



**UNIVERSIDADE FEDERAL DO PARÁ
CENTRO DE GEOCIÊNCIAS
CURSO DE PÓS-GRADUAÇÃO EM GEOFÍSICA**

**MÉTODOS DE OTIMIZAÇÃO GLOBAL APLICADOS
NA BUSCA DOS PARÂMETROS SRC**

Klaus Mairan Laurido do Carmo



**Belém – Pará
2004**

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO DE GEOCIÊNCIA
CURSO DE PÓS-GRADUAÇÃO EM GEOFÍSICA

DISSERTAÇÃO DE MESTRADO POR
Klaus Mairan Laurido do Carmo

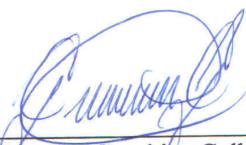
MÉTODOS DE OTIMIZAÇÃO GLOBAL APLICADOS
NA BUSCA DOS PARÂMETROS SRC

COMO REQUISITO PARCIAL Á OBTENÇÃO DE GRAU DE MESTRE EM
CIÊNCIAS NA ÁREA DE GOFÍSICA

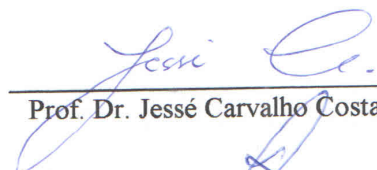
Data de Aprovação: 21.06.2004

Nota: 89% (EXCELENTE)

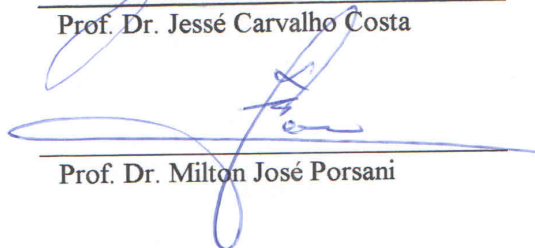
Comitê de Tese:



Prof. Dr. German Garabito Callapino



Prof. Dr. Jessé Carvalho Costa



Prof. Dr. Milton José Porsani

Belém – Pará
2004

À minha família

AGRADECIMENTOS

Em primeiro lugar a Deus, por estar sempre presente na minha vida, principalmente nas horas mais difíceis e até aqui tem me ajudado.

Ao Conselho nacional de Desenvolvimento Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro enquanto bolsista sem o qual não seria possível a realização deste trabalho.

Ao meu orientador Prof. Dr. German Garabito pela sugestão do tema de pesquisa e pelo apoio durante este período de trabalho.

À minha família pela compreensão e apoio.

A todo corpo de professores e funcionários do Curso e do Departamento de Geofísica.

Aos colegas do curso pelo companheirismo e esclarecimentos nos momentos de dificuldade.

Ao Curso de Pós-graduação em Geofísica (CPGf) coordenado pelo Prof. Dr. João Batista Correa da Silva pela infra-estrutura de laboratório e espaço físico de estudo.

À Benildes Lopes, secretária do CPGf, por fazer as coisas andarem e desatar os nós da burocracia e por sua paciência, apoio e incentivo durante todo esse período.

*“Dá instrução ao sábio, e ele se fará mais sábio,
ensina ao justo, e ele crescerá em entendimento.
O temor do Senhor é o princípio da sabedoria,
e a ciência do Santo, a prudência”.*

Pv. 9.9,10, Bíblia Sagrada.

SUMÁRIO

	p.
DEDICATÓRIA	i
AGRADECIMENTOS	ii
EPÍGRAFE	iii
LISTA DE ILUSTRAÇÕES	vi
RESUMO	1
ABSTRACT	3
1 INTRODUÇÃO	5
1.1 DADOS SÍSMICOS	5
1.2 EMPILHAMENTO SÍSMICO	7
1.3. OBJETIVOS	8
2 EMPILHAMENTO SÍSMICO POR SUPERFÍCIE DE REFLEXÃO COMUM ...	10
2.1 ONDA HIPOTÉTICAS E SEUS ATRIBUTOS	10
2.2 APROXIMAÇÃO HIPERBÓLICA DOS TEMPOS DE TRÂNSITO	12
2.3 EMPILHAMENTO SRC	16
3 OTIMIZAÇÃO	18
3.1 OTIMIZAÇÃO (ASPECTOS GERAIS)	18
3.2 OTIMIZAÇÃO DE FUNÇÕES MULTIMODAIS	19
3.3 OTIMIZAÇÃO GLOBAL DOS PARÂMETROS DO EMPILHAMENTO SRC	21
3.3.1 Função Objetivo do Método SRC	22
3.3.2 Exemplo da função objetivo SRC para três parâmetros (β_0, R_{PIN}, R_N)	22
3.3.3 Exemplo da função objetivo SRC para dois parâmetros (β_0, R_{PIN})	24
4 ALGORITMOS DE OTIMIZAÇÃO GLOBAL	26
4.1 ALGORITMO <i>SIMULATED ANNEALING</i>	26
4.2 ALGORITMO <i>VERY FAST SIMULATED ANNEALING</i>	30
4.3 ALGORITMO <i>DIFFERENTIAL EVOLUTION</i>	33

4.4 ALGORITMO <i>CONTROLLED RANDOM SEARCH – 2 (CRS2)</i>	35
4.5 DISCURSÃO DOS ALGORITMOS	38
5 EMPILHAMENTO SRC POR MEIO DE OTIMIZAÇÃO GLOBAL	39
5.1 O MODELO MARMOUSI	39
5.2. DADOS SÍSMICOS	40
5.3 ALGORITMO DE EMPILHAMENTO SRC	41
5.4 RESULTADOS DO EMPILHAMENTO SRC	42
6 CONCLUSÕES	45
REFERÊNCIAS	46
ANEXOS	
ANEXO A – ALGORITMO <i>SIMULATED ANNEALING</i>	51
ANEXO B – ALGORITMO <i>VERY FAST SIMULATED ANNEALING</i>	56
ANEXO C – ALGORITMO <i>DIFFERENTIAL EVOLUTION</i>	60
ANEXO D – ALGORITMO <i>CONTROLLED RANDOM SEARCH</i>	64

LISTA DE ILUSTRAÇÕES

	p
Figura 1.1 – Experimento de tiro comum. O asterisco representa a fonte e os triângulos os grupos de receptores. A seta indica o sentido da translação do conjunto fonte-receptores, durante o levantamento sísmico, sobre a linha de aquisição.	6
Figura 2.1 – Modelo composto por três camadas homogêneas separadas por interfaces curvadas. a) Onda Ponto de Incidência Normal b) Onda Normal.	11
Figura 2.2 – Representação esquemática do modelo sísmico com um refletor Σ , um raio central de incidência normal $X_0 R X_0$ e um raio paraxial de reflexão primária $S R' G$. A curva azul representa a frente de onda PIN e a curva vermelha representa a frente de onda N.	12
Figura 2.3 – Parte inferior: Modelo composto por três camadas homogêneas com um raio de afastamento nulo ou de incidência normal (linha vermelha). Parte superior: Superfície de cobertura múltipla (linha de cor azul) corresponde as reflexões da segunda interface. As linhas de cor vermelha definem a superfície de empilhamento SRC correspondente ao ponto de reflexão R e calculada pela expressão (2.1).	14
Figura 2.4 - Parte inferior: Modelo composto por três camadas homogêneas e um raio de afastamento nulo (linha de cor verde). <i>Parte superior</i> : Superfície de cobertura múltipla (linhas de cor azul) correspondente ao ponto de reflexão R calculado pela expressão (2.2)..	15
Figura 3.1 - Pontos de mínimos e máximos, globais e locais para uma função $f(\mathbf{p})$, com $\mathbf{p} \in P \subseteq R$	20
Figura 3.2 - Valores de coerência para o ponto P_0 . O valor mínimo de coerência está localizado em $b_0 = -3.03^\circ$, $R_{PIN} = 16245$ m e $K_N = 0.00054545^{-1}$ m.	23
Figura 3.3 - Mapa de coerência para valores de β_0 e R_{PIN} , associados a um ponto de amostragem da seção AN, de um modelo sintético, onde R_N é fixo.	25
Figura 4.1a - Desempenho do algoritmo SA após várias avaliações da função para a busca de três parâmetros.	29

Figura 4.1b - Desempenho do algoritmo SA após várias avaliações da função para busca de dois parâmetros.	29
Figura 4.2a - Desempenho do algoritmo VFSA após várias avaliações da função para a busca de três parâmetros.	32
Figura 4.2b - Desempenho do algoritmo VFSA após várias avaliações da função para busca de dois parâmetros.	32
Figura 4.3a – Desempenho do algoritmo DE após várias avaliações da função na busca de três parâmetros.	34
Figura 4.3b - Desempenho do algoritmo DE após várias avaliações da função na busca de dois parâmetros.	34
Figura 4.4a - Desempenho do algoritmo CRS2 após várias avaliações da função na busca de três parâmetros.	37
Figura 4.4b - Desempenho do algoritmo CRS2 após várias avaliações da função na busca de três parâmetros.	37
Figura 5.1 – Modelo de velocidade dos dados Marmousi.	41
Figura 5.2 - Seção com afastamento nulo, simulada usando o operador de empilhamento, definida por dois parâmetros (β_0 , R_{NIP}).	43
Figura 5.3. Seção com afastamento nulo, simulada usando o operador de empilhamento, definida por três parâmetros (β_0 , R_{NIP} , R_N).	43
Figura 2A. Determinação da temperatura inicial do algoritmo SA para a busca dos parâmetros: a) ângulo de emergência (β_0) e b) raios de curvatura da onda PIN (R_{PIN}) e da onda N (R_N).	54
Figura 1C. Exemplo em duas dimensões de uma função objetivo mostrando suas linhas de contorno e o processo para geração de $v_i^{(G+1)}$ no esquema ED-1.	60
Figura 2C. Ilustração do cruzamento do processo para $n_{param} = 7$, $k = 2$ e $L = 3$	61
Figura 1D. Fluxograma para o algoritmo CRS2.	65
Figura 2D. Uma ilustração do algoritmo gerando pontos testes em 2D.	66

RESUMO

O método de empilhamento por Superfície de Reflexão Comum (SRC) produz seções simuladas de afastamento nulo (AN) por meio do somatório de eventos sísmicos dos dados de cobertura múltipla contidos nas superfícies de empilhamento. Este método não depende do modelo de velocidade do meio, apenas requer o conhecimento a priori da velocidade próxima a superfície. A simulação de seções AN por este método de empilhamento utiliza uma aproximação hiperbólica de segunda ordem do tempo de trânsito de raios paraxiais para definir a superfície de empilhamento ou operador de empilhamento SRC. Para meios 2D este operador depende de três atributos cinemáticos de duas ondas hipotéticas (ondas PIN e N), observados no ponto de emergência do raio central com incidência normal, que são: o ângulo de emergência do raio central com fonte-receptor nulo (β_0), o raio de curvatura da onda ponto de incidência normal (R_{PIN}) e o raio de curvatura da onda normal (R_N).

Portanto, o problema de otimização no método SRC consiste na determinação, a partir dos dados sísmicos, dos três parâmetros (β_0 , R_{PIN} , R_N) ótimos associados a cada ponto de amostragem da seção AN a ser simulada. A determinação simultânea destes parâmetros pode ser realizada por meio de processos de busca global (ou otimização global) multidimensional, utilizando como função objetivo algum critério de coerência. O problema de otimização no método SRC é muito importante para o bom desempenho no que diz respeito a qualidade dos resultados e principalmente ao custo computacional, comparado com os métodos tradicionalmente utilizados na indústria sísmica.

Existem várias estratégias de busca para determinar estes parâmetros baseados em buscas sistemáticas e usando algoritmos de otimização, podendo estimar apenas um parâmetro de cada vez, ou dois ou os três parâmetros simultaneamente. Levando em conta a estratégia de busca por meio da aplicação de otimização global, estes três parâmetros podem ser estimados através de dois procedimentos: no primeiro caso os três parâmetros podem ser estimados simultaneamente e no segundo caso inicialmente podem ser determinados simultaneamente dois parâmetros (β_0 , R_{PIN}) e posteriormente o terceiro parâmetro (R_N) usando os valores dos dois parâmetros já conhecidos. Neste trabalho apresenta-se a aplicação e comparação de quatro algoritmos de

otimização global para encontrar os parâmetros SRC ótimos, estes são: *Simulated Annealing* (SA), *Very Fast Simulated Annealing* (VFSA), *Differential Evolution* (DE) e *Controlled Random Search - 2* (CRS2).

Como resultados importantes são apresentados a aplicação de cada método de otimização e a comparação entre os métodos quanto a eficácia, eficiência e confiabilidade para determinar os melhores parâmetros SRC. Posteriormente, aplicando as estratégias de busca global para a determinação destes parâmetros, por meio do método de otimização VFSA que teve o melhor desempenho foi realizado o empilhamento SRC a partir dos dados Marmousi, isto é, foi realizado um empilhamento SRC usando dois parâmetros (β_0, R_{PIN}) estimados por busca global e outro empilhamento SRC usando os três parâmetros (β_0, R_{PIN}, R_N) também estimados por busca global.

ABSTRACT

The Common Reflection Surface Stack (CRS) method provides the simulation of zero-offset (ZO) sections by means of the summing seismic events of the multicoverage data contained in the stacking surface. This method does not depend on the velocity macro-model of medium; it only requires a priori knowledge of the near-surface velocity. The simulation of ZO sections for this stacking method uses a hyperbolic second-order travelttime approximation of the paraxial rays to define the CRS stacking surface or CRS stack operator. For 2D media, this operator depends on three kinematic attributes of two hypothetical waves (*NIP* and *N* waves), observed in the point of emergency of the central ray with normal incidence, namely, the angle of emergency of the central ZO ray (β_0), the radius of curvature of the Normal Incidence Point Wave (R_{NIP}) and the radius of curvature of the Normal Wave (R_N). Therefore, the optimization problem in the CRS method consists in the determination, from the seismic data, of the three optimal parameters (β_0, R_{NIP}, R_N) associated to each sample point of ZO section to be simulated. The simultaneous determination of these parameters can be made by means of multidimensional global search process (or global optimization), using as objective function some coherence criterion. The optimization problem in CRS method is very important for the good performance with respect to quality of the results and mainly to computational cost, compared with the methods traditionally used in the seismic industry. There are several search strategies to determine these parameters, based on systematic searches and using optimization algorithms, where only one parameter at each time can be estimated, or the two or three parameters simultaneously. Taking in to account the search strategy by means of the application of global optimization, these three parameters can be estimated through of procedures: in the first case the three parameters can be simultaneously estimated and in second case initially two parameters can be determined simultaneously (β_0, R_{NIP}), and subsequently the third parameter (R_N), using the values of the two parameters already known. In this work it is presented the application and comparison of four algorithms of global optimization to find the CRS optimal parameters: *Simulated Annealing* (SA), *Very Fast Simulated Annealing* (VFSA), *Differential Evolution* (DE) and *Controlled Random Search – 2* (CRS2). As importants results of the application of each optimization method, as well as between the methods regarding the effectiveness, efficiency and

reliability to determine the best CRS parameters are presented. Subsequently, applying the global search strategies for the determination of these parameters, by means of the optimization method VFSA that presented the best performance, the CRS stacking was applied to the Marmousi dataset, one stacking using two parameters (β_0, R_{NIP}), estimated by global search, and another CRS stacking using the three parameters (β_0, R_{NIP}, R_N), also estimated by global search.

1 – INTRODUÇÃO

Neste capítulo introdutório definiremos os dados sísmicos de reflexão e o procedimento para aquisição desses dados, e ainda falaremos sobre o empilhamento sísmico e os objetivos deste trabalho de dissertação.

1.1 – DADOS SÍSMICOS

Um dado sísmico é o registro da propagação de ondas sísmicas na subsuperfície da Terra. Na sísmica de prospecção, as ondas elásticas utilizadas são geradas artificialmente por meio de fontes como Dinamite ou outro explosivo. À aquisição dos dados sísmicos dá-se o nome de levantamento sísmico. Os levantamentos sísmicos podem ser feitos em terra ou em mar. Problemas relacionados com a proteção do meio ambiente e segurança levaram ao desenvolvimento de fontes sísmicas de baixa energia, dentre as quais podem-se mencionar o Air Gun, Sparker, Aquapulse, Vaporchoc, Flexotir, Dinoseis, além de sistemas vibratórios (Vibroseis) usados largamente pela indústria. Uma parte da energia dessas ondas é absorvida pelo meio e outra parte é transmitida e refletida nas estruturas geológicas da terra retornando á superfície, sendo registrada por vários receptores dispostos sistematicamente com relação a fonte. Cada experimento de campo desse tipo constituído de uma fonte e vários receptores, é denominado experimento de fonte ou tiro comum (FC), pois todos os receptores registram a resposta do meio à uma mesma fonte.

Esse registro é discreto e feito a intervalos regulares de tempo, determinados pela razão amostral ou intervalo de amostragem. É comum que esse intervalo seja de 2ms, sendo que intervalos ainda menores já são possíveis, apesar de ainda não rotineiros. Com base no tempo decorrido entre a emissão do pulso sísmico e o registro de uma determinada reflexão, pode-se determinar a profundidade do refletor, desde que se conheça a velocidade média do meio. Os receptores representam, na realidade, grupos de transdutores (geofones) e estes grupos podem ser constituídos por mais de 100 geofones individuais dispostos na superfície, segundo vários padrões chamados de arranjos onde os mais comuns são: os arranjos lineares, isto é, com os

geofones dispostos em linha reta (levantamento 2D) e os arranjos em área (levantamento 3D), isto é, com os geofones dispostos em forma de malha.

A Figura 1.1 ilustra o levantamento sísmico 2D, onde uma fonte gera um onda que será registrada pelos receptores. Após cada experimento o conjunto fonte-receptores é transladado sobre uma curva, denominada linha de aquisição sísmica, e então o experimento é realizado novamente nesta nova posição.

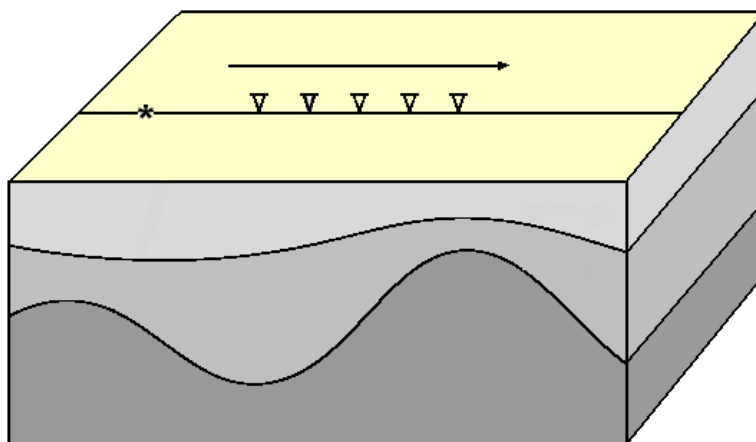


Figura 1.1. Experimento de tiro comum. O asterisco representa a fonte e os triângulos os grupos de receptores. A seta indica o sentido da translação do conjunto fonte-receptores, durante o levantamento sísmico, sobre a linha de aquisição.

A distância entre os grupos de geofones pode variar entre 15 e 300 metros. Cada vez que um pulso sísmico é produzido pela fonte, registram-se simultaneamente, um certo número de canais ou traços sísmicos, sendo que cada canal corresponde a um grupo de geofones. Durante o registro os sinais sísmicos captados pelos geofones são digitalizados e gravados em uma fita magnética. Os geofones que participam de cada registro cobrem uma extensão superficial da ordem de 1 Km a 2 Km, e uma única linha pode ultrapassar 100 Km de comprimento.

A coleção de vários traços sísmicos recebe o nome de seção sísmica. Finalmente, a coleção de várias seções sísmicas forma os dados sísmicos de cobertura múltipla.

1.2 – EMPILHAMENTO SÍSMICO

A sísmica de reflexão tem como principal objetivo produzir imagens das regiões de interesse em subsuperfícies, mapeando estruturas geológicas e extraíndo parâmetros que caracterizam sua litologia e propriedades petrofísicas. Portanto, o processamento sísmico de reflexão é uma ferramenta essencial às atividades de exploração e monitoramento de reservatórios de petróleo.

Então, com o intuito de obter imagens das regiões em subsuperfície, os dados são submetidos a uma série de processos os quais levam à obtenção de uma seção sísmica final que será interpretada. Uma das etapas aplicadas com esse objetivo simula uma seção de afastamento nulo (AN), ou seja, uma seção composta de traços sísmicos com afastamento fonte-receptor nulo.

No processamento sísmico convencional utilizado na indústria, a simulação de uma seção AN é realizada por meio da técnica CMP cuja seqüência de processamento compreende as seguintes etapas principais: a análise de velocidade, silenciamento ou muting, a correção normal moveout (NMO), correção dip moveout (DMO) e o empilhamento horizontal de traços. O processo de empilhamento CMP, que depende do modelo de velocidade, não produz resultados satisfatórios quando é aplicado em meios complexos, como por exemplo, em meios com fortes variações laterais de velocidades.

