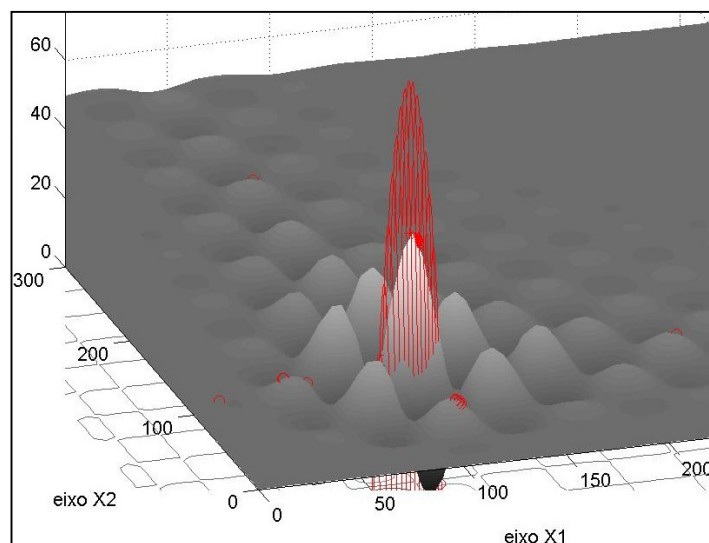


Fig. 2 – Exemplo de superfície gerada

Como se pode notar, ela não é exatamente igual à superfície gerada pela função original. No entanto, ela possui a curvatura no mesmo sentido que a curvatura da superfície original, e as coordenadas de seu ponto de máximo estão bastante próximas das coordenadas do ótimo global. A coordenada obtida é aplicada na função original. Caso a avaliação obtida seja melhor, ela substitui o indivíduo com pior avaliação da população. Mostra-se na Figura 3 como cada algoritmo contribui para a solução geral. Nesta figura, os pontos marcados com um 'o' são os pontos gerados pelo AG, os pontos marcados com um '*' são os pontos gerados pelo politopo, e o ponto marcado com um '+' é o ponto de máximo obtido com a identificação.

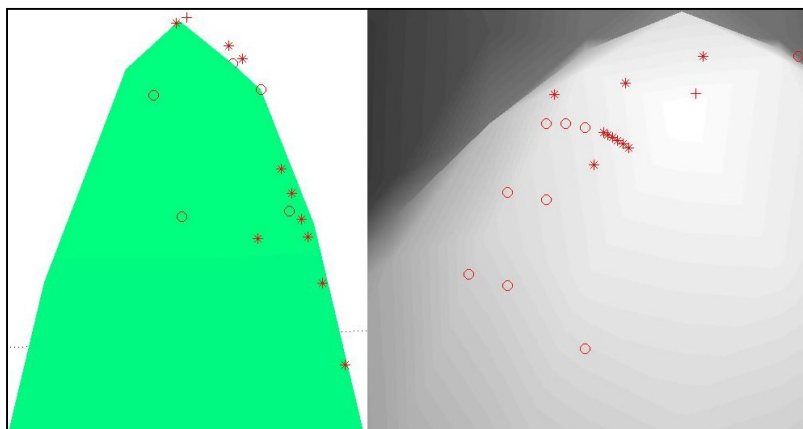


Fig. 3 – Vista Lateral e Vista Superior das contribuições de AG (o), AP (*) e AI (+)

Em 90% dos casos o valor obtido com a identificação esteve dentro de uma faixa de 5% de diferença para o valor do ótimo global, e em 70% dos casos dentro de uma faixa de 2%, sendo que por duas vezes o máximo global foi atingido.

3. Conclusões

A inclusão do algoritmo de identificação contribuiu de maneira significativa na aceleração da velocidade de convergência, sem a penalização do acesso à f_{obj} . Em outros trabalhos foi utilizado apenas o AP na hibridização, o que já produziu uma melhora na convergência. No entanto, cabe lembrar que este é muito custoso em termos de avaliações da f_{obj} . Busca-se aqui, com seu uso, apenas gerar pontos mais bem distribuídos para que a técnica de identificação seja mais precisa. A grande vantagem da técnica de identificação é que, atendidas as restrições, o ponto de máximo obtido a partir dela se encontra bastante próximo do ótimo global, e praticamente obtém-se o ótimo global após uma iteração. Isto aumenta consideravelmente a velocidade de convergência e a qualidade da resposta final.

A f_{obj} utilizada neste trabalho é bastante peculiar o que torna os resultados mais interessantes ainda. Ela possui um certo número de picos com ótimos locais em torno do ótimo global, e a diferença entre o valor dos ótimos locais e do ótimo global não é tão grande, o que faz com que o algoritmo tenha que ser mais seletivo, o que algumas vezes impediu o AG de convergir. Em outros testes feitos com funções menos 'seletivas', foram necessárias poucas iterações para se atingir o ótimo global.

Pôde observar-se também que a convergência do AG tem uma grande dependência com o número de gerações, havendo um limite inferior abaixo do qual todas as técnicas e configurações de ajustes aqui testados não produziram resultados significativos.

4. Referências

- NELDER, J., e MEAD, R. A Simplex Method for Function Minimization. *Computer Journal*, n 7, pp. 308 – 313, 1965.
- BARIS, G. e HORNE, R. Optimization of well placement. Stanford University, 2000.
- BARIS, G. e N. HORNE, R. Uncertainty Assessment of Well Placement Optimization. *SPE Annual Technical Conference and Exhibition*, New Orleans, U.S.A., Setembro, 2001.
- WANG, P.; LITVAK, M.P. e AZIZ, K. Optimization of Production from Mature Fields. Stanford University, 2002.
- TÚPAC, J. Y; VELLASCO, M.R. M. e PACHECO, M.M.A.C. Seleção de alternativas para o desenvolvimento de campos de petróleo por algoritmo genéticos. PUC-RJ, *Congresso Brasileiro de Petróleo - Natal*, Brasil. 2001.
- JUNIOR, P. C. Otimização de locações de poços usando simulação numérica de reservatórios. Unicamp, *Dissertação de Mestrado*, pp 141, 1999.
- CHRISTIE, M.; SUBBEY, S. e SAMBRIDGE, M. Prediction under Uncertainty in Reservoir Modeling. *8th European Conference on the Mathematics of Oil Recovery*, Freiberg, Alemanha, Setembro, 2002.

5. Agradecimentos

Agradeço o apoio financeiro da Agência Nacional do Petróleo - ANP, e da Financiadora de Estudos e Projetos - FINEP, por meio do Programa de Recursos Humanos da ANP para o Setor Petróleo e Gás PRH-34 ANP/MCT. Agradeço também ao professor Eduardo Camponogara pelo apoio e pelas dicas.