Em contraste ao empilhamento CMP, existem outros métodos de imageamento que não dependem de modelo de velocidade e que podem ser aplicados em dados de meios heterogêneos produzindo resultados satisfatórios. Podemos citar os métodos: Imageamento Homeomórfico Multifoco (Gelchinsky et al., 1997), Elemento de Reflexão Comum (ERC), Elemento de Evoluta Comum (EEC), Ponto de Tiro Comum (PTC) (Gelchinsky & Keydar 1999; Gelchinsky et al., 1999a,b) e Empilhamento por Superfície de Reflexão Comum (SRC) (Müller et al., 1998; Jäger et al., 2001; Garabito et al., 2001a; Garabito et al., 2003). Essas alternativas podem fornecer melhorias em relação ao empilhamento convencional e fornecem também atributos de frentes de ondas hipotéticas que podem ser usados para a determinação do modelo de velocidade, Zona de Fresnel, etc.

Para a simulação de seções AN através do método de Empilhamento por Superfícies de Reflexão Comum (SRC), a fórmula de aproximação hiperbólica dos tempos de transito é usada para calcular as superfícies ou operadores de empilhamento. Para meios 2D o operador de

empilhamento SRC depende de três parâmetros que são: o ângulo de emergência do raio central com fonte-receptor nulo (β_0), o raio de curvatura da onda ponto de incidência normal (R_{PIN}) e o raio de curvatura da onda normal (R_N). Estes três parâmetros, que estão associados a cada ponto de amostragem da seção AN, são determinados, a partir dos dados sísmicos de cobertura múltipla, por meio de processos de otimização e análise de coerência utilizando estratégias que envolvem processos de busca de um, dois ou três parâmetros. Portanto, o problema de otimização no método SRC consiste na busca do trio de parâmetros (β_0, R_{PIN}, R_N) , a partir dos dados sísmicos, que definem a superfície de empilhamento que melhor se ajusta aos eventos sísmicos.

A importância da otimização no método SRC esta ligada diretamente à precisão dos resultados a serem obtidos e principalmente ao maior custo computacional comparado com os métodos tradicionais usados na indústria sísmica. Por isso existe a necessidade de pesquisar o desempenho dos métodos de otimização global tradicionais e outros, no processo de determinação dos parâmetros do método SRC.

1.3. OBJETIVOS

O objetivo deste trabalho consiste do estudo, aplicação e comparação de algoritmos de otimização global aplicados no problema de otimização do método de empilhamento SRC. Tomando como base a estratégia de busca dos parâmetros SRC proposta em Garabito et al., (2001a,b) que consiste na busca global de dois parâmetros e, posteriormente, de um parâmetro, por meio da aplicação do algoritmo Simulated Annealing (SA) (Corana et al., 1987). Neste trabalho apresenta-se um estudo comparativo de três algoritmos de otimização global aplicados na busca dos parâmetros SRC, que são: Very Fast Simulated Annealing (VFSA) (Ingber, 1989, 1993, 1996), Differential Evolution (DE) (Storn & Price 1996) e Controlled Random Search - 2 (CRS2) (Price, 1983).

Nesta oportunidade apresentaremos os resultados da comparação destes quatro métodos quanto a eficiência, eficácia e confiabilidade para a determinar os melhores parâmetros SRC. Uma vez encontrado o algoritmo com melhor desempenho na busca dos parâmetros, o mesmo será implementado para realizar o empilhamento SRC por meio da otimização de dois parâmetros e dos três parâmetros simultaneamente.

ESTRUTURA DA TESE:

Além deste capítulo introdutório, o corpo desta tese é composto pelos seguintes capítulos:

No capítulo 2 é feita uma breve apresentação do método SRC, dos atributos de frentes de ondas e também são apresentadas as principais fórmulas de aproximação dos tempos de transito usados para o cálculo das superfícies de empilhamento.

No capítulo 3 descreveremos alguns aspectos gerais sobre otimização, bem como a otimização dos parâmetros SRC e a função objetivo usada pelo empilhamento SRC.

No capítulo 4 apresentamos a aplicação na busca dos parâmetros SRC dos quatro algoritmos de otimização global Simulated Annealing (SA), Very Fast Simulated Annealing (VFSA), Differential Evolution (DE) e Controlled Random Search - 2 (CRS2) e faremos a comparação entre os métodos quanto à eficiência, eficácia e confiabilidade para determinação do trio de parâmetro SRC.

No capítulo 5 apresentaremos os resultados da aplicação do empilhamento SRC obtidos, nos dados marmusi, usando a determinação de dois e três parâmetros por meio de otimização global

No capítulo 6 são dadas as conclusões.

Nos anexos A, B, C e D são dados mais detalhes sobre os algoritmos de otimização global utilizados aqui.

2 – EMPILHAMENTO SÍSMICO POR SUPERFÍCIE DE REFLEXÃO COMUM

Neste capítulo é apresentada a descrição dos três atributos de frentes de ondas correspondentes a onda Ponto de Incidência Normal (PIN) e a onda Normal (N), bem como as principais fórmulas hiperbólicas de aproximações dos tempos de trânsito usadas para o cálculo das curvas e das superfícies de empilhamento e é apresentada também a técnica de empilhamento SRC.

2.1. ONDAS HIPOTÉTICAS E SEUS ATRIBUTOS

As trajetórias ou superfícies de empilhamento usadas no método de empilhamento SRC são definidas com base em atributos cinemáticos das duas frentes de ondas hipotéticas: a onda PIN e a onda N, como definidas por Hubral (1983). A Figura 2.1 ilustra a propagação de duas ondas hipotéticas num meio heterogêneo, composto por três camadas homogêneas, associados a um raio com afastamento fonte-receptor nulo, desde o ponto de incidência normal R sobre o refletor até o ponto de observação x_0 na linha sísmica. A onda Ponto de Incidência Normal (PIN) é uma onda hipotética que se propaga de forma ascendente a partir de uma fonte localizada no ponto R . A Figura 2.1a mostra a propagação desta onda em diferentes instantes de tempo até atingir o ponto de emergência x_0 . A onda normal (N) é uma onda hipotética originada a partir de um experimento do refletor explodindo em profundidade e se propaga de forma ascendente, sendo que a curvatura da frente de onda no ponto R é igual à curvatura local do refletor em R . A Figura 2.1b mostra a propagação desta onda em diferentes instantes de tempo associada ao raio normal.

Observa-se que a propagação destas ondas hipotéticas está associada com a trajetória do raio com incidência normal, sendo que as curvaturas finais das frentes de ondas emergentes, tanto da onda PIN como da onda N, são medidas no ponto x_0 , e o ângulo de emergência comum (β_0) das duas ondas hipotéticas coincide com o ângulo de emergência do raio com incidência normal em x_0 . Os três atributos cinemáticos das ondas hipotéticas PIN e N usados no empilhamento SRC são: o ângulo de emergência do raio com afastamento nulo β_0 , a curvatura da onda PIN K_{NIP}

(ou raio de curvatura $R_{NIP} = 1/K_{NIP}$) e a curvatura da onda N K_N (ou raio de curvatura $R_N = 1/K_N$), ambos medidos no ponto x_0 .

Em um meio heterogêneo, como ilustrado pela Figura 2.1, estes três atributos cinemáticos (β_0 , R_{NIP} e R_N) estão associados com a orientação, distância e curvatura do refletor respectivamente. Ao longo da trajetória do raio normal R até o ponto de interseção (T) do raio com a interface curva, a onda hipotética PIN se propagam com frente de onda circular, mas logo após a transmissão através da interface curva, a frente de onda em geral não é mais circular. Em um meio heterogêneo, as curvaturas K_{NIP} e K_N medidas no ponto de emergência x_0 , são aproximações circulares das frente de ondas PIN e N, respectivamente.

Como mencionado anteriormente, os três atributos cinemáticos (β_0 , R_{NIP} e R_N) definem as curvas ou superfícies de empilhamento do método SRC, usando para isto a fórmula de aproximação hiperbólica dos tempos de trânsito a ser descrita adiante. Portanto, para um determinado ponto de observação (x_0) em um dado tempo (t_0) do raio com afastamento nulo, estes três atributos de frentes de ondas podem ser determinados a partir dos dados sísmicos por meio de medidas de coerência do sinal sísmico.

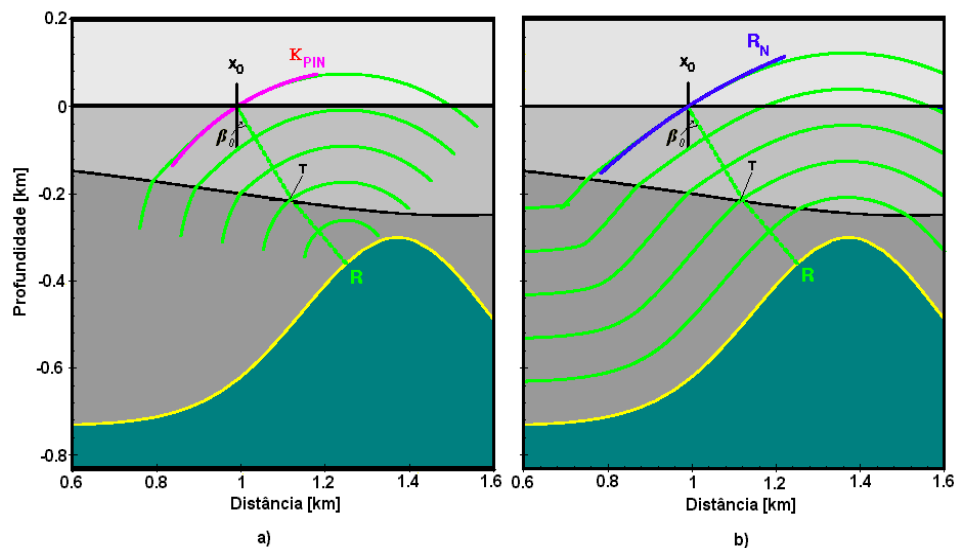


Figura 2.1. Modelo composto por três camadas homogêneas separadas por interfaces curvadas. a) Onda Ponto de Incidência Normal b) Onda Normal

2.2 – APROXIMAÇÃO HIPERBÓLICA DOS TEMPOS DE TRÂNSITO

A expansão hiperbólica de segunda ordem em série de Taylor que aproxima os tempos de trânsito dos raios na vizinhança de um raio central fixo é obtida por meio da teoria paraxial do raio (Schleicher et al., 1993). Para meios 2D, Tygel et al. (1997) obtém essa aproximação em função dos atributos cinemáticos das ondas hipotéticas PIN e N. Como mostrado na Figura 2.2, considera-se um refletor curvo Σ em profundidade e um raio central fixo com afastamento fonte-receptor nulo $X_0 R X_0$, onde R é o ponto de incidência normal. Para um par fonte-receptor (S, G) na vizinhança do ponto de emergência X_0 do raio central, considera-se um raio de reflexão primária $S R' G$ com relação ao mesmo refletor Σ .

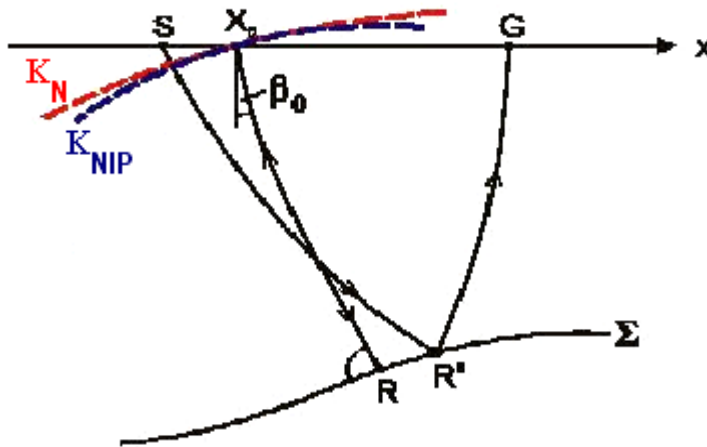


Figura 2.2. Representação esquemática do modelo sísmico com um refletor Σ , um raio central de incidência normal $X_0 R X_0$ e um raio paraxial de reflexão primária $S R' G$. A curva azul representa a frente de onda PIN e a curva vermelha representa a frente de onda N.

A expansão da aproximação hiperbólica dos tempos de trânsito dos raios de reflexões primárias na vizinhança de um raio central normal, aplicável para configurações irregulares de fonte e receptores, é dada por

$$t^2(x_m, h) = \left(t_0 + \frac{2 \operatorname{sen} \beta_0}{v_0} (x_m - x_0)^2 \right)^2 + \frac{2 t_0 \cos^2 \beta_0}{v_0} \left((x_m - x_0)^2 K_N + h^2 K_{PIN} \right) \quad (2.1)$$

onde t_0 é o tempo duplo do raio central de afastamento nulo e v_0 a velocidade próxima à superfície entorno do ponto de emergência do raio central. As coordenadas $x_m = (x_G + x_S)/2$ e $h = (x_G - x_S)/2$ são o ponto médio e o meio-afastamento entre a fonte e o receptor, respectivamente, sendo x_S e x_G as coordenadas horizontais do par fonte-receptor (S, G) respectivamente. A coordenada $X_0 = (x_0, h = 0)$ sobre a linha sísmica é o ponto de emergência do raio central com fonte e receptor coincidentes. Como explicado e ilustrado na Figura 2.2, β_0 é o ângulo de emergência do raio normal com relação a normal da superfície, K_{NIP} e K_N são, respectivamente, as curvaturas das ondas PIN e da onda normal N, ambas medidas no ponto de emergência do raio central. Devido ao fato de a equação (2.1) esta expressa em função de v_0 que é conhecida a priori, esta aproximação hiperbólica dos tempos de trânsito é independente do modelo de velocidades do meio e portanto aplicável a meios heterogêneos.

Denota-se por $P_0 = (t_0, x_0)$ a coordenada de um ponto de amostragem (ou ponto imagem) na seção AN a ser simulada. Para valores conhecidos dos três atributos das onda hipotéticas $(\beta_0, R_{NIP}$ e $R_N)$ associados ao ponto P_0 , pode-se calcular a superfície de empilhamento associada a esse ponto usando a expressão (2.1).

Para ilustração de uma superfície de empilhamento definida por essa expressão, considera-se o modelo sintético composto por três camadas homogêneas como apresentadas na parte inferior da Figura 2.3. Na parte superior desta figura, no domínio (x_m, h, t) , mostra-se as curvas dos tempos de trânsito das reflexões primárias (curvas de cor azul) correspondente ao segundo refletor, medidas ao longo da linha sísmica com geometria de afastamento fonte-receptor comum. As linhas curvas de cor vermelha formam a superfície de empilhamento do ponto de amostragem P_0 , corresponde a uma reflexão primaria no ponto R localizado sobre a segunda interface. Neste caso, o trio de atributos das ondas PIN e N associados ao raio normal $X_0 R X_0$, foram calculados por modelamento direto. Pode-se então afirmar que a superfície de empilhamento SRC, também denominada de operador de empilhamento SRC, é uma aproximação da resposta cinemática de uma reflexão sobre uma interface num meio heterogêneo.

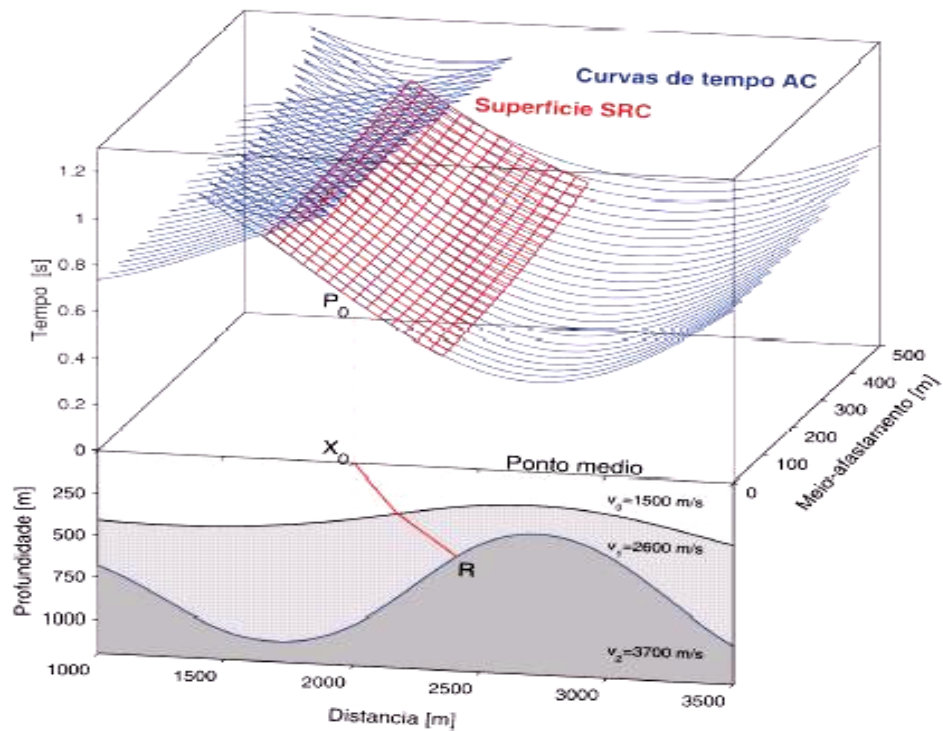


Figura 2.3. *Parte inferior:* Modelo composto por três camadas homogêneas com um raio de afastamento nulo ou de incidência normal (linha vermelha). *Parte superior:* Superfície de cobertura múltipla (linha de cor azul) corresponde as reflexões da segunda interface. As linhas de cor vermelha definem a superfície de empilhamento SRC correspondente ao ponto de reflexão R e calculada pela expressão (2.1)

Considerando um meio homogêneo com um refletor, os três atributos cinemáticos das duas ondas hipotéticas fornecem informações sobre a orientação angular do refletor no ponto de incidência normal, a posição do ponto de reflexão na subsuperfície e a curvatura local do refletor no ponto de incidência do raio normal. Seguindo essa interpretação, se considerarmos o caso particular em que $K_{PIN} = K_N$ significa que o raio de curvatura da onda hipotética N não fornece mais a informação da curvatura do refletor no ponto de incidência normal, isto significa que os atributos cinemáticos não estão vinculados de forma unívoca a um ponto de reflexão em subsuperfície. Aplicando esta condição em (2.1), obtemos a expressão

$$t^2(x_m, h) = \left(t_0 + \frac{2 \operatorname{sen} \beta_0}{v_0} (x_m - x_0) \right)^2 + \frac{2 t_0 \cos^2 \beta_0}{v_0} K_{PIN} \left((x_m - x_0)^2 + h^2 \right), \quad (2.2)$$

resultando assim uma expressão dependente de dois atributos cinemáticos (β_0 e R_{NIP}) associados a onda hipotética PIN. Devido ao fato de que β_0 e R_{NIP} fornecerem informações sobre a posição de um ponto na subsuperfície, os tempos de trânsito calculados por (2.1) podem se interpretados como aproximações dos tempos de trânsito de uma difração, em outras palavras, a expressão (2.2) é uma aproximação na vizinhança do raio central do operador de migração pré-empilhamento. Por outro lado, como mostrado na parte superior da Figura 2.4, mesmo dependendo apenas de dois atributos cinemáticos, a superfície de empilhamento (curvas de cor verde) definida pela Expressão (2.2) também é uma aproximação dos tempos de trânsito associados à reflexão em R. Portanto, de forma similar ao método de empilhamento SRC que simula uma seção AN com base na formula (2.1), também pode ser simulada uma seção AN usando a expressão (2.2), que define as superfícies de empilhamento associadas tanto a pontos de reflexões como a difratores em subsuperfície.

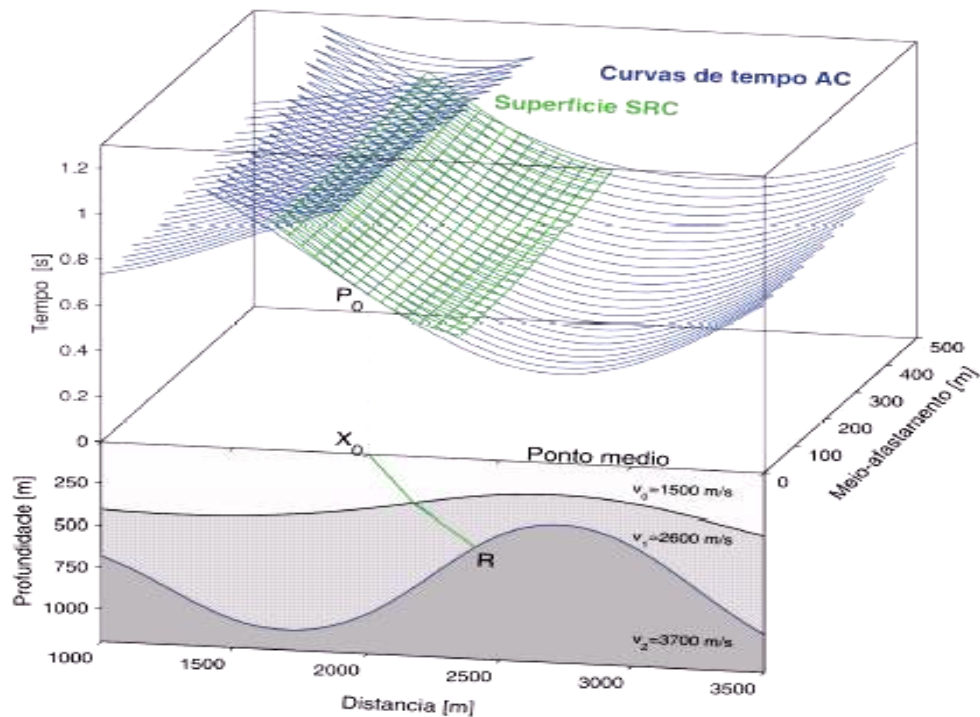


Figura 2.4: *Parte inferior:* Modelo composto por três camadas homogêneas e um raio de afastamento nulo (linha de cor verde). *Parte superior:* Superfície de cobertura múltipla (linhas de cor azul) correspondente ao ponto de reflexão R calculado pela expressão (2.2)

Vale mencionar que a expressão (2.1) pode ser particularizada para as diferentes configurações de fontes e receptores, usadas com maior frequência no processamento sísmico convencional, tais como o Ponto Médio Comum (PMC), Fonte Comum (FC), Receptor Comum (RC), Afastamento fonte-receptor nulo (AN). A expressão dos tempos de trânsito para a configuração AC coincide com a fórmula geral (2.1), mas sob a consideração de $h = constante$. Neste trabalho utilizaremos as fórmulas gerais (2.1) e (2.2) que podem ser usadas para configurações irregulares de fontes e receptores.

2.3 – EMPILHAMENTO SRC:

O empilhamento por Superfície de Reflexão comum (SRC) (Müller et al., 1998, Mann et al., 1999) é um processo de empilhamento sísmico multiparamétrico que não depende de um macro modelo de velocidade.

O método de empilhamento SRC simula uma seção AN por meio do somatório de eventos sísmicos dos dados de cobertura múltipla contido nas superfícies de empilhamento associados a cada um dos pontos imagem da seção AN a ser simulada. Neste método, para calcular uma determinada superfície de empilhamento, ou operador de empilhamento SRC, utiliza-se a fórmula de aproximação hiperbólica dos tempos de trânsito (2.1). Assim sendo, a definição do melhor operador SRC associado a um certo ponto de amostragem (P_0) da seção AN, requer a determinação a partir dos dados de cobertura múltipla dos três atributos cinemáticos (β_0 , R_{NIP} e R_N) que produzam um operador SRC com melhor ajuste possível aos eventos sísmicos associados ao ponto P_0 . Portanto, o problema principal da simulação de uma seção AN pelo método de empilhamento SRC consiste na determinação dos três atributos cinemáticos, para cada um dos pontos de amostragem da seção AN a ser simulada, envolvendo medidas de coerência dos dados sísmicos.

Estes três atributos, também denominados como trio de parâmetros do operador de empilhamento SRC, podem ser determinados simultaneamente por meio de processos de otimização global multidimensional, usando como função objetivo uma medida de coerência do sinal sísmico (por exemplo, semblance) realizada ao longo de vários operadores SRC testados. Assim, a determinação do trio de parâmetros independentes (β_0 , R_{NIP} e R_N) pode ser formulada como um problema de maximização (ou minimização) da medida de coerência *semblance*

(Neidell & Taner, 1997), avaliada para pontos dentro de um espaço de busca tridimensional definido pelos intervalos: $-\pi/2 < \beta_0 < +\pi/2$ e $-\infty < R_{NIP}, R_N < +\infty$.

Uma vez encontrado o trio de parâmetros ótimos associado a um determinado ponto de amostragem da seção AN, é realizado o somatório dos eventos sísmicos contidos no operador SRC definido pelo trio de parâmetros que produzem o maior valor de coerência. Portanto, a seção AN simulada é obtida pela repetição deste procedimento para cada um dos pontos de amostragem desta seção.

Em geral, o processo de otimização global multidimensional é um processo que consome muito tempo computacional e como regra, tem-se que, quanto maior o número de parâmetros maior o tempo computacional. Na otimização do trio de parâmetros SRC este problema torna-se ainda mais crítico devido ao fato de que a função objetivo é uma medida de coerência que envolve grande volume de dados sísmicos. Devido a este fato, surge a necessidade de determinarmos o trio de parâmetros SRC por meio de estratégias de busca de um parâmetro e dois parâmetros envolvendo menores volumes de dados sísmicos.

Atualmente, na literatura existem dois algoritmos de empilhamento SRC que usam diferentes estratégias de busca do trio de parâmetros. O primeiro denominado “Extended Pragmatic Search Strategy” proposto por Müller, 1999; Jäger et al., 2001; Mann et al. 2002, realiza a busca sistemática de um parâmetro por vez e somente na etapa final utiliza o algoritmo de otimização simplex (Nelder & Mead, 1965) para determinar simultaneamente o trio de parâmetros finais, tomando como aproximações iniciais os valores dos parâmetros resultantes das buscas unidimensionais. A segunda é denominada “Global Search Strategy” proposta por Garabito et al., 2001a, onde é usado um algoritmo global para a busca inicial de dois parâmetros e posteriormente de um parâmetro. Similar ao anterior, também na etapa final é usado um algoritmo de otimização local para determinar o trio de parâmetros finais, usando como aproximações iniciais os valores dos parâmetros resultantes das buscas globais.

Neste trabalho com base na estratégia proposta por Garabito et al., 2001a, realizamos um estudo comparativo de quatro métodos de otimização global (SA, VFSA, DE e CRS2) aplicados na busca simultânea de dois parâmetros (β_0, R_{PIN}) e de três parâmetros (β_0, R_{PIN}, R_N).

3 – OTIMIZAÇÃO

Neste capítulo apresentaremos alguns aspectos gerais sobre otimização, a otimização dos parâmetros SRC e a função objetivo usada pelo empilhamento SRC.

3.1. OTIMIZAÇÃO (ASPECTOS GERAIS):

No sentido clássico, otimização pode ser definida como o processo de encontrar a melhor solução (ou solução ótima) de um conjunto de soluções para um problema. O desenvolvimento de uma variedade de técnicas matemáticas especializada e o uso de computadores de alta velocidade permitem a muitos problemas complexos, em diversas áreas da ciência, serem otimizados. Estas técnicas (sua fundação lógica, suas características e seus métodos de aplicação) compreendem o domínio da teoria de otimização.

As técnicas de otimização devem ser utilizadas quando não existe uma solução simples e diretamente calculável para um problema. Isso geralmente ocorre quando a estrutura do problema é complexa, ou existem várias soluções possíveis. Neste caso, é possível que não exista nenhum procedimento direto de solução, de forma que as técnicas de otimização podem ser utilizadas na busca pela melhor solução do problema.

Geralmente a função a ser otimizada é da forma $f : R^n \rightarrow Y \subseteq R$ onde Y é chamado espaço de critério (Delaurenti, M., 1999). Dada uma função f , dependendo de várias variáveis, o problema de otimização consiste em minimizarmos (ou maximizarmos) a função $f(\mathbf{p})$ otimizando os valores de seus parâmetros $\mathbf{p} = (p_1, \dots, p_n)$, com $p_i \in R$ e $i \in \{1, \dots, n\}$, em que \mathbf{p} denota o vetor composto por n parâmetros da função, ou seja, consiste em:

$$\min_{\mathbf{p} \in R^n} f(\mathbf{p}).$$

Esta formulação também é conhecida como problema de otimização inrestrita em que não há restrições dos valores dos parâmetros \mathbf{p} . Um problema de otimização restrita consiste em minimizar (ou maximizar) $f(\mathbf{p})$ com $\mathbf{p} \in R^n$, sujeito a $g_k(\mathbf{p}) \leq h_k$, $k = 1, \dots, q$, onde as q equações $g_k(\mathbf{p}) \leq h_k$ constituem o conjunto das restrições da otimização. A função f é também

chamada de função objetivo, ou função custo. Maximizar $f(\mathbf{p})$ em uma região qualquer do espaço de dimensão n é equivalente a minimizar $-f(\mathbf{p})$ na mesma região. Estes problemas são problemas de otimização clássica.

A otimização pode ser dividida em 2 classes: global e local. A otimização global encontra a melhor solução do conjunto de "todas" as soluções possíveis, ou seja, dada uma função contínua $f(\mathbf{p})$, encontrar \mathbf{p}^* tal que $f(\mathbf{p}^*) \leq f(\mathbf{p})$ para todo \mathbf{p} pertencente ao espaço de dimensão n . A otimização local encontra a melhor solução dentro de um conjunto de soluções que está próximo ao ponto inicial, isto é, dada uma função contínua $f(\mathbf{p})$ tenta encontrar um vetor solução \mathbf{p}^* tal que $f(\mathbf{p}^*) \leq f(\mathbf{p})$ para todo \mathbf{p} próximo a \mathbf{p}^* , pois na otimização local, a solução encontrada depende do ponto de início do processo de busca. A otimização global sempre encontrará a melhor solução possível, independentemente das condições de início do processo de busca, porém, geralmente, requisita um maior esforço computacional. Nesta dissertação trabalharemos apenas com a otimização global de funções de extremos múltiplos.

3.2. OTIMIZAÇÃO DE FUNÇÕES MULTIMODAIS

Problemas de otimização global em espaços contínuos estão presentes em diversas áreas da ciência. Em geral, a tarefa é otimizar certas propriedades de um sistema por uma escolha pertinente de seus parâmetros. A aproximação padrão para um problema de otimização começa designando uma função objetivo que possa modelar os objetivos do problema enquanto é incorporado algumas restrições (otimização restrita).

Em muitos casos a função objetivo está designada a transformar o problema de otimização em uma tarefa de minimização, isto é, o objetivo da otimização seria encontrar os valores ótimos que minimizem da função objetivo.

Em algumas aplicações, a função objetivo pode ser bastante complicada e caracterizada pela presença de vários máximos e mínimos. Nesse tipo de função objetivo, chamada multimodal, o mínimo de todos os mínimos é chamado de mínimo global e todos os outros são chamados mínimos locais. A Figura 3.1 ilustra os pontos de mínimos e máximo, globais e locais, para uma função contínua $f : P \subseteq R \rightarrow R$ com $p \in P$, onde os pontos p_1, p_3 são pontos de máximos locais e os pontos p_4, p_6 são pontos de mínimos locais, por outro lado os pontos p_2 e p_5 são os pontos de mínimo global e máximo global respectivamente.

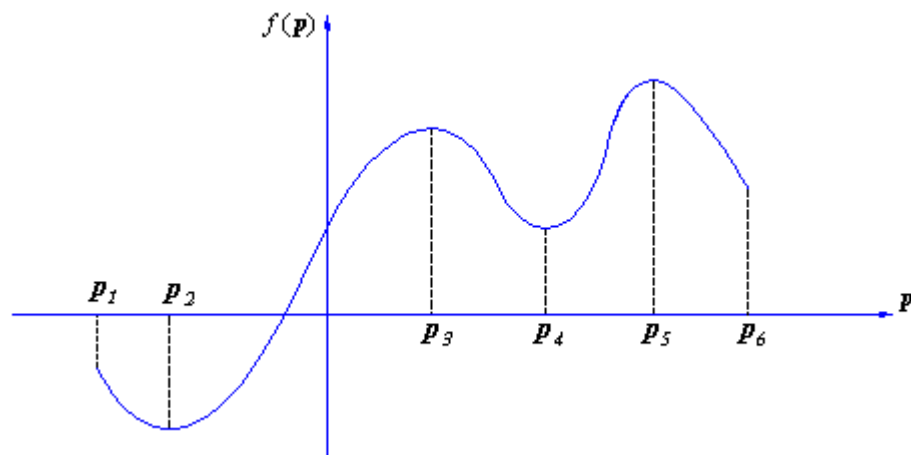


Figura 3.1. Pontos de mínimos e máximos, globais e locais para uma função $f(p)$, com $p \in P \subseteq R$.

Um algoritmo de otimização local tenta encontrar um mínimo local nas proximidades da solução inicial ou de partida. Quase todos os métodos de busca local são determinísticos e usam propriedades locais da função objetivo para calcular uma atualização da solução atual e procurar na direção descendente a solução final. Esses tipos de algoritmos podem não encontrar o mínimo global dependendo da aproximação inicial. Um algoritmo de otimização global tenta encontrar o mínimo global de uma função objetivo multimodal. A maioria dos métodos de buscas globais são estocásticos e usam informações globais sobre a função objetivo para atualizar uma posição corrente. Em geral, a convergência destes métodos a uma solução global ótima não é garantida; no entanto, os algoritmos de busca global têm maior probabilidade de encontrar o mínimo global independentemente da aproximação inicial.

O comportamento da função objetivo também influi na convergência a uma solução global ótima, isto é, a facilidade de se determinar uma solução global aumenta com a suavidade da função objetivo e diminui com a presença de um número elevado de mínimos locais.

Portanto, o objetivo da otimização global consiste em encontrar os mínimos (ou máximos) de uma função de varias variáveis com valores dentro de uma determinada região do espaço multidimensional. Em outras palavras, a tarefa de otimização é encontrar os vetores solução \mathbf{p}^* que minimizem a função objetivo $f(\mathbf{p})$. Os vetores \mathbf{p}^* são chamados de pontos ótimos e os valores correspondentes $f(\mathbf{p}^*)$ são chamados valores ótimos, logo o par \mathbf{p}^* e $f(\mathbf{p}^*)$ constitui uma solução ótima.

3.3 – OTIMIZAÇÃO GLOBAL DOS PARÂMETROS DO EMPILHAMENTO SRC

O problema crucial para a implementação do método de empilhamento SRC, consiste na determinação, a partir dos dados sísmicos, dos três parâmetros (β_0, R_{PIN}, R_N) ótimos associados a cada ponto de amostragem da seção AN a ser simulada. A determinação simultânea destes três parâmetros pode ser realizada por meio de processos de otimização utilizando como função objetivo a medida de coerência *semblance*, aplicada nos dados sísmicos. No processo de otimização este trio ótimo de parâmetros é o que produz o maior valor da função objetivo, isto é, o maior valor de coerência *semblance*. Por outro lado, um trio ótimo de parâmetros associado a um ponto de amostragem da seção AN, define a superfície ou operador de empilhamento SRC que tem o melhor ajuste com os eventos sísmicos dos dados de cobertura múltipla. Neste trabalho, a otimização dos parâmetros SRC é descrita como sendo um processo de minimização, considerando a função *semblance* negativa e conseqüentemente um trio ótimo de parâmetros estará associado ao menor valor da medida de coerência.

Achar o mínimo global em geral é um problema difícil e o sucesso do algoritmo de busca global depende muito do comportamento da função de coerência que neste caso depende de três parâmetros (β_0, R_{PIN}, R_N) , restritos aos seguintes intervalos $-\pi/2 \leq \beta_0 \leq \pi/2$, $0 \leq R_{PIN} \leq +\infty$, $-\infty \leq R_N \leq +\infty$. Por outro lado, na otimização dos parâmetros SRC, a determinação do mínimo global também dependerá das características do método de otimização utilizado, cuja

convergência para mínimo global pode ser comprometida pelo elevado custo de avaliação da função objetivo e o tempo computacional do algoritmo de otimização.

Para explicar de maneira objetiva o problema da otimização dos parâmetros SRC, nos próximos itens serão apresentados a expressão matemática da função objetivo SRC (*semblance*) e também será mostrado uma representação gráfica desta função. Para o cálculo da função objetivo associado a um ponto de amostragem da seção AN foi utilizado os dados Marmusi-2D que são conhecidos na literatura sísmica e bastante usados para testes de novas metodologias de imageamento e inversão de dados sísmicos. A descrição tanto do modelo quanto dos dados de cobertura múltipla Marmusi serão fornecidos no capítulo 5.

3.3.1 – Função Objetivo do Método SRC

O critério de coerência a ser usado neste trabalho é o semblance (Neidell & Taner, 1971) que representa uma medida da razão entre a energia do sinal depois da somatória dos traços e a energia acumulada de todos os traços envolvidos na somatória, expressa como:

$$S = \frac{\sum_t \left(\sum_{i=1}^M U_{i,t(i)} \right)^2}{M \sum_t \sum_{i=1}^M U_{i,t(i)}^2} \quad (3.1)$$

onde $U_{i,t(i)}$ é a amplitude do sinal sísmico indexado pelo número de traços sísmicos $i = 1, \dots, M$ e a trajetória de empilhamento $t(i)$ da expressão (2.1) ou (2.2). O sobrescrito M é o número de traços e \sum_t indica que a somatória é realizada numa janela temporal em relação à trajetória central de empilhamento. O semblance é uma medida de coerência normalizada com valores entre 0 e 1, podendo atingir um valor próximo à unidade somente quando todos os valores do sinal sísmico, ao longo da trajetória de empilhamento, são idênticos.

3.3.2 – Exemplo da função objetivo SRC para três parâmetros (β_0, R_{PIN}, R_N)

Com o propósito de explicar melhor o problema de otimização dos três parâmetros SRC foi constituído um gráfico tridimensional com valores de coerência (Figura 3.2), avaliados para valores dos parâmetros $(\beta_0, R_{PIN}, K_N = 1/R_N)$ sistematicamente distribuídos nos seguintes

intervalos: $-50^\circ \leq \beta_0 \leq 50^\circ$, $0 \leq R_{PIN} \leq 10000m$, $-0.002^{-1}m \leq K_N \leq +0.002^{-1}m$. Portanto, o volume de coerência definido por estes intervalos define o espaço de busca para a otimização dos parâmetros SRC.

O cubo de coerência da Figura 3.2 corresponde ao ponto de amostragem $P_0(t_0 = 1.27s, x_0 = 7375m)$ da seção de afastamento nulo dos dados Marmusi. Os valores de coerência foram calculados usando o negativo da função *semblance* (3.1) avaliada pelo somatório dos eventos sísmicos ao longo das superfícies de empilhamento definidas por (2.1). Devido ao fato de que o ponto de amostragem considerado encontra-se exatamente na interseção de dois eventos sísmicos, também denominados de ventos com mergulhos conflitantes, o cubo de coerência mostra dois mínimos bem definidos que correspondem as duas zonas que se destacam por ter os valores ótimos de coerência. O mínimo global que tem o valor ótimo de coerência está localizado em $\beta_0 = -3.03$ graus, $R_{PIN} = 1624.5$ m e $K_N = 0.00054545^{-1}$ m. Este trio de parâmetros define a superfície de empilhamento SRC que melhor ajusta os eventos sísmicos dos dados de cobertura múltipla correspondentes apenas a um dos eventos com mergulho conflitante da seção AN. Associado ao mesmo ponto de amostragem da seção AN tem-se outro trio de parâmetros correspondente ao mínimo local, e portanto outra superfície de empilhamento SRC correspondente ao segundo evento sísmico com mergulho conflitante.

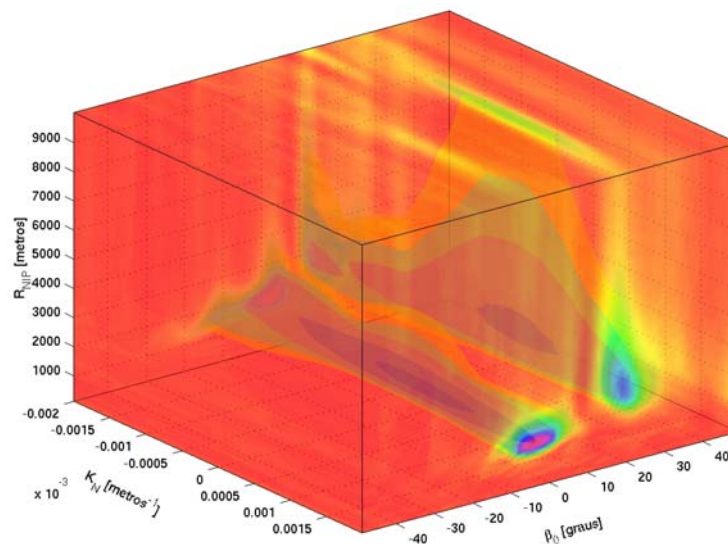


Figura 3. 2. Valores de coerência para o ponto P_0 . O valor mínimo de coerência está localizado em $\beta_0 = -3.03^\circ$, $R_{PIN} = 1624.5$ m e $K_N = 0.00054545^{-1}$ m

Quando existe a interseção de eventos sísmicos, para a simulação correta da interferência entre esses eventos na seção AN, devem ser determinados os mínimos associados a cada um dos eventos conflitantes. No caso da função coerência da Figura 3.2 ambos os mínimos devem ser determinados. No entanto, neste trabalho trataremos apenas da determinação do mínimo global. Pode-se dizer, portanto, que o problema de otimização do empilhamento SRC consiste em encontrar o mínimo global por meio da aplicação de técnicas de otimização global. Em geral a otimização global é um processo dispendioso, e neste caso particular torna-se ainda mais dispendioso por causa do custo da avaliação da função coerência que envolve grandes volumes de dados sísmicos. Neste sentido existe a necessidade do estudo dos métodos de otimização que tenham melhor desempenho quando aplicados a determinação dos parâmetros SRC.

3.3.3 – Exemplo da função objetivo SRC para dois parâmetros (β_0 , R_{PIN})

Em geral, o custo computacional de um processo de otimização é maior quanto maior o número de parâmetros a serem procurados. Portanto, com o objetivo de determinar os parâmetros SRC com um menor custo computacional, foram desenvolvidas algumas estratégias de busca que não determinam simultaneamente os três parâmetros. Por exemplo, usando a expressão simplificada (2.2) o problema de otimização do empilhamento SRC reduz-se á busca global de dos parâmetros β_0 e R_{PIN} .

Também para o ponto $P_0(t_0 = 1.27s, x_0 = 7375m)$ da seção NA dos dados Marmousi, na Figura 3.3 mostra o mapa de coerência cujos valores foram calculados com a função coerência (3.1) usando a função de tempo de transito (2.2) dependente de dois parâmetros (β_0 , R_{PIN}). Devido aos mergulhos conflitantes também neste mapa de coerência existem dois mínimos bem definidos. Neste caso, o problema de otimização do empilhamento SRC consiste na busca, no espaço de busca bidimensional mostrado na Figura 3.3, de dois parâmetros ótimos correspondentes ao mínimo global.

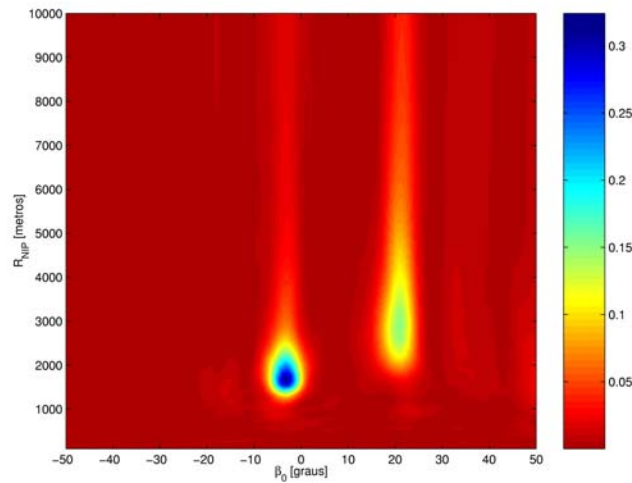


Figura 3.3. Mapa de coerência para valores de β_0 e R_{PIN} , associados a um ponto de amostragem da seção AN, de um modelo sintético, onde R_N é fixo.

Vale mencionar que em alguns dados sísmicos com baixa razão sinal-ruído existem muitos mínimos locais com valores de coerências relativamente baixos comparados com os valores de mínimos locais bem definidos que correspondem aos eventos sísmicos. Assim sendo, a função coerência do empilhamento SRC é uma função multimodal ou de extremo múltiplo e, portanto, a busca simultânea de três parâmetros ($\beta_0, R_{PIN}, K_N = 1/R_N$) ou de dois parâmetros (β_0, R_{PIN}) deve ser realizado utilizando um algoritmo de otimização global.

4 - ALGORITMOS DE OTIMIZAÇÃO GLOBAL

Neste capítulo será dada uma breve apresentação dos métodos de otimização que propomos adaptar para a busca dos parâmetros do empilhamento SRC e faremos uma comparação entre os desempenhos dos mesmos. Uma primeira comparação de métodos de otimização global aplicados no problema de otimização do método SRC foi feita em Carmo & Garabito (2003). Nesta oportunidade, determinaremos os parâmetros do operador SRC aplicando quatro métodos de otimização global: *Simulated Annealing* (SA), *Very Fast Simulated Annealing* (VFSA), *Differential Evolution* (DE) e *Controlled Random Search – 2* (CRS2).

Serão efetuados testes comparativos referentes aos seguintes indicadores de desempenho:

- **Eficácia:** consiste da avaliação se o algoritmo atingiu um ponto de mínimo (critério de parada).
- **Eficiência:** é uma medida do número de avaliações da função que foram necessárias. Também está relacionado com o tempo de funcionamento do algoritmo.
- **Confiabilidade ou robustez:** é medida pelo número de acertos em achar o mínimo global, ou pelo menos chegar suficientemente próximo.

Neste capítulo os testes de otimização são realizados nos dados de cobertura múltipla do modelo Marmousi que serão apresentados no Capítulo 5.

A seguir serão apresentados os resultados da aplicação dos quatro métodos de otimização aplicados tanto para a busca simultânea de três parâmetros ($\beta_0, R_{PIN}, K_N = 1/R_N$) como para a busca de dois parâmetros (β_0, R_{PIN}). Cada um dos gráficos mostrados nos seguintes itens, representa vários testes realizados com cada um dos algoritmos de otimização estudados, onde o eixo horizontal representa o número de avaliações da função objetivo e o eixo vertical o valor da função objetivo (negativo do *semblance*). Estes gráficos mostram a evolução dos algoritmos até atingir o mínimo global.

4.1. ALGORITMO *SIMULATED ANNEALING*

Simulated annealing (SA) é um método de otimização global que imita o processo físico pelo qual um cristal se desenvolve (ou cristaliza) a partir do esfriamento lento de um cristal fundido, até atingir o estado de energia mínima. Flutuações aleatórias na energia permitem ao sistema de esfriamento lento escapar de um estado de energia mínima local para alcançar o estado de energia mínima global. Neste trabalho será utilizado o algoritmo *Simulated Annealing* modificado por Corana et al. (1987). Mais detalhes deste método encontra-se no apêndice A.

No método SA o processo de minimização começa com um dado valor inicial de temperatura ($T = T_0$) e, em geral, a partir de um vetor de parâmetros iniciais ($\mathbf{p} = \mathbf{p}_0$) que é gerado aleatoriamente, uma vez que este não tem influência decisiva na determinação final dos parâmetros ótimos. Então, dado um ponto inicial \mathbf{p} e o respectivo valor da função objetivo $f(\mathbf{p})$, a geração de um ponto teste \mathbf{p}' , é realizada por meio da aplicação de uma perturbação em cada elemento de \mathbf{p} , segundo a relação (Corana et al., 1987):

$$p'_i = p_i + \lambda v_i, \quad i = 1, 2, \dots, n \quad (4.1.1)$$

onde λ é um número aleatório uniforme distribuído entre $[-1, 1]$ e v_i é o i -ésimo elemento do vetor dos comprimentos do passo \mathbf{v} . A variável n representa o número de parâmetros a serem otimizados. Um ponto teste \mathbf{p}' é aceito incondicionalmente se o valor de $f(\mathbf{p}')$ for menor que $f(\mathbf{p})$ (i.e., $\Delta f = f(\mathbf{p}') - f(\mathbf{p}) < 0$). Então, o ponto \mathbf{p} é substituído por \mathbf{p}' e o algoritmo tem um deslocamento descendente. Se o novo $f(\mathbf{p})$ tiver o menor valor registrado, este valor e o vetor \mathbf{p} correspondente são guardados como os atuais valores ótimos. Quando o valor $f(\mathbf{p}')$ do ponto teste tem um valor igual ou maior a $f(\mathbf{p})$ (i.e., $\Delta f \geq 0$), um critério de probabilidade decide se esse ponto teste ascendente é aceito ou não, com base na relação de probabilidade conhecida como critério de Metropolis:

$$P = e^{-\Delta f / T} \quad (4.1.2)$$

em que T é a temperatura. Se o novo ponto teste for aceito, então \mathbf{p} e $f(\mathbf{p})$ são substituídos por \mathbf{p}' e $f(\mathbf{p}')$, respectivamente, e o algoritmo tem um deslocamento ascendente.

A principal característica deste algoritmo é a perturbação do vetor dos parâmetros \mathbf{p} , controlado por um vetor dos comprimentos dos passos, \mathbf{v} , que é ajustado automaticamente após

um determinado número de ciclos de geração e aceitação de pontos testes. O ajuste dos vetores dos comprimentos dos passos é realizado procurando uma melhor amostragem da função objetivo.

Depois de realizar vários ciclos de geração de pontos testes a temperatura T é diminuída segundo a seqüência de esfriamento dada por:

$$T_{i+1} = r_T T_i, \quad (4.1.3)$$

em que r_T é definido no intervalo (0,1). Conforme diminui a temperatura os comprimentos dos passos declinam e o algoritmo restringe sua busca na área mais promissora para a otimização, sendo atingido assim o mínimo global. O algoritmo é interrompido quando o critério de parada é cumprido.

A Figura 4.1a ilustra o desempenho algorítmico do SA para a busca de três parâmetros (β_0, R_{PIN}, R_N), onde podemos observar que o algoritmo fica preso no mínimo local devido a rápida queda de temperatura e também por causa deste mínimo local ser bem definido. No entanto, aproximadamente 70% dos testes realizados atingem o mínimo global por volta de 380 avaliações da função objetivo. A Figura 4.1b ilustra o desempenho algorítmico do SA para busca de dois parâmetros (β_0, R_{PIN}), assim como na busca de três parâmetros, para a busca dos dois parâmetros o algoritmo SA também fica preso no mínimo local, devido a queda rápida de temperatura e ao mínimo local ser bem definido, tendo a maioria dos testes atingido o mínimo global por volta de 340 avaliações da função objetivo.

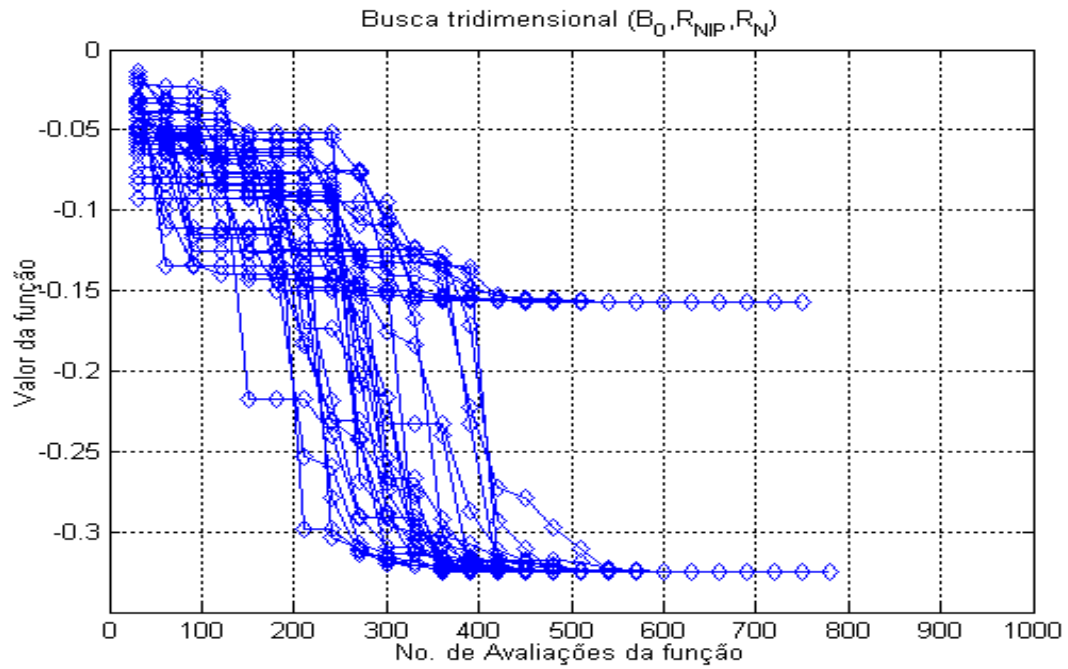


Figura 4.1a. Desempenho do algoritmo SA após várias avaliações da função para busca de três parâmetros.

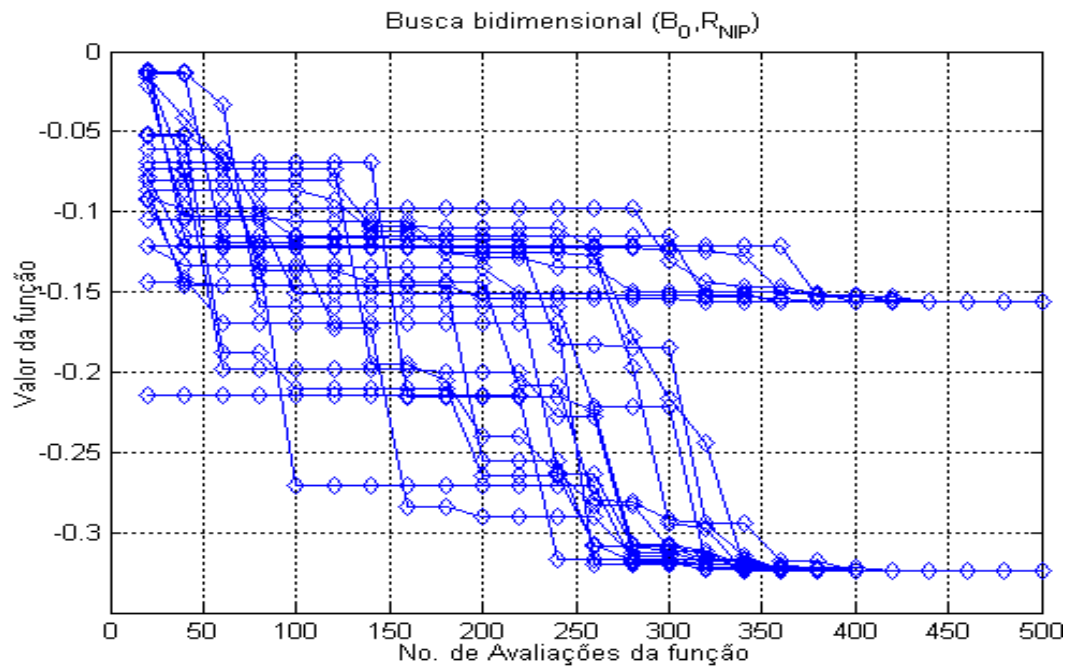


Figura 4.1b. Desempenho do algoritmo SA após várias avaliações da função para busca de dois parâmetros.

4.2. ALGORITMO *VERY FAST SIMULATED ANNEALING*

Objetivando melhorar o desempenho do algoritmo SA padrão também denominado de algoritmo *Boltzmann Annealing* (BA) vários autores propuseram variantes deste algoritmo. Uma destas variantes é o algoritmo SA modificado proposto por Corana et al., (1987) que foi apresentada no anterior item. Como outro exemplo podemos citar o algoritmo *Fast Simulated Annealing* (FSA) que é considerado mais eficiente que o algoritmo SA padrão e usa uma seqüência de esfriamento de temperatura mais rápida (Szu & Hartley, 1987). O algoritmo *Very Fast Simulated Annealing* (VFSA) proposto por Ingber (1989) introduz varias modificações al algoritmo padrão SA.

No algoritmo VFSA a perturbação de cada elemento de α_k^i da dimensão i , é realizada segundo a relação:

$$\alpha_{k+1}^i = \alpha_k^i + y^i (B_i - A_i), \quad y^i \in [-1, 1], \quad (4.2.1)$$

com

$$\alpha_k^i, \alpha_{k+1}^i \in [A_i, B_i], \quad (4.2.2)$$

onde os novos parâmetros α_{k+1}^i são gerados a partir dos parâmetros antigos α_k^i calculados com a variável aleatória y^i , que é gerada segundo a seguinte expressão:

$$y^i = \text{sgn}(u^i - 1/2) T_i [(1 + 1/T_i)^{2|u^i - 1|} - 1], \quad (4.2.3)$$

onde o número aleatório u^i é obtido a partir da distribuição uniforme $U[0,1]$. Em Ingber (1989) foi mostrado que a convergência para o mínimo global pode ser atingido usando a seqüência de esfriamento dada pela expressão:

$$T_i(k) = T_{0i} \exp(-c_i k^{1/n}), \quad (4.2.4)$$

onde T_{0i} é a temperatura inicial para o parâmetro i do modelo e c_i é o parâmetro livre a ser usado para controlar o decaimento da temperatura e ajudar a sintonizar o algoritmo para problemas específicos. Vale ressaltar que no presente trabalho foi usada uma única temperatura para todos os componentes do vetor de parâmetros e conseqüentemente um único valor para o parâmetro livre. Mais detalhes sobre este método encontra-se no apêndice B.

Neste trabalho utilizamos para a busca bidimensional deste algoritmo $T_0 = 1.5$, $k = 20$ e $c_i = 0.25$, para a busca tridimensional usamos $T_0 = 1.5$, $k = 30$ e $c_i = 0.25$.

A Figura 4.2a ilustra o desempenho algorítmico do algoritmo VFSA para a busca dos três parâmetros SRC (β_0, R_{PIN}, R_N). Podemos observar que o algoritmo também fica preso no mínimo local, devido à rápida queda de temperatura e devido este mínimo ser bem definido ou forte, mas grande parte dos testes realizados, aproximadamente 80%, atingiram o mínimo global por volta de 340 avaliações da função objetivo. A Figura 4.2b mostra o desempenho algorítmico deste método para a busca de dois parâmetros SRC (β_0, R_{PIN}) e notamos que aproximadamente 30% dos testes ficam presos no mínimo local, pelo mesmo motivo citado acima, com os outros 70% dos testes alcançando o mínimo global por volta de 200 avaliações da função objetivo.

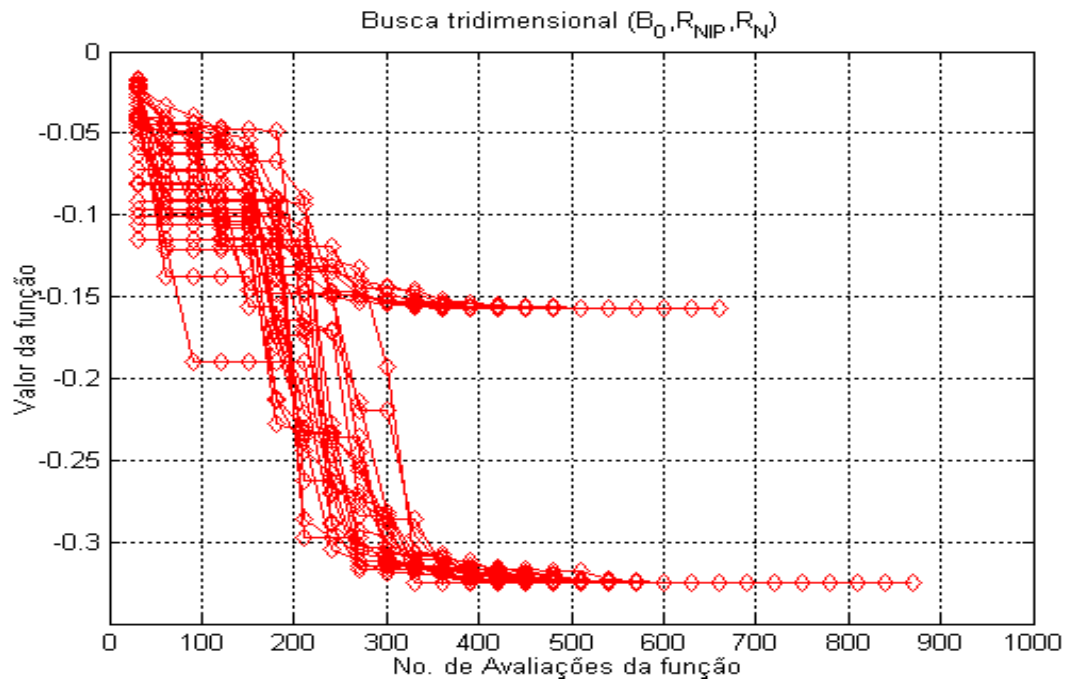


Figura 4.2a. Desempenho do algoritmo VFSA após várias avaliações da função para busca de três parâmetros.

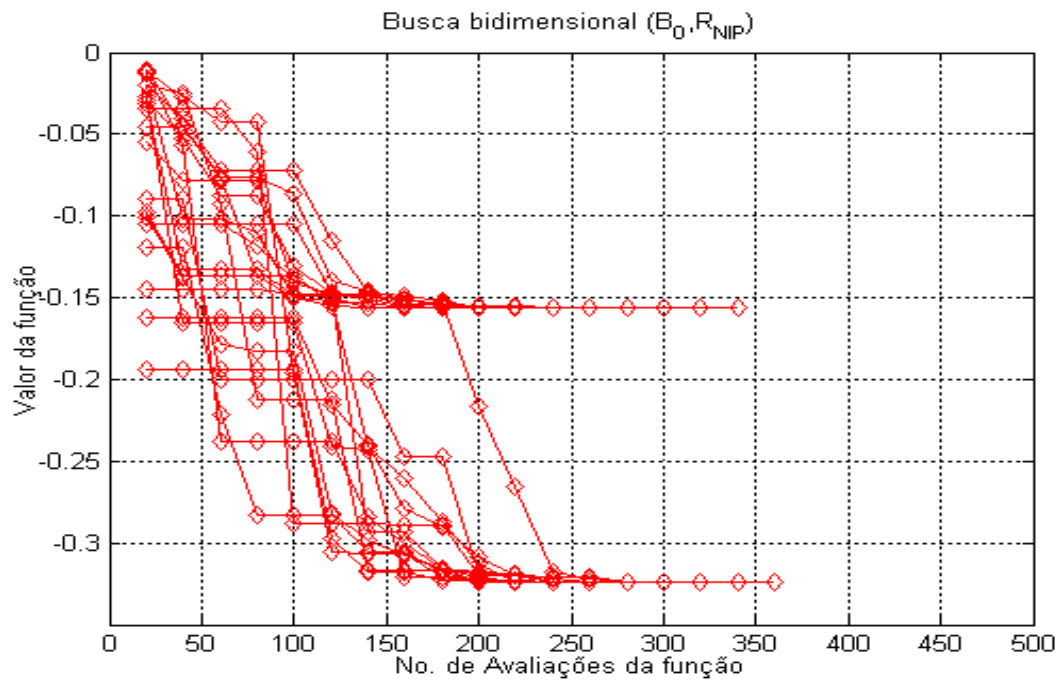


Figura 4.2b. Desempenho do algoritmo VFSA após várias avaliações da função para a busca de dois parâmetros.

4.3. ALGORITMO *DIFFERENTIAL EVOLUTION*

O algoritmo *Differential Evolution* (DE) pertence à classe de algoritmos evolucionários. Os algoritmos evolucionários são variantes de processos de busca aleatória e utilizam a terminologia da biologia e genética, isto é, para uma amostra aleatória em uma dada iteração, pares de indivíduos “pais” (pontos) selecionados com base nas suas “boas condições” (valor da função) recombina e geram uma nova descendência, onde os melhores desta descendência são selecionados para a próxima geração (Solomatine, 1999a). O DE é um método de busca direta que utiliza m vetores de parâmetros de n dimensões, em que a idéia por trás deste processo é um esquema para gerar vetores de parâmetros testes. Basicamente, DE gera novos vetores de parâmetros pela adição a um terceiro vetor, de uma diferença ponderada entre dois vetores da população escolhidos aleatoriamente. Se o vetor resultante produzir um valor menor da função objetivo do que um membro da população predeterminado, o vetor gerado recentemente substituirá o vetor com o qual foi comparado na geração seguinte, caso contrário o vetor antigo é conservado (Storn & Price, 1996). Este processo é repetido até alcançar o mínimo global. Mais detalhes deste método encontra-se no apêndice C.

No presente trabalho foram utilizados uma população $NP = 50$, um fator de amplificação $F = 0.3$ e um cruzamento $CR = 0.8$.

A Figura 4.3a ilustra o desempenho algorítmico do DE para a busca dos três parâmetros SRC (β_0, R_{PIN}, R_N). Podemos notar que o algoritmo não fica preso no mínimo local, embora este ponto de mínimo seja bem definido, necessitando de um número elevado de avaliações da função objetivo para atingir o mínimo global, que neste caso foi de aproximadamente 900 avaliações. A Figura 4.3b mostra o desempenho algorítmico deste método para a busca de dois parâmetros SRC (β_0, R_{PIN}) onde observamos que o algoritmo DE já fica preso no mínimo local poucas vezes, com aproximadamente 98% dos testes atingindo o mínimo global em torno de 700 avaliações da função objetivo.

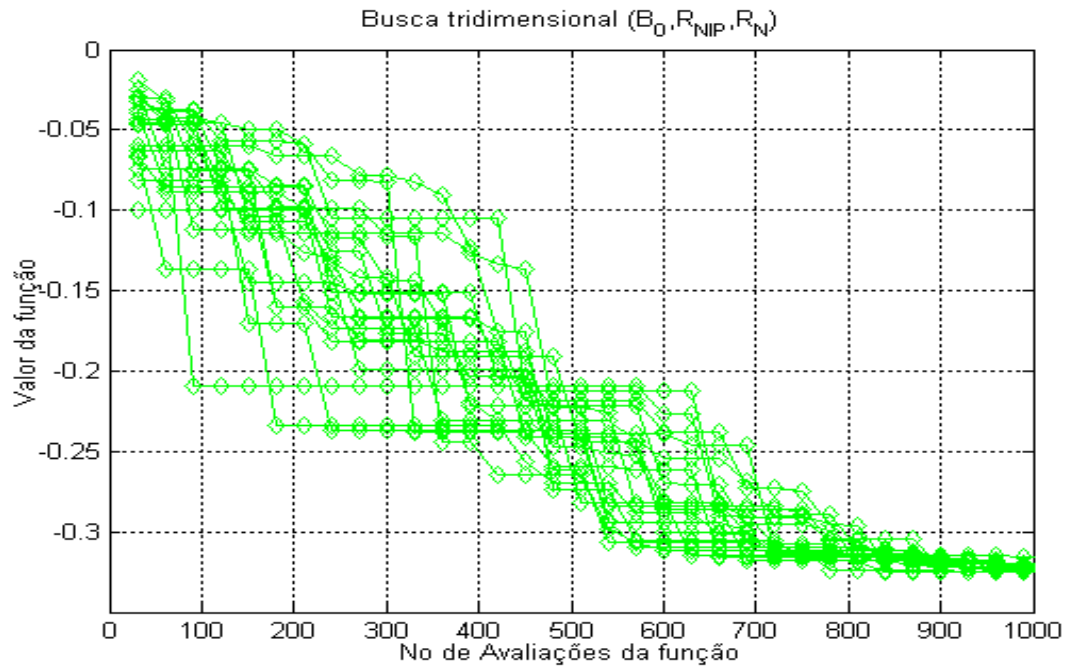


Figura 4.3a. Desempenho do algoritmo DE após várias avaliações da função na busca de três parâmetros.

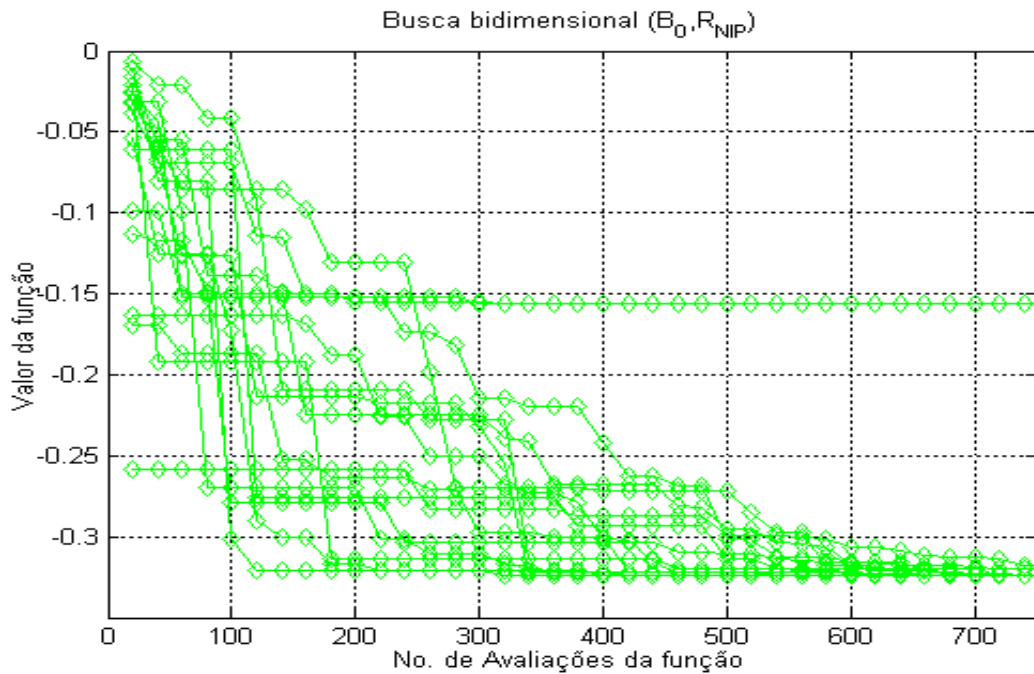


Figura 4.3b. Desempenho do algoritmo DE após várias avaliações da função na busca de dois parâmetros.

4.4. ALGORITMO CONTROLLED RANDOM SEARCH - 2 (CRS2)

O algoritmo *Controlled Random Search 2* (CRS2) é uma variante do algoritmo CRS1 que pertence à classe de algoritmos Monte Carlo (Price, 1983). No algoritmo CRS, dada uma função f de n variáveis independentes, um domínio de busca D_f é definido especificando os limites de cada variável. Um conjunto de N pontos testes é escolhido aleatoriamente sobre D_f , conforme as restrições adicionais. A função é avaliada em cada ponto teste, a posição e o valor da função correspondente a cada ponto são armazenados em uma matriz A . Do conjunto de N pontos é escolhido o ponto \mathbf{p} que tem o menor valor da função objetivo ($f(\mathbf{p})$). Em cada iteração, um novo ponto teste \mathbf{p}' é selecionado aleatoriamente de um certo subconjunto de pontos testes pertencentes aos N pontos, onde $N \gg n$. Desde que a posição de \mathbf{p}' satisfaça as restrições, a função é avaliada em \mathbf{p}' e o valor da função $f(\mathbf{p}')$ é comparado com $f(\mathbf{p})$. Se $f(\mathbf{p}') < f(\mathbf{p})$, então \mathbf{p} é trocado (em A) por \mathbf{p}' . Do contrário, se $f(\mathbf{p}') > f(\mathbf{p})$, então o teste é descartado e um novo ponto é escolhido do conjunto potencial de pontos testes. Na medida que o algoritmo evolui, o subconjunto atual de N pontos armazenados tende a se agrupar ao redor do mínimo global, cujos valores da função são menores que o atual valor de $f(\mathbf{p})$.

Em cada iteração, são escolhidos aleatoriamente n pontos do conjunto de N pontos (R_2, \dots, R_{n+1}). No algoritmo CRS2 o ponto R_1 não é escolhido aleatoriamente e toma sempre o valor do ponto com o menor valor da função do conjunto de N pontos. Estes pontos constituem um *simplex* no espaço n -dimensional, onde o ponto R_{n+1} é o pólo aleatoriamente escolhido do *simplex*, e o próximo ponto teste P , ponto imagem do pólo com respeito ao centróide G dos restantes n pontos (Price, 1983), é gerado pela seguinte relação:

$$\bar{P} = 2\bar{G} - \bar{R}_{n+1}, \quad (4.1)$$

os pontos teste gerados por esse processo são chamados pontos testes primários.

Para este trabalho utilizamos os valores de $N = 30$ para $n = 2$ e $N = 40$ para $n = 3$. A Figura 4.4a ilustra o desempenho algorítmico do CRS2 para a busca dos três parâmetros SRC (β_0, R_{PIN}, R_N), onde podemos notar que o algoritmo fica preso ao mínimo local devido a boa definição deste ponto, porém 80% dos testes atingem o mínimo global com aproximadamente 970 avaliações da função objetivo.

A Figura 4.4b, ilustra o desempenho algorítmico deste método para a busca de dois parâmetros SRC (β_0 , R_{PIN}), e podemos observar que o algoritmo fica poucas vezes preso no mínimo local, com a grande maioria dos testes atingindo o mínimo global em torno de 450 avaliações da função objetivo.

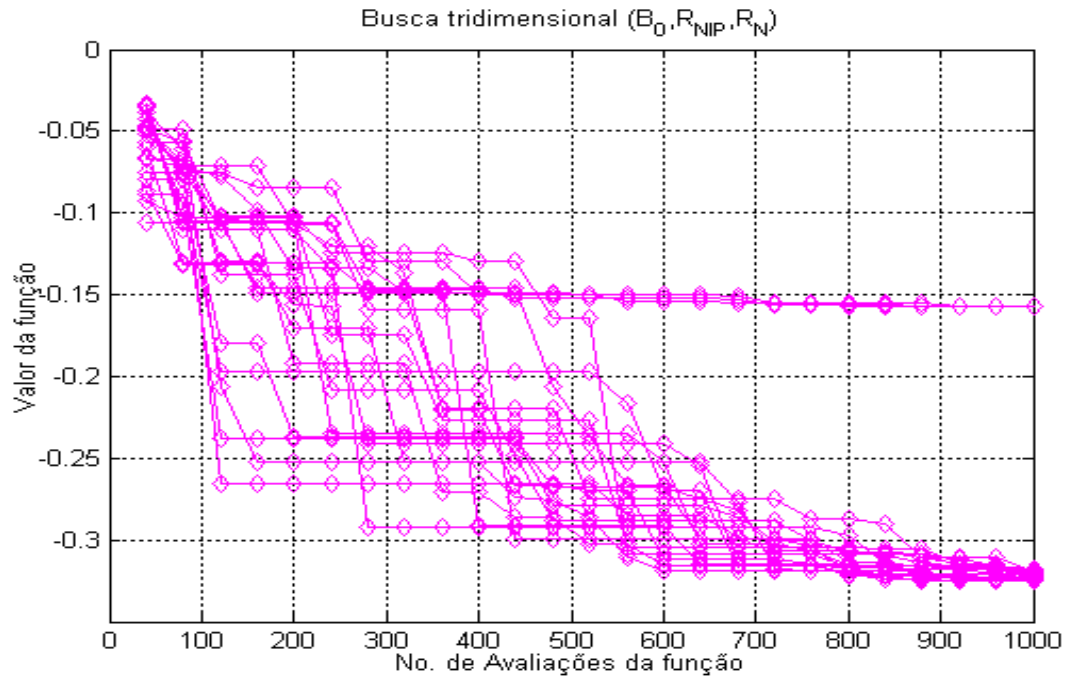


Figura 4.4a. Desempenho do algoritmo CRS2 após várias avaliações da função na busca de três parâmetros.

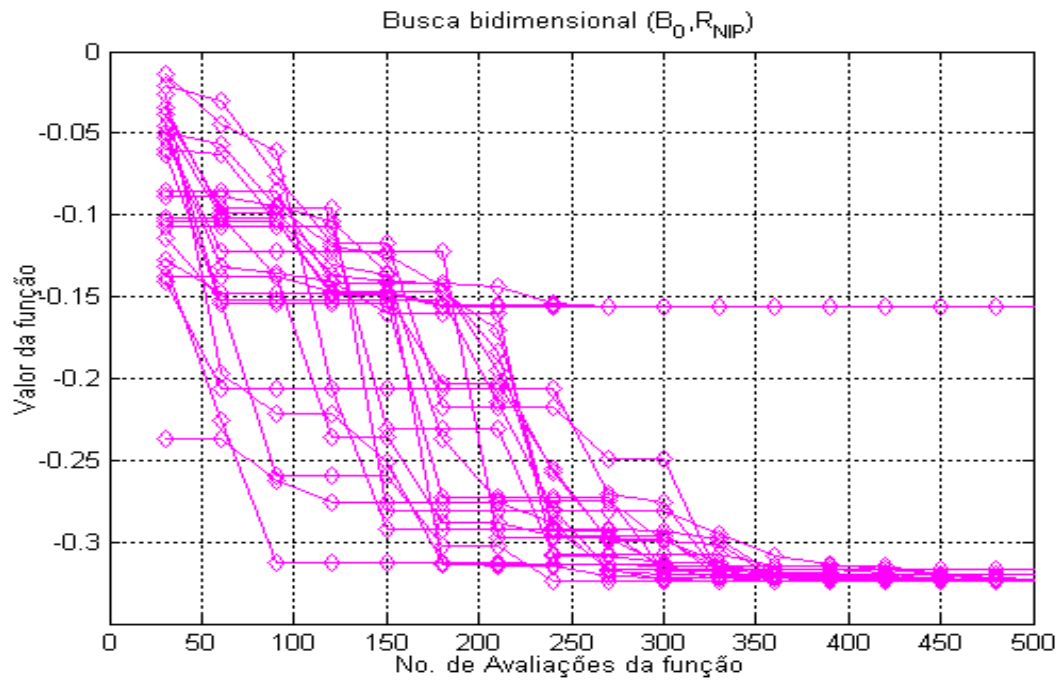


Figura 4.4b. Desempenho do algoritmo CRS2 após várias avaliações da função na busca de dois parâmetros.

4.5. DISCURSÃO DOS ALGORITMOS

Analisando os resultados acima podemos observar que o algoritmo VFSA mostrou ser o mais eficaz e eficiente, isto é, o mais rápido em encontrar o valor mínimo da função objetivo tanto para a busca de dois parâmetros como para a busca de três parâmetros, com uma média de 200 e 330 avaliações da função, respectivamente. No entanto, devido à rápida diminuição da temperatura este algoritmo mantém-se preso no mínimo local e em alguns casos consegue sair do mínimo após um certo numero de iterações.

O algoritmo SA tem uma diminuição de temperatura mais lenta do que o VFSA ficando em segundo lugar quanto à eficácia e eficiência, com uma média de 340 avaliações da função para a busca de dois parâmetros e 370 avaliações para a busca de três parâmetros. Devido à diminuição rápida de temperatura, observa-se em vários testes este algoritmo também fica preso no mínimo local, diminuindo assim o grau de confiabilidade ou robustez.

Em seguida temos os algoritmos DE e CRS2. O CRS2 tem uma boa eficiência para a busca de dois parâmetros com uma média de 450 avaliações da função e poucas vezes fica preso no mínimo local, enquanto que o algoritmo DE, para esta mesma busca, necessita em média de 700 avaliações da função para encontrar o mínimo global, ficando também poucas vezes preso no mínimo local, tendo assim um maior grau de confiabilidade.

Embora o CRS2 seja mais eficiente do que o DE para busca de dois parâmetros, torna-se extremamente lento para a busca de três parâmetros considerando que o algoritmo DE também teve uma considerável queda de eficiência. Com uma média de 900 avaliações da função para DE e 970 avaliações para o CRS2, notamos também que DE não fica preso no mínimo local enquanto que o CRS2 permanece preso apenas em poucos teste.

A confiabilidade ou robustez pode ser medida pelo número de êxitos em achar o mínimo global com exatidão pré-definida. Por causa do caráter aleatório da busca, nenhum algoritmo pode ser 100% de confiança. Portanto, como mostrado nos resultados acima, os algoritmos CRS2 e DE são muito robustos para a busca de três parâmetros, enquanto que SA e VFSA podem ser considerados de menor grau de confiança ou robustez. Por outro lado, para a busca de dois parâmetros todos os algoritmos podem ser considerados de confiança.

5 - EMPILHAMENTO SRC POR MEIO DE OTIMIZAÇÃO GLOBAL

Neste capítulo será apresentada uma descrição do modelo Marmousi, que representa uma situação geológica real, e também os dados de reflexão sísmica gerados a partir desse modelo. Além disso, também serão mostrados os resultados do empilhamento SRC feitos para dois e três parâmetros estimados anteriormente por meio de busca global.

5.1. O MODELO MARMOUSI

O modelo Marmousi é uma série de dados de reflexão sísmica produzida computacionalmente através de diferenças finitas, a partir de um modelo geologicamente representativo. Estes dados foram produzidos no Instituto Francês do Petróleo com o objetivo de servir como referência para testes com diferentes técnicas de inversão e imageamento de reflexão sísmica, tendo sido apresentado e discutido no 52º congresso de Associação Européia de Geofísica em Compenhagen no ano de 1990.

O modelo Marmousi foi criado com base em um perfil geológico localizado em Quengela Norte da bacia Cuanza em Angola. Com base neste perfil foi construído um modelo geométrico bidimensional (Figura 5.1) com 9200 metros de extensão horizontal e 3000 metros na vertical, constituído por 160 camadas onde se tem distribuições de velocidades e densidades com gradientes horizontais e verticais realísticos.

Geologicamente o modelo consiste de um intervalo de sedimentos deltaicos com espaçamentos de oeste para leste, depositado sobre uma série de evaporitos salíferos, onde a parte mais oriental do modelo parece ter sido afetada pela movimentação lateral do deslizamento da camada de sal, dando lugar ao desenvolvimento de falhas normais. Também, têm-se ocorrência de depósitos de carbonatos plataformais dobrados, nos quais uma armadilha estrutural de hidrocarbonetos é esperada.

A história geológica com base nesta seção possui duas fases distintas:

- a) A primeira fase corresponde a uma contínua sedimentação plataformal. No final desta sedimentação os depósitos foram suavemente dobrados e então erodidos.

- b) A Segunda fase iniciou-se com a deposição de uma série de evaporitos salíferos. Nesta série uma outra rica em argilas e margas com matéria orgânica foi depositada. Em seguida foram depositados sedimentos detríticos de folhelhos e arenitos, de que o espaçamento é controlado pela ocorrência de deslocamentos laterais contínuos no sal. Tais deslizamentos na camada de sal também são responsáveis pelo aparecimento de falhas e ativamento das mesmas durante toda a deposição dos sedimentos detríticos.

5.2. DADOS SÍSMICOS

Os dados sísmicos foram gerados segundo a seguinte disposição geométrica:

- a) O arranjo das fontes é constituído de 6 canhões de água espalhados em uma extensão de 40 metros, profundidade de 8 metros e com um intervalo de 8 metros entre cada um deles. Os disparos são realizados simultaneamente e a fonte colocada coincidentemente com o ponto médio do arranjo das fontes. O pulso inicial usado para simular a fonte foi aproximado a partir da digitalização de um pulso real utilizado no campo, ao qual foi aplicado um filtro trapezoidal nas frequências (0, 10, 40, 60 Hz).
- b) O arranjo de receptores é composto de 96 grupos de hidrofones, sendo a distância entre dois grupos consecutivos de 25 metros, com afastamento inicial e final de 200 e 2575 metros respectivamente. Cada grupo de hidrofone consiste de 5 hidrofones, com uma distância de 4 metros entre cada um desses. O traço para cada grupo é obtido a partir da soma de 5 traços registrados. A profundidade dos sensores é de 12 metros.
- c) Com o dispositivo descrito antes, um total de 240 disparos foram simulados de oeste para leste, estando o primeiro e último disparos localizados a 3000 e 8975 metros de extremidade oeste do modelo e com uma distância de 25 m entre dois disparos consecutivos.

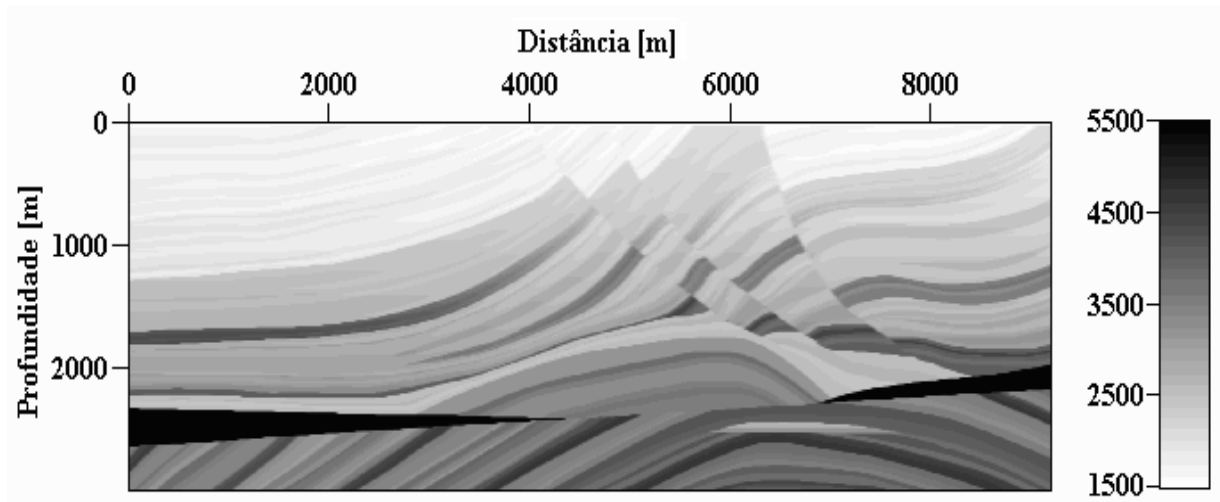


Figura 5.1. Modelo de velocidade dos dados Marmousi

5.3 ALGORITMO DE EMPILHAMENTO SRC

A simulação de uma seção AN por meio do método de empilhamento SRC é realizada através do somatório de eventos sísmicos dos dados de cobertura múltipla contidos nos operador de empilhamento SRC associado a cada um dos pontos imagem da seção AN a ser simulada. Para calcular uma dada superfície ou operador de empilhamento SRC é necessário a determinação, a partir dos dados de cobertura múltipla, dos três parâmetros SRC (β_0 , R_{PIN} , R_N). Existem várias estratégias de busca para se determinar os três parâmetros ótimos que definem um operador SRC com melhor ajuste possível aos eventos sísmicos associados a um dado ponto P_0 . Recentemente, Garabito (2001) apresentou uma estratégia de busca destes parâmetros aplicando os algoritmos de otimização global SA e de otimização local quasi-Newton (QN), que compreende as seguintes etapas: 1) determinação de dois parâmetros (β_0 , R_{PIN}) a partir dos dados pré-empilhados usando o algoritmo SA; 2) determinação R_N a partir dos dados pós-empilhados usando o algoritmo SA; e 3) nesta etapa, usando como solução inicial os resultados das duas etapas anteriores, é realizado o refinamento dos três parâmetros a partir dos dados pré-empilhados usando o algoritmo QN. Aplicando esta estratégia de busca para cada ponto de amostragem P_0 da seção AN, são determinados os três parâmetros SRC e posteriormente, fazendo o somatório dos eventos sísmicos dos dados de cobertura múltipla contidos nessas superfícies de empilhamento, é obtida a seção AN simulada.

Neste trabalho, para a simulação de seções de afastamento nulo pelo método de empilhamento SRC, será aplicado apenas o algoritmo de otimização global VFSA para a busca dos parâmetros que definem a superfície de empilhamento. O primeiro empilhamento será realizado determinando os três parâmetros SRC para cada ponto de amostragem P_0 da seção AN a ser simulada. O segundo empilhamento será realizado determinando somente dois parâmetros SRC para cada ponto de amostragem P_0 da seção AN a ser simulada. Vale salientar que, em ambos os casos o empilhamento dos eventos sísmicos serão realizados ao longo das superfícies definidas com os parâmetros obtidos apenas pela otimização global, ou seja, não será aplicada a etapa de refinamento dos parâmetros usando o algoritmo de otimização local.

5.3 RESULTADOS DO EMPILHAMENTO SRC:

Aqui são apresentados os resultados das seções de afastamento nulo simuladas a partir dos dados Marmousi descritos no início deste capítulo. Como mencionado no item anterior, foi realizado o empilhamento SRC inicialmente usando três parâmetros (β_0 , R_{PIN} , R_N) e posteriormente usando apenas dois parâmetros (β_0 , R_{PIN}). Em ambos casos, foi utilizado o algoritmo de otimização global VFSA que teve o melhor desempenho na busca dos parâmetros SRC (Capítulo 3). Nas Figuras 5.2 e 5.3 são mostrados os resultados dos empilhamentos SRC obtidos através do somatório dos eventos sísmicos contidos nas superfícies de empilhamento SRC definidos por três parâmetros e dois parâmetros, respectivamente.

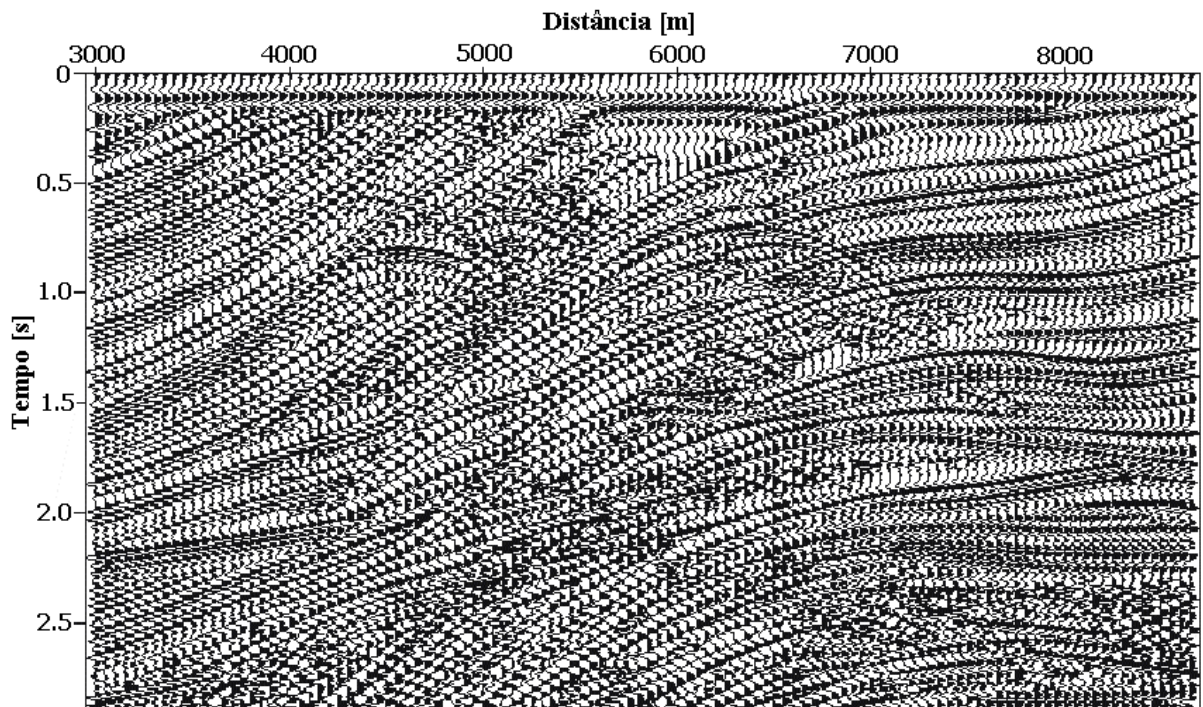


Figura 5.2 - Seção com afastamento nulo, simulada usando o operador de empilhamento, definida por três parâmetros (β_0, R_{PIN}, R_N).

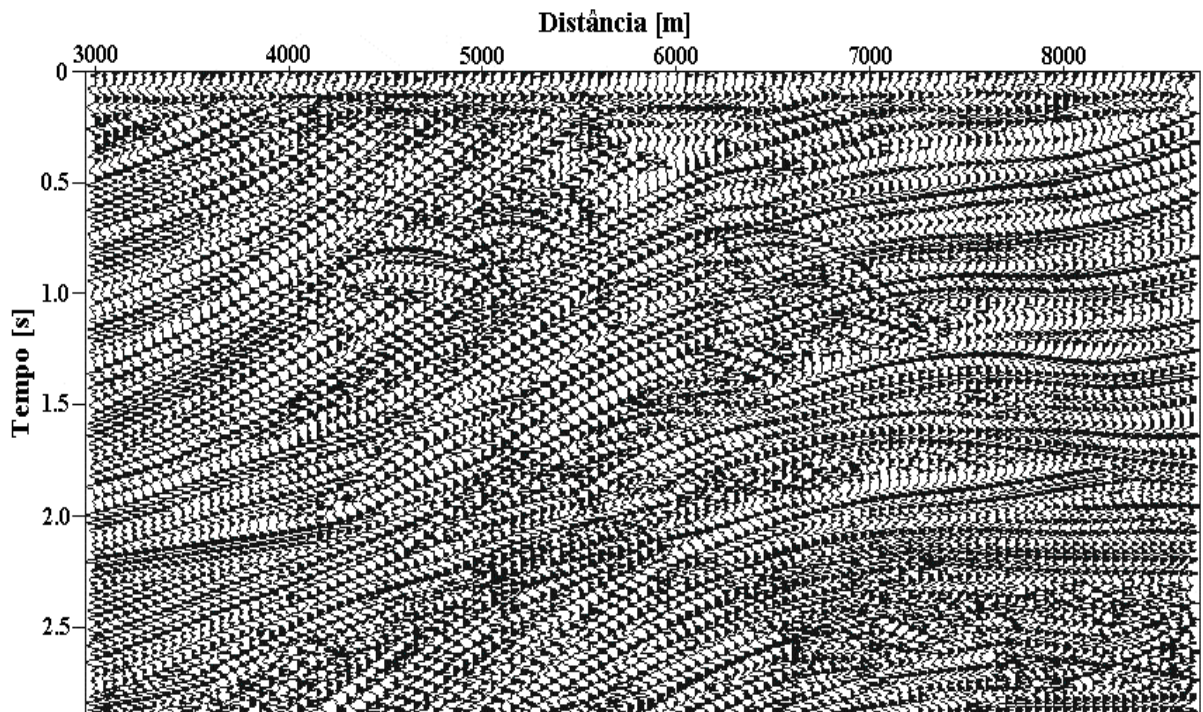


Figura 5.3 - Seção com afastamento nulo, simulada usando o operador de empilhamento, definida por dois parâmetros (β_0, R_{PIN}).

Pode-se dizer que de uma maneira geral estes resultados nos mostram que o algoritmo de otimização global VFSA é robusto, já que as seções AN simuladas em ambos os casos mostram uma boa definição dos eventos sísmicos. No entanto, observa-se que a seção AN obtida usando três parâmetros tem um aumento na razão sinal-ruído e uma melhor resolução dos eventos sísmicos, principalmente em algumas partes correspondentes a tempos maiores. Isto pode ser explicado pelo fato de ter um melhor ajuste, com eventos de reflexão, do operador definido por três parâmetros, o qual implica no uso de maior número de traços no processo de empilhamento. Quanto ao tempo computacional, foi verificado que o processo de busca e empilhamento usando três parâmetros apresenta em média um aumento de 1/3 no tempo de processamento, com relação ao tempo gasto pelo mesmo processamento usando apenas dois parâmetros.

6 – CONCLUSÕES

Neste capítulo são indicadas as conclusões obtidas da aplicação e comparação dos quatro algoritmos de otimização global, descritos anteriormente, no problema de otimização do empilhamento SRC.

CONCLUSÕES

As conclusões descritas a seguir resultam das numerosas avaliações da função objetivo (semblance) pelos algoritmos de otimização global. Foram feitas buscas para dois parâmetros e também para os três parâmetros simultaneamente. Os resultados obtidos foram os seguintes:

O algoritmo CRS2 mostrou ter boa eficiência, eficácia e confiabilidade na busca de dois parâmetros do operador SRC. Mas mostrou uma grande queda de eficiência para a busca de três parâmetros simultaneamente.

Dentre os quatro algoritmos tratados aqui, o algoritmo DE mostrou ser o mais lento para este problema com baixa eficiência tanto para a busca bidimensional como para a busca tridimensional dos parâmetros do operador SRC.

Já o algoritmo SA apresentou muito boa eficiência e eficácia para a busca dos parâmetros SRC, tanto com a busca bidimensional como para a busca tridimensional, tornando-se assim uma boa alternativa para o problema de otimização do método SRC.

No entanto, foi o algoritmo VFSA que apresentou o melhor comportamento para este problema de otimização apresentando alta eficiência e eficácia tanto na busca de dois parâmetros como na busca simultânea do trio de parâmetros do operador SRC. Embora o algoritmo SA também tenha se destacado na busca global destes parâmetros com bons resultados quanto às medidas de eficácia e eficiência, foi o algoritmo VFSA que teve a melhor performance nesses indicadores, apesar de ambos algoritmos apresentarem aproximadamente o mesmo grau de confiabilidade ou robustez.

REFERÊNCIAS BIBLIOGRÁFICAS

- BACK, T. & Schwefel, H.-P. 1993. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, (1): pp. 1-23.
- BIRGIN, G. Ernesto; Biloti, Ricardo; Tygel, Martin; Santos, T. Lúcio.1999. Restricted optimization: a clue to a fast and accurate implementation of the Common Reflection Surface Stack method. *Journal of Applied Geophysics*, 42, 143-155.
- CARMO, L. M. K. & Garabito, G., 2003. Métodos de otimização global aplicados na busca dos parâmetros SRC – 2D. In: *INTERNATIONAL CONGRESS OF THE BRAZILIAN GEOPHYSICAL SOCIETY*, 8., Rio de Janeiro.
- CORANA, A.; Marchesi, M.; Martini, C.; Ridela, S. 1987. Minimizing multimodal functions of continuous variables with “Simulated Annealing algorithm”. *ACM Transactions on Mathematical Software*, 13(3): 262-280.
- DELAURENTI, M., 1999. *Design and optimization techniques of high-speed VLSI circuits*. (PhD thesis). Disponível em: <<http://www.equars.com/~marco/poli/phd/>>. Acesso em: agosto de 2002.
- GARABITO, G., 2001. *Empilhamento Sísmico Por Superfície de Reflexão Comum: Um novo algoritmo usando otimização global e local*. Belém, Universidade Federal do Pará, Centro de Geociências, Pós-Graduação em Geofísica, Belém. (Tese de Doutorado)
- GARABITO, G.; Cruz, J. C. R.; Hubral, P.; Costa, J., 2001a. Common reflection surface stack by global optimization. In: *ANNUAL INTERNAT. MTG., SOC. EXPL. GEOPHYS. EXPANDED ABSTRACTS*, 71.

- GARABITO, G.; Cruz, J. C. R.; Hubral, P. and Costa, J., 2001b. Empilhamento de superfície de reflexão comum com mergulhos conflitantes. Resumo Expandido, In: *CONGRESSO INTERNACIONAL DA SOCIEDADE BRASILEIRA DE GEOFÍSICA*, 7.
- GARABITO, G.; Paschoal, W. G., Lima, A.W., 2003. Comparação dos resultados dos métodos de empilhamentos SRC e CMP (NMO/DMO) aplicados nos dados Marmousi. In: *INTERNATIONAL CONGRESS OF THE BRAZILIAN GEOPHYSICAL SOCIETY*, 8., Rio de Janeiro.
- GELCHINSKY, B. & Keydar, S., 1999. Homeomorphic imaging approach – theory and practice. *J. Appl. Geoph.* 42: 169-228 (Special issue on Karlsruhe workshop on macro model independent seismic reflection imaging).
- GELCHINSKY, B.; Bercovitch, A.; Keydar, S., 1997. *Multifocusing homeomorphic imaging: Parts 1 and 2.* (Special Course on Homeomorphic Imaging, Seeheim, Germany, Course Notes).
- GELCHINSKY, B.; Bercovitch, A.; Keydar, S., 1999a. Multifocussing homeomorphic imaging Part 1. Basic concepts and formulas. *J. Appl. Geoph.* 42: 229-242 (Special issue on Karlsruhe workshop on macro model independent seismic reflection imaging).
- GELCHINSKY, B.; Bercovitch, A.; Keydar, S., 1999b. Multifocussing homeomorphic imaging Part 2. Multifold data set and multifocusing. *J. Appl. Geoph.* 42: 243-260 (Special issue on Karlsruhe workshop on macro model independent seismic reflection imaging).
- GERMAN, S. & German, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images, *IEEE Trans. Patt. Anal. Mac. Int.* 6(6), 721-741.
- HUBRAL, P., 1983. Computing true amplitude reflections in a laterally inhomogeneous earth. *Geophysics*, 48: 1051-1062.

- INGBER, L., 1989. Very fast simulated re-annealing, *Math1. Comput. Modelling*, 12 (8): 967-973.
- INGBER, L., 1993. Simulated annealing: Practice versus theory, *Math1. Comput. Modelling*. 18 (11): 29-57.
- INGBER, L., 1996. Adaptive simulated annealing (ASA): Lessons learned. *Journal Control and Cybernetics*. 25 (1): 33-54.
- JÄGER, R., 1999. *The Common-Reflection-Surface Stack – Theory and Application*. Institute of the Geophysics, University of Karlsruhe. 121p. (Master Thesis).
- JÄGER, R., Mann, J., Höcht, G., Hubral, P., 2001. Common-reflection-surface stack: Image and attributes. *Geophysics*, 66(1): 97-109.
- MAJER, P., 2000. *Inversion of seismic parameters: Determination of the 2-D iso-velocity layer model*. Germany, Karlsruhe University, 75p. (Master Thesis)
- MANN, J., Bergler, S., Zhang, Y., Chira, P., Hubral, P., 2002. Generalizations of the Common-Reflection-Surface Stack. In: *MTG. EUR. ASSN. GEOSCI. ENG.*, 64., Extended Abstracts, Session E023.
- MANN, J., Jäger, R., Höcht, G., Hubral, P., 1999. Common-reflection-surface stack and wavefield attributes. Extended abstracts. In: *INTERNAT. CONGRESS, SOCIEDADE BRASILEIRA DE GEOFÍSICA*, 6.
- METROPOLIS, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.* 21 (6): 1087-1092.
- MÜLLER, T., 1999. *The Common Reflection Surface Stack Method – Seismic imaging without explicit knowledge of the velocity model*. Germany, Karlsruhe University, 131p. (PhD. Thesis)

- MÜLLER, T., Jeager, R., Höcht, G., 1998. Common Reflection Surface Stack Method – Imaging with an unknown velocity model. *SEG ANNUAL CONFERENCE, Ext. Abstr., New Orleans*.
- NEIDELL, N. & Taner, M. 1971. Semblance and other coherency measures for multichannel data. *Geophysics*, 36, 482-497.
- NELDER, J.A., & Mead, R., 1965. A simplex method for function minimization. *Computer Journal*, 7:308-313.
- PRICE, L., W., 1977. Controlled random search procedure for global optimization. *The Computer Journal*, 20 (4): 367-370.
- PRICE, L., W., 1983. Global Optimization by Controlled Random Search. *Journal of Optimization Theory and Applications*, 40 (3):333-348.
- SCHLEICHER, J., Tygel, M., Hubral, P., 1993. Parabolic and hyperbolic paraxial two-point traveltimes in 3D media. *Geophysical Prospecting*, 41(4): 495-514.
- SOLOMATINE, D.P., 1998. Genetic and other global optimization algorithms – comparison and use in calibration problems. In: *PROC. INTERN. CONFERENCE ON HYDROINFORMATICS*, Copenhagen: 1021-1027.
- SOLOMATINE, D.P., 1999a. Two Strategies of Adaptive Cluster Covering with Descent and Their Comparison to Other Algorithms. *Journal of Optimization Global*, 14, (1): 55-78.
- SOLOMATINE, D.P., 1999b. Random search methods in model calibration and pipe network desing. *Water Industry Systems: Modelling and Optimization Applications. Research Studies Press, Ltd. Baldock, UK*, 2: 317-332.

- STORN, R. & Price, K., 1995. *Differential Evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces* (Technical Report TR-95-012, ICSI). Disponível em: <<ftp.icsi.berkeley.edu>>. Acesso em: março de 2002.
- STORN, R. & Price, K., 1996. “Minimizig the real functions of the ICEC’ 96 contest by Differential Evolution”. IEEE CONFERENCE ON EVOLUTIONARY COMPUTAION, Nagoya: 842-844.
- SZU, H. & Hartley, R., 1987. Fast Simulated Annealing, *Phys. Lett. A* 122, (3-4): 157-162.
- TRAPPE, H., Gierse, G., Pruessmann, J. 2001. Case studies show potential of Common reflection Surface Stack – structural resolution in the time domain beyond the conventional NMO/DMO stack. *First Break*, 19: 625 – 633.
- TYGEL, M., Müller, T., Hubral, P., Schleicher, J., 1997. Eigenwave based multiparameter travelttime expansions. In: *ANNUAL MEETING OF SEG, 67., Expanded Abstracts*, 1770-1773.

ANEXOS

A - ALGORITMO SIMULATED ANNEALING

Simulated annealing (SA) é um método de otimização global que imita o processo físico pelo qual um cristal se desenvolve (ou cristaliza) a partir do esfriamento lento (annealing) de um cristal fundido, até atingir o estado de energia mínima. Flutuações aleatórias na energia permitem ao sistema de esfriamento lento escapar de um estado de energia mínima local para alcançar o estado de energia mínima global.

Dada uma função f de n variáveis independentes e seja $\mathbf{p} \in P$, com $P \subset R^n$, deseja-se

$$\min_{\mathbf{p} \in P} f(\mathbf{p})$$

com $a_i \leq p_i \leq b_i$ para $i = 1, 2, \dots, n$, onde $\mathbf{p} = (p_1, p_2, \dots, p_n)$, sendo a_i e b_i , os limites inferiores e superiores respectivamente, da i -ésima componente de \mathbf{p} .

No algoritmo SA o processo de minimização começa com um dado valor inicial de temperatura ($T = T_0$) e, em geral, a partir de um vetor de parâmetros iniciais ($\mathbf{p} = \mathbf{p}_0$) que é gerado aleatoriamente, uma vez que este não tem influência decisiva na determinação final dos parâmetros ótimos. Então, dado um ponto inicial \mathbf{p} e o respectivo valor da função objetivo $f(\mathbf{p})$, a geração de um ponto teste \mathbf{p}' , numa vizinhança de \mathbf{p} , é realizada por meio da aplicação de uma perturbação aleatória ao longo de cada direção de p_i , segundo a relação (Corana et al., 1987):

$$p_i' = p_i + (2\lambda - 1)v_i, \quad i = 1, 2, \dots, n \quad (\text{A.1})$$

onde λ é um número aleatório uniforme distribuído entre $[0, 1]$ e v_i é o i -ésimo elemento do vetor dos comprimentos do passo $\mathbf{v} = (v_1, v_2, \dots, v_n)$, onde ficam registrados os tamanhos dos intervalos ao longo de cada coordenada p_i . A variável n representa o número de parâmetros a serem otimizados. Se o ponto teste \mathbf{p}' cair fora dos limites do espaço de busca definido por \mathbf{a} e \mathbf{b} , um novo ponto teste é gerado por:

$$p_i' = p_i + \lambda(b_i - a_i), \quad i = 1, 2, \dots, n \quad (\text{A.2})$$

até ficar dentro dos limites do espaço de busca.

Um ponto teste \mathbf{p}' é aceito incondicionalmente se o valor de $f(\mathbf{p}')$ for menor que $f(\mathbf{p})$ (*i.e.*, $\Delta f = f(\mathbf{p}') - f(\mathbf{p}) < 0$). Então, o ponto \mathbf{p} é substituído por \mathbf{p}' e o algoritmo tem um deslocamento descendente. Se o novo $f(\mathbf{p})$ tiver o menor valor registrado, este valor e o vetor \mathbf{p}

correspondente são guardados como os atuais valores ótimos. Quando o valor $f(\mathbf{p}')$ do ponto teste tem um valor igual ou maior a $f(\mathbf{p})$ (i.e., $\Delta f = f(\mathbf{p}') - f(\mathbf{p}) \geq 0$), um critério de probabilidade (critério de Metropolis) decide se esse ponto teste ascendente é aceito ou não, com base na relação de probabilidade:

$$P = e^{-\Delta f / T} \quad (\text{A.3})$$

em que T é a temperatura. O valor de P é comparado com P' que é um número aleatório no intervalo $[0,1]$. Se $P > P'$, o novo ponto teste é aceito, isto é, \mathbf{p} e $f(\mathbf{p})$ são substituídos por \mathbf{p}' e $f(\mathbf{p}')$, respectivamente, e o algoritmo tem um deslocamento ascendente. Dois fatores diminuem a probabilidade de um deslocamento ascendente ser aceito: valores baixos de temperatura e valores elevados de Δf .

A principal característica deste algoritmo é a perturbação do vetor dos parâmetros \mathbf{p} , controlado por um vetor dos comprimentos dos passos que é ajustado automaticamente após um determinado número de ciclos de geração pontos testes N_s . O ajuste dos vetores dos comprimentos dos passos \mathbf{v} , que objetiva uma melhor amostragem da função objetivo levando em conta o equilíbrio entre o número de deslocamentos aceitos e rejeitados, é realizado para manter a porcentagem média dos deslocamentos aceitos, ao redor de 50% do número total de deslocamentos (Corana et al., 1987):

$$v'_i = v_i \left[1 + \frac{c_i}{0.4} \left(\frac{N_i}{N_s} - 0.6 \right) \right], \quad \text{se } N_i > 0.6N_s, \quad (\text{A.4a})$$

$$v'_i = \frac{v_i}{1 + \frac{c_i}{0.4} \left(0.4 - \frac{N_i}{N_s} \right)}, \quad \text{se } N_i < 0.4N_s, \quad (\text{A.4b})$$

$$v'_i = v_i, \quad \text{se } 0.4N_s \leq N_i \leq 0.6N_s \quad (\text{A.4c})$$

em que v_i é o comprimento do passo atual, c_i é um parâmetro que controla a variação do passo da i -ésima componente e N_i é o número de deslocamentos aceitos na i -ésima direção após N_s tentativas. A razão N_i/N_s é restrita ao intervalo $[0, 1]$.

Um grande número de deslocamentos aceitos com relação àqueles rejeitados significa que a função objetivo é amostrada (ou explorada) com passos muito pequenos e, contrariamente, um grande número de deslocamentos rejeitados significa que a função objetivo é amostrada com passos muito longos, isto é, os novos pontos testes são gerados muito longe do ponto atual.

Uma boa proporção 1:1 entre deslocamentos aceitos e rejeitados garante que o algoritmo faça um bom mapeamento das feições da função.

Para um determinado valor de N_T que regula o número de gerações de pontos-testes numa dada temperatura, a temperatura T diminui depois de $N_T N_S$ iterações segundo a relação:

$$T_{j+1} = r_T \cdot T_j \quad (\text{A.5})$$

Em que r_T é o fator de decaimento da temperatura e está restrito ao intervalo (0,1). Conforme diminui a temperatura os comprimentos dos passos declinam e o algoritmo restringe sua busca na área mais promissora para a otimização, sendo atingido assim o mínimo global.

O algoritmo é interrompido quando as diferenças, resultantes da comparação do valor final da função objetivo da temperatura atual com os valores finais da função objetivo das últimas N_ε temperaturas e com o valor ótimo da função objetivo, são menores do que o valor de tolerância do erro ε .

O algoritmo SA, nas temperaturas iniciais, primeiro faz um reconhecimento ou uma amostragem ampla da função objetivo, usando para isto comprimento dos passos longos. A medida que a temperatura diminui, os comprimentos dos passos também diminuem, e o algoritmo reduz sua busca para as regiões promissoras. Por outro lado, devido ao fato de que a probabilidade de aceitar um passo em uma direção ascendente é sempre maior que zero, o algoritmo SA pode sair de um mínimo local.

Portanto, uma apropriada seleção da temperatura inicial $T = T_0$ é crucial para a eficiência do algoritmo SA. Se T_0 for muito baixo os comprimentos dos passos podem ser muito pequenos e a área contendo o mínimo global pode não ser visitada pelo algoritmo. Isto porque a solução atual pode não conseguir escapar do mínimo local. Por outro lado, se T_0 for muito alto os comprimentos dos passos podem ser bastante grandes, e uma área ampla demais é visitada, tendo portanto, um elevado número de aceitações dos pontos testes, o que implica num incremento do tempo computacional.

A.1 - DETERMINAÇÃO DA TEMPERATURA INICIAL E CRÍTICA PARA OS PARÂMETROS SRC

Como mencionado anteriormente, um dos parâmetros chave para o controle da eficiência de um algoritmo SA é a temperatura inicial T_0 , que deve ser estimada antes da aplicação do mesmo. Se T_0 for muito baixo os comprimentos dos passos podem ser muito pequenos e a área contendo o mínimo global pode não ser visitada pelo algoritmo. Isto porque a solução atual pode não conseguir escapar do mínimo local. Por outro lado, se T_0 for muito alto os comprimentos dos passos podem ser muito grande, e uma área ampla demais é visitada, tendo portanto, um elevado número de aceitação dos pontos testes, o que implica num incremento do tempo computacional.

Teste item apresenta-se a estimativas das temperaturas iniciais e críticas para os parâmetros do empilhamento CRS. Para isto foi escolhido um ponto de amostragem da seção afastamento nulo dos dados de cobertura múltipla de um modelo sintético.

No algoritmo SA aplicado neste trabalho, a temperatura inicial pode ser estimada com base nas variações dos comprimentos dos passos de cada parâmetro, observados com ajuda de alguns poucos testes. Para um dado ponto P_0 da seção afastamento nulo, foram realizados testes com uma temperatura inicial $T_0 = 1,0$, e foram monitoradas as variações dos comprimentos dos passos com a diminuição da temperatura para os três parâmetros. As relações dos comprimentos dos passos com a temperatura são mostrados na Figura 1A. Para o parâmetro β_0 (Figura 1A-a) observa-se que os comprimentos dos passos praticamente não mudam para temperaturas maiores que 0,06, porém havendo uma diminuição dos passos para temperaturas menores a 0,06. Esta temperatura ($T = 0,06$) é chamada temperatura crítica, a partir da qual começa uma busca eficiente do algoritmo. Antes da temperatura crítica muitos pontos são testados e muitos deles são aceitos, produzindo assim um desperdício considerável de tempo computacional. Na Figura 1A-b, observa-se que as temperaturas críticas de R_{PIN} e R_N são diferentes, cujos valores são aproximadamente 0,06 e 0,01, respectivamente. No entanto, as temperaturas críticas de β_0 e R_{PIN} são similares. Observa-se que as temperaturas críticas podem ser diferentes para cada um dos parâmetros, então no caso da otimização simultânea dos três parâmetros a temperatura inicial deve ser estabelecida, tomando como referência o valor máximo das temperaturas críticas. Com base nestes resultados, pode-se tomar como temperatura inicial $T_0 = 0,1$.

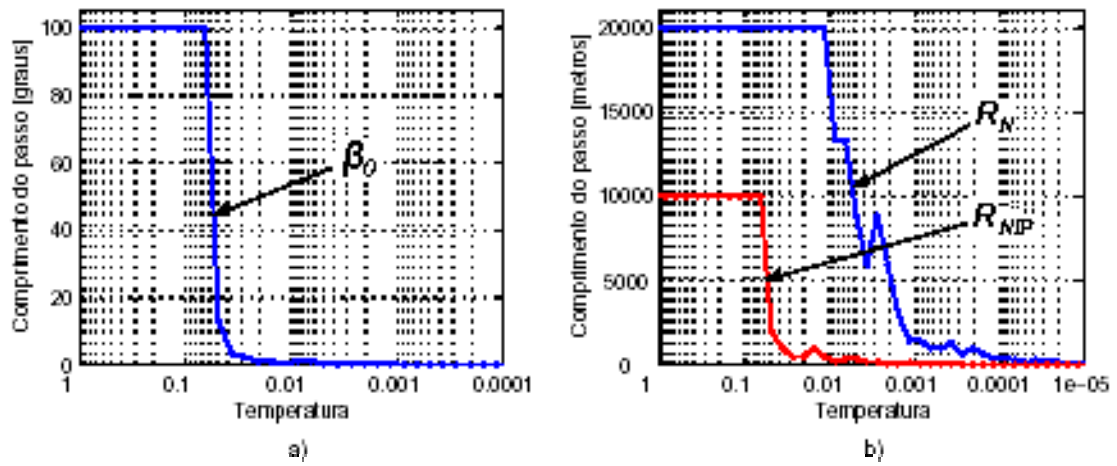


Figura 2A. Determinação da temperatura inicial do algoritmo SA para a busca dos parâmetros: a) ângulo de emergência (β_0) e b) raios de curvatura da onda NIP (R_{NIP}) e da onda N (R_N).

B - ALGORITMO *VERY FAST SIMULATED ANNEALING*

Ao longo do tempo alguns autores, assim como Corana et al., (1987), têm estudado o algoritmo padrão do SA, fazendo modificações para melhorar sua confiabilidade quanto à garantia de convergência para o mínimo global e também a eficiência do algoritmo. Em Corana et al., (1987) foi modificado o algoritmo *Boltzmann Annealing* acelerando o esfriamento (*quenching*) através da introdução de uma seqüência de esfriamento exponencial dada por: $T_{i+1} = r_T T_i$, com r_T entre 0 e 1, onde o valor do fator de esfriamento r_T é escolhido de forma a favorecer a convergência ao mínimo global.

O grande desafio do *simulated annealing* é esfriar a temperatura tão rápido quanto seja possível sem que o algoritmo fique preso num mínimo local. Em outras palavras, é desejada a mais rápida seqüência de esfriamento que preserve a probabilidade igual a um de atingir o mínimo global. A seguir apresenta-se uma breve descrição do *simulated annealing* padrão conhecido como *Boltzmann Annealing* e das variantes conhecidas como *Fast Simulated Annealing* e *Very Fast Simulated Annealing*, frisando que neste trabalho de dissertação foi utilizado o último algoritmo.

B.1 - *BOLTZMANN ANNEALING*

Boltzmann Annealing (BA) foi essencialmente introduzido como uma importante técnica de amostragem de Monte Carlo para calcular integrais de contorno de grandes dimensões surgidos em problemas de física estatística (Metropolis et al., 1953). Este método consiste de três relações funcionais (Szu & Hartley, 1987):

- $g(x)$ - densidade da probabilidade do espaço de n parâmetros $x = \{x^i; i = 1, n\}$;
- $h(x)$ - densidade de probabilidade para aceitar um novo valor da função objetivo apenas a partir do valor anterior.
- $T(k)$ - a seqüência de esfriamento (*annealing*) da temperatura T em k iterações de esfriamento, isto é, mudança da instabilidade ou flutuações das duas densidades de probabilidades anteriores.

Com base num funcional derivado para vários sistemas físicos pertencentes à classe dos sistemas Gauss-Markov, a distribuição de probabilidade g pode ser escolhido como sendo:

$$g(x) = (2\pi T)^{-n/2} \exp[-\Delta x^2 / (2T)], \quad (1B)$$

onde $\Delta x = x - x_0$ é o desvio ou perturbação de x a x_0 , sendo x_0 normalmente o ponto teste previamente escolhido, e T é uma medida das flutuações da distribuição Boltzmann g no espaço de parâmetros x de dimensão n .

A probabilidade de aceitação de um ponto teste está baseada na probabilidade de se obter um novo estado de energia E_{k+1} relativo ao estado de energia anterior E_k ,

$$h(\Delta E) = \exp(-\Delta E/T) \quad (2B)$$

onde ΔE representa a diferença de energia entre os valores atual e anterior da função objetivo apropriados ao problema físico, isto é, $\Delta E = E_{k+1} - E_k$. Este critério de aceitação é conhecido como critério de Metropolis.

Dada a distribuição $g(x)$, para obter um mínimo global de $E(x)$ o decaimento da temperatura T para uma iteração k deve ser obtida conforme a relação (Geman & Geman, 1984):

$$T(k) = \frac{T_0}{\ln k} \quad (3B)$$

com a temperatura inicial T_0 bastante alta. Esta seqüência de esfriamento logarítmico é bastante lenta e portanto precisa de um tempo longo de convergência para o mínimo global.

B.2 - FAST SIMULATED ANNEALING

Embora existam muitas variantes e melhoramentos feitos no algoritmo BA, descrito acima, especificamente foi notado que distribuição de Cauchy tem algumas vantagens definidas sobre a forma da distribuição de Boltzmann (Szu & Harley, 1987). A distribuição de Cauchy

$$g(x) = \frac{T}{(\Delta x^2 + T^2)^{(n+1)/2}}, \quad (4B)$$

tem uma extremidade maior que a forma Gaussiana da distribuição de Boltzmann, permitindo acesso mais fácil para testar o mínimo local na busca do mínimo global desejado, em outras palavras permite ter maior chance para conseguir sair dos mínimos locais. Embora o critério de geração dos parâmetros do modelo seja diferente, é usado o mesmo critério de aceitação (2B) do algoritmo BA. Em adição á nova função para geração dos parâmetros do modelo (4B) a seqüência de esfriamento requerida para a convergência ao mínimo global desejado é dada por:

$$T(k) = \frac{T_0}{k}. \quad (5B)$$

O método do *Fast Simulated Annealing* (FSA) é estatisticamente visto para ter uma seqüência de esfriamento da temperatura exponencialmente mais rápido que o método BA.

B.3 - *VERY FAST SIMULATED ANNEALING*

Os métodos BA e FSA têm distribuições g que amostram intervalos infinitos e não existe previsão para considerar diferenças em cada dimensão de parâmetros, por exemplo, diferentes sensibilidades exigiriam diferentes seqüências de esfriamento da temperatura. No caso do método FSA, não existe um algoritmo rápido suficiente para calcular um gerador aleatório Cauchy de dimensão n . O problema da distribuição de construir uma distribuição de Cauchy de dimensão n , pode ser evitado usando um produto de n distribuições de Cauchy unidimensionais, isto porque a distribuição de Cauchy unidimensional tem alguns algoritmos rápidos. Neste caso também pode ser usado temperaturas diferentes, cada uma correspondendo a diferentes parâmetros do modelo. No entanto o uso de um produto de n distribuições de Cauchy unidimensionais como uma função geradora, ainda precisa de uma seqüência de esfriamento muito lenta para a convergência ao mínimo global.

O problema descrito acima promoveu motivações para o desenvolvimento de um novo algoritmo denominado *Very Fast Simulated Annealing* (Ingber, 1989). Neste algoritmo é proposta uma nova distribuição de probabilidade para geração do modelo, de forma que não é mais exigida uma seqüência de esfriamento lenta. Considere um parâmetro do modelo α_k^i na dimensão i gerado na iteração (ou no instante de esfriamento) k com o limite

$$\alpha_k^i \in [A_i, B_i], \quad (6B)$$

calculado com a variável aleatória y^i ,

$$\alpha_{k+1}^i = \alpha_k^i + y^i(B_i - A_i), \quad y^i \in [-1, 1]. \quad (7B)$$

sendo A_i e B_i os limites inferiores e superiores, respectivamente. O parâmetro y^i é gerado conforme a seguinte distribuição

$$g_T(y) = \prod_{i=1}^n \frac{1}{2(|y^i| + T_i) \ln(1 + 1/T_i)} \equiv \prod_{i=1}^n g_T^i(y^i), \quad (8B)$$

que tem a seguinte distribuição de probabilidade acumulativa

$$G_T^i(y^i) = \frac{1}{2} + \frac{\text{sgn}(y^i) \ln(1 + |y^i|/T_i)}{2 \ln(1 + 1/T_i)}. \quad (9B)$$

A variável aleatória y^i é gerada de um número aleatório u^i obtida a partir da distribuição uniforme $U[0,1]$. Então, o número aleatório u^i pode ser mapeado dentro da anterior distribuição com a seguinte expressão

$$y^i = \text{sgn}(u^i - 1/2) T_i [(1 + 1/T_i)^{|2u^i - 1|} - 1]. \quad (10B)$$

Em Ingber (1989) foi mostrado que para essa distribuição, o mínimo global pode ser estatisticamente obtido usando a seqüência de esfriamento T_i feito através da expressão

$$T_i(k) = T_{0i} \exp(-c_i k^{1/n}). \quad (11B)$$

onde T_{0i} é a temperatura inicial para o parâmetro i do modelo e c_i é o parâmetro livre a ser usado para controlar o decaimento da temperatura e ajudar a sintonizar o algoritmo para problemas específicos.

Na implementação deste algoritmo, novos parâmetros α_{k+1}^i são gerados a partir dos parâmetros antigos α_k^i como segue

$$\alpha_{k+1}^i = \alpha_k^i + y^i (B_i - A_i), \quad (12B)$$

restringidos por

$$\alpha_{k+1}^i \in [A_i, B_i]. \quad (13B)$$

ou seja, as variáveis aleatórias y^i são geradas até que um conjunto n seja obtido satisfazendo estas restrições.

De forma similar ao BA e FSA, o critério de aceitação deste algoritmo também é realizado com a expressão (2B) denominado critério de Metropolis.

C - ALGORITMO *DIFFERENTIAL EVOLUTION*

O algoritmo *Differential Evolution* (DE), recentemente introduzido, tem provado ser um eficiente método para otimizar funções objetivo multimodais. Além da sua boa propriedade de convergência, suas principais vantagens são: simplicidade de conceito e facilidade de uso. O algoritmo DE pertence à classe de algoritmos evolucionários, e é um método de busca direta que utiliza m vetores de parâmetros de n dimensões,

$$\mathbf{p}_i^{(G)}, \quad i = 1, \dots, m \quad (1C)$$

com uma população para cada geração G , isto é, para cada iteração da otimização, onde m não muda durante o processo de minimização.

A população inicial é escolhida aleatoriamente e deve tentar cobrir uniformemente todo o espaço de busca. A idéia crucial por trás do DE é o esquema para gerar vetores de parâmetros testes. Basicamente, DE gera novos vetores de parâmetros pela adição de uma diferença ponderada entre dois vetores da população escolhidos aleatoriamente, a um terceiro vetor. Se o vetor resultante produzir um valor da função objetivo menor do que o valor da função de um membro da população predeterminado, o vetor gerado recentemente substituirá o vetor com qual foi comparado na geração seguinte, caso contrário o vetor antigo é conservado. Este princípio básico pode ser estendido para outras variantes do DE, neste trabalho trataremos da variante DE-1 que é abordada a seguir.

Aplicando uma perturbação em cada vetor $\mathbf{p}_i^{(G)}$, $i = 1, \dots, m$, um novo vetor $\mathbf{v}_i^{(G+1)}$ é gerado conforme a expressão

$$\mathbf{v}_i^{(G+1)} = \mathbf{p}_{r_1}^{(G)} + F(\mathbf{p}_{r_2}^{(G)} - \mathbf{p}_{r_3}^{(G)}) \quad (2C)$$

(Storn & Price, 1996) com $r_1, r_2, r_3 \in [1, m]$, inteiros e mutuamente diferentes, e $F > 0$. Os inteiros r_1 , r_2 e r_3 são escolhidos aleatoriamente do intervalo $[1, m]$. A quantidade $F \in [0, 2]$ é um fator constante e real que controla a amplificação da variação da diferença $(\mathbf{p}_{r_2}^{(G)} - \mathbf{p}_{r_3}^{(G)})$. O vetor $\mathbf{p}_{r_1}^{(G)}$ que é perturbado para produzir $\mathbf{v}_i^{(G+1)}$ não tem nenhuma relação com $\mathbf{p}_i^{(G)}$ mas é um membro da população escolhido aleatoriamente. O esquema DE-1 especifica que o vetor a ser perturbado é escolhido aleatoriamente e que a perturbação consiste de um vetor diferença ponderada. A

Figura 1C ilustra em duas dimensões os diferentes vetores que participam do esquema de geração de vetores.

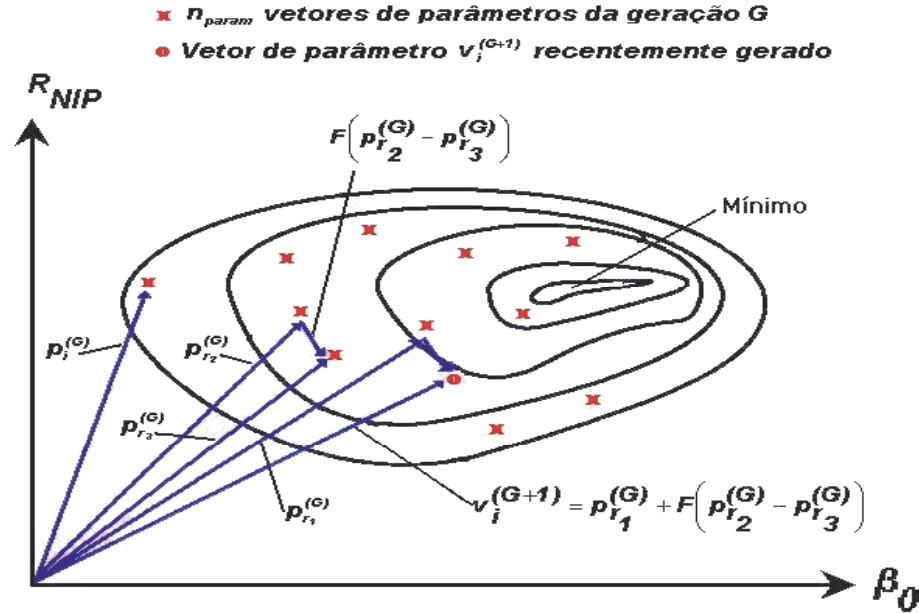


Figura 1C. Exemplo em duas dimensões de uma função objetivo mostrando suas linhas de contorno e o processo para geração de $v_i^{(G+1)}$ no esquema ED-1.

Para aumentar a diversidade do novo vetor de parâmetros, um cruzamento é introduzido. Para este fim, o vetor:

$$\mathbf{u}_i^{(G+1)} = (u_{oi}^{(G+1)}, u_{li}^{(G+1)}, \dots, u_{(n-1)i}^{(G+1)}) \quad (3C)$$

com

$$\mathbf{u}_{ji}^{(G+1)} = \begin{cases} \mathbf{v}_{j,i}^{(G+1)} & \text{para } j = |k|_n, |k+1|_n, \dots, |k+L-1|_n \\ p_{j,i}^{(G)} & \text{para todos os outros } j \in [0, n-1] \end{cases} \quad (4C)$$

é formado em que o símbolo $| \cdot |_n$ denota o módulo. O índice de partida k em (4C) é um inteiro escolhido aleatoriamente do intervalo $[0, n-1]$. O inteiro L , pertencente a $[1, n]$, denota o número de parâmetros que estão sendo trocados. A relação de probabilidade utilizada neste algoritmo é dada por $\Pr(L = v) = (CR)^{v-1}$, $v > 0$, onde $CR \in [0, 1]$ é a constante de cruzamento e constitui uma variável de controle no desenvolvimento do processo. A decisão aleatória para k e L é feita para cada vetor $u_i^{(G+1)}$ gerado recentemente. A Figura 2 ilustra o mecanismo do cruzamento DE.

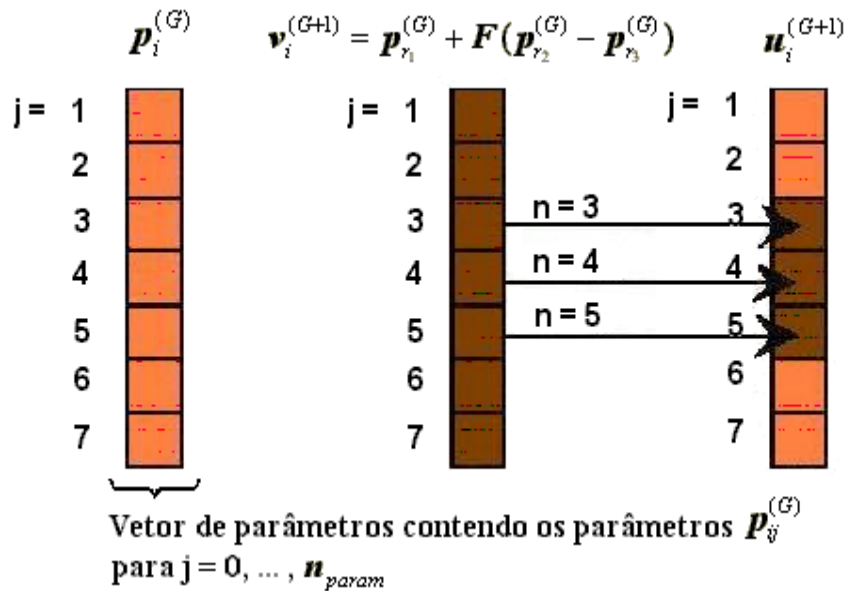


Figura 2C. Ilustração do cruzamento do processo para $n_{param} = 7$, $k = 2$ e $L = 3$.

Para decidir se deve ou não ser um membro da geração $G+1$, o novo vetor $\mathbf{u}_i^{(G+1)}$ é comparado com $\mathbf{p}_i^{(G)}$. Se o vetor $\mathbf{u}_i^{(G+1)}$ produzir um valor da função objetivo menor do que o valor produzido por $\mathbf{p}_i^{(G)}$, então $\mathbf{p}_i^{(G+1)}$ é substituído por $\mathbf{u}_i^{(G+1)}$. Caso contrário o vetor antigo $\mathbf{p}_i^{(G)}$ é conservado.

No DE a capacidade de achar o mínimo global e uma rápida taxa de convergência é muito sensível às variáveis de controle NP , F e CR . Algumas regras para a escolha destes parâmetros são dadas abaixo:

- Tamanho da população NP : Este valor depende do custo da função objetivo e dos outros parâmetros do algoritmo. Uma grande população tem maior probabilidade de achar o mínimo global para funções multimodais. Utilizamos aqui $NP = 50$.
- Fator de amplificação F : F não pode ser menor do que um certo valor que impede a convergência prematura. Para o problema de otimização do método SRC uma boa escolha é $F = 0.3$.
- Constante de Cruzamento CR : Para um valor alto de CR tem-se uma convergência rápida. Entretanto, a partir de um certo valor em diante a velocidade de convergência pode decrescer ou a população pode convergir prematuramente. Este valor depende do custo da função

objetivo, uma boa escolha é um valor entre $CR = 0.3$ e $CR = 0.9$. Neste trabalho utilizamos $CR = 0.8$

D - ALGORITMO CONTROLLED RANDOM SEARCH

Descreveremos aqui o processo *Controlled Random Search* (CRS) para otimização global. O algoritmo CRS pertence à classe de algoritmos Monte Carlo. É um processo otimização que não requer que a função objeto seja diferenciável ou que as variáveis sejam contínuas, e é aplicado a otimização restrita. As principais características do algoritmo são descritas abaixo.

Seja f uma função de n variáveis, um domínio de busca inicial D é definido especificando os limites de cada variável, isto é, $a_i \leq p_i \leq b_i$ para $i = 1, 2, \dots, n$ sendo a_i e b_i os limites inferiores e superiores respectivamente. Um conjunto de N pontos testes são escolhidos aleatoriamente de D , conforme as restrições adicionais, onde $N \gg n$. A função é avaliada em cada ponto e a posição e o valor da função correspondente a cada ponto são armazenados em uma matriz A .

Do conjunto de N pontos são escolhidos os pontos p_L que tem o menor valor da função objetivo (f_L) e p_M que tem o maior valor da função objetivo (f_M). Em cada iteração, um novo ponto teste p' é selecionado aleatoriamente de um certo subconjunto de pontos testes pertencentes aos N pontos. Desde que a posição de p' satisfaça as restrições, a função é avaliada em p' e o valor da função $f(p')$ é comparado com f_M . Se $f(p') < f_M$, então p_M é trocado por p' na matriz A . Caso contrário, se $f(p') > f_M$, então o teste é descartado e um novo ponto é escolhido. Ou seja, o conjunto atual de N pontos armazenados tende a agrupar ao redor do mínimo onde estão os valores de f menores do que f_M , isto é, o algoritmo reduz sua busca para a região promissora.

O método de otimização CRS proposto em Price (1977), faz um compromisso razoável entre o conflito de exigências da busca e da convergência, pela definição do conjunto de pontos de avaliação em termos da configuração dos N pontos atualmente armazenados. Em cada iteração, são escolhidos aleatoriamente $n+1$ pontos de N , isto é, R_1, R_2, \dots, R_{n+1} . O ponto R_1 não é escolhido aleatoriamente e sempre assume o valor do ponto p_L que tem o menor valor da função (f_L). Estes pontos constituem um *simplex* no espaço n -dimensional, onde o ponto R_{n+1} é o pólo ou denominado vértice do *simplex*, e o próximo ponto teste P , ponto imagem do pólo com respeito ao centróide G dos restantes n pontos e definido pela seguinte relação:

$$\bar{P} = 2\bar{G} - \bar{R}_{n+1}, \quad (1D)$$

onde \bar{P} , \bar{G} , \bar{R}_{n+1} são vetores posição no espaço n -dimensional, dos pontos testes, centróide e pólo do *simplex*, respectivamente.

Conservando o princípio essencial do CRS (i.e., a definição do próximo ponto teste em relação a um *simplex* escolhido aleatoriamente de uma configuração de pontos armazenados), os detalhes do procedimento podem ser modificados em uma variedade de caminhos. Usando pontos testes gerados por 1D, por exemplo, o processo CRS é uma busca altamente orientada, mas para algumas aplicações a convergência é inaceitavelmente lenta. Muitas variantes têm sido exploradas tentando aumentar a convergência sem reduzir significativamente a capacidade de busca global. Uma outra versão do CRS que foi desenvolvida e freqüentemente tem mostrado ser superior a versão anterior descrita é a chamada de CRS2 sendo que a anterior é referida como CRS1.

O algoritmo CRS2 difere do CRS1 apenas na escolha do ponto do *simplex* R_1 que não é mais escolhido aleatoriamente, sendo então tomado sempre o ponto p_L que tem o menor valor da função dos N pontos armazenados. A Figura 1D mostra o fluxograma das etapas do algoritmo CRS2.

No CRS1, $n+1$ pontos discretos são escolhidos aleatoriamente de N ; e por causa da escolha arbitraria de R_{n+1} como o pólo do *simplex* (os $n+1$ pontos podem ser gerados em uma única ordem aleatória), o número total dos pontos testes é $(n+1) \cdot \binom{N}{n+1}$. No CRS2, entretanto, como R_1 é sempre o ponto p_L , os n pontos são escolhidos aleatoriamente de $N-1$ pontos. Além disso, p_L nunca pode ser o pólo do *simplex*. Assim, o número dos potenciais pontos testes em CRS2 é $n \cdot \binom{N-1}{n}$, e estes pontos forma um subconjunto dos correspondentes pontos teste do CRS1.

Como uma simples ilustração do princípio CRS2, a Figura 2D mostra uma configuração de seis pontos em algum estagio da busca do mínimo global de uma função de duas variáveis. Com $N = 6$ e $n = 2$, o número de pontos CRS1 primários gerados por essa configuração é

$$3 \cdot \binom{6}{3} = 60 \quad (2D)$$

Dado que o ponto p_L , identificado na Figura 2D, tem o menor valor da função da configuração, o conjunto teste correspondente para CRS2 é o subconjunto dos

$$2 \cdot \binom{5}{2} = 20 \quad (3D)$$

pontos testes indicados no diagrama. Na prática, N pode ser escolhido maior do que 6 (tipicamente, $N = 30$ para buscas bidimensionais usando CRS2).

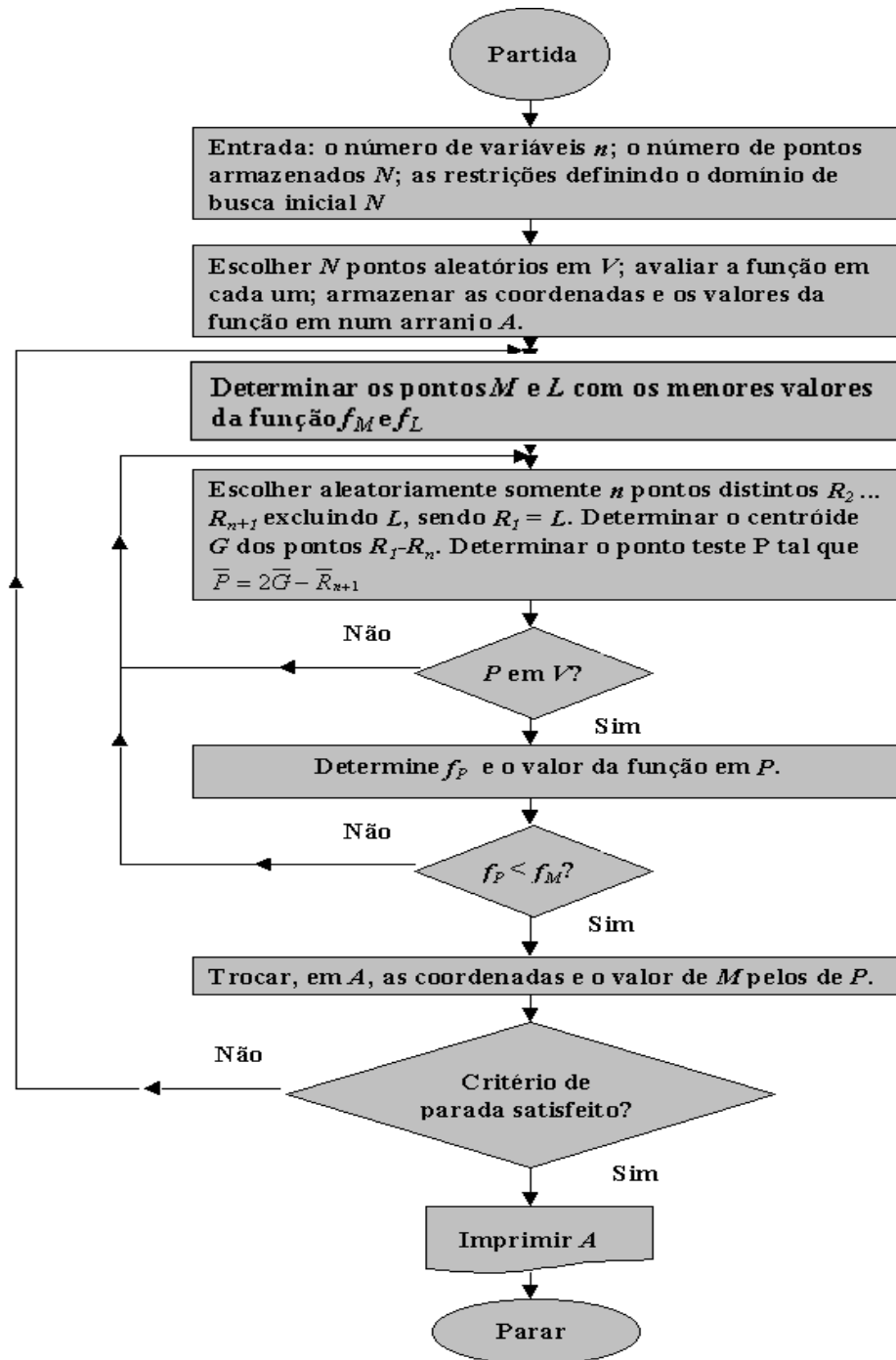


Figura 1. Fluxograma para o algoritmo CRS2

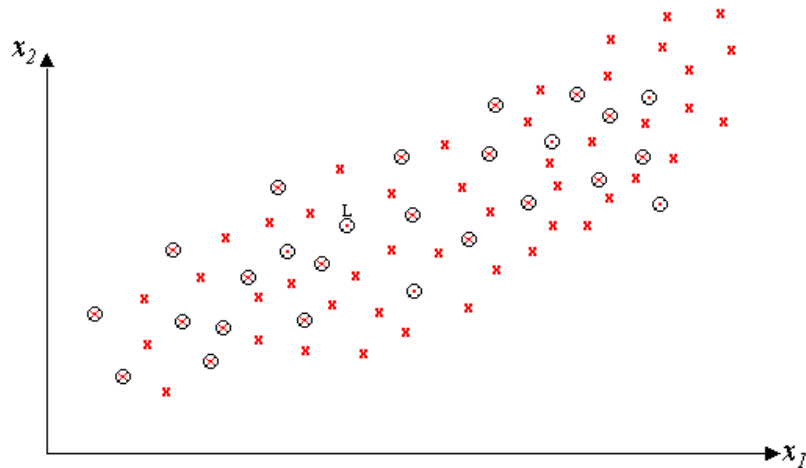


Figura 2. Uma ilustração do algoritmo gerando pontos testes em 2D.

- configuração de seis pontos armazenados
- L tem o menor valor da função
- x conjunto de 60 possíveis pontos testes para CRS1
- ⊙ subconjunto de 20 pontos testes para CRS2

Aparentemente o conjunto teste do CRS2 tende a influenciar a busca para a região ao redor de p_L . Se p_L está na vizinhança do mínimo global, então claramente a consequência desta influencia está numa convergência mais rápida do que a atingida no CRS1. Cada vez que um novo “melhor” ponto é gerado pelo processo, o conjunto teste do CRS2 é consequentemente mudado. O perigo, como em todo processo de busca global, é que a convergência da busca não escape do mínimo local e falhe em achar o mínimo global.

D.1 - A ESCOLHA DE N

Certamente que um grande valor de N aumenta a probabilidade de se encontrar o mínimo global. Por outro lado um grande valor de N exige maior esforço computacional e torna a convergência do algoritmo lenta. A escolha apropriada para N depende de várias experiências. Estudos anteriores mostram que uma regra adotada é fazer

$$N = 25.n \quad (4D)$$

a menos que exista uma boa razão para outra escolha. Para muitos problemas o mínimo global pode ser localizado com um pequeno valor de N . Em algumas circunstâncias falha utilizar o valor recomendado de N resultando na convergência para um mínimo local ao invés do mínimo

global. Experiências com o CRS2, entretanto, sugere que este processo suportará um pequeno valor de N . A regra recomendada para o CRS2 é

$$N = 10.(n+1). \quad (5D)$$

No entanto, um valor adequado de N pode não ser obtido com as anteriores regras recomendadas, neste caso pode-se determinar um valor específico para o problema em particular a partir de vários testes.