

## Modelagem da estrutura crustal média em Goiás com algoritmo genético

Afonso E.V. Lopes\*, Marcelo Assumpção, Jose R. Barbosa, IAG-USP, Brasil

Copyright 2003, SBGF - Sociedade Brasileira de Geofísica

This paper was prepared for presentation at the 8<sup>th</sup> International Congress of The Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 14-18 September 2003.

Contents of this paper were reviewed by The Technical Committee of The 8<sup>th</sup> International Congress of The Brazilian Geophysical Society and does not necessarily represents any position of the SBGF, its officers or members. Electronic reproduction, or storage of any part of this paper for commercial purposes without the written consent of The Brazilian Geophysical Society is prohibited.

### Abstract

Genetic Algorithms (GA) have been playing an increasingly important role on geophysical problem solving. GA was used in the inversion of hypocenter determination residues of seven regional seismic events in an structural model of the crust and upper mantle at Goiás region. The determination of the hypocenters was made with the arrival times of P and S waves of the seven events, according to crustal models evaluated by a computational implementation of GA. A brief discussion is made about the advantages and disadvantages of a blind search method (GA) compared with a full search in controlled space. These two methods present equivalent results, which show a 38 km thick crust, with P wave velocity of 8.2 km in the upper mantle and Vp/Vs ratio of 1.75. These values are consistent with a Pre-Cambrian shield region.

### Introdução

A Faixa Sísmica Goiás-Tocantins (Berrocal et al. 1984; Assumpção et al. 1986) é uma notável feição sismotectônica dentro do domínio geotectônico da Província Tocantins (Almeida 1977). A província Tocantins abrange quase totalmente os estados de Goiás, Tocantins e o Distrito Federal e parte de outros estados. É delimitada a NW pelo Cráton Amazônico, a sudeste pelo Cráton do São Francisco, a norte pela Bacia do Parnaíba e a sul pela Bacia do Paraná (Fuck et al. 1994).

Na literatura encontram-se diversos estudos direcionados a certos sismos da Faixa Sísmica Goiás-Tocantins como: sismos do município de Araguaiana-MT (1996-1997) por Veloso et al. 1997; o sismo de Araquapaz-GO (14/01/1986) por Assumpção et al.; o sismo de Brasília (20/11/2000) por Marza et al. 2001 e outros. Todos estes estudos adotam um modelo crustal médio a-priori sem testá-lo, o qual pode não ser o modelo que melhor representa a região. Logo, a caracterização de um modelo crustal médio para a região pode melhorar as determinações hipocentrais, dando informações mais precisas para estudos da sismicidade e caracterização de mecanismos focais na região.

Utilizamos 95 leituras dos tempos de chegada de ondas P e S de sete sismos regionais (figura 1) e determinamos seus hipocentros com o programa HYPO71 (Lee & Lahr

1978) para diferentes modelos crustais. Os melhores modelos foram selecionados por um Algoritmo Genético implementado pelos autores no programa GALU, e por um método de busca em espaço controlado (Método da Tentativa e Erro com uma varredura uniforme no domínio dos parâmetros a serem testados).

Tabela 1: Dados dos sismos utilizados na modelagem

Ano	DiaJ	Hora	Lat (S)	Lon (W)	Prof. (km)	Mag. (Mr)	
1	2002	238	02:29:20	15.08	46.86	3.5	3.3
2	2001	306	07:27:31	14.33	50.87	2.2	2.3
3	2001	300	18:42:08	15.10	51.27	0.4	2.8
4	2001	246	20:22:16	16.51	49.54	0.0*	2.2
5	2001	075	19:07:10	14.24	48.90	0.0*	2.2
6	2000	325	09:36:34	16.08	47.82	1.0	3.7
7	2000	257	07:39:47	11.66	49.79	0.0*	3.7

\*Profundidade Focal Fixada na modelagem

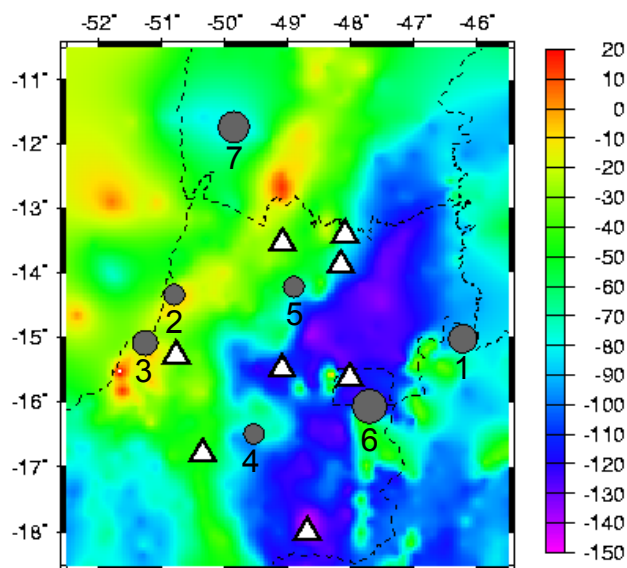


Figura 1: Epicentros dos sismos utilizados na modelagem crustal (círculos escuros), estações sismográficas (triângulos) e índices dos eventos referidos na tabela 1. A Anomalia Bouguer, representada em *mGal* pela paleta de cores, e dados de Função do Receptor (França 2003) nos mostram que a compensação isostática em Goiás se dá principalmente pelo mecanismo de Airy (variação da espessura crustal).

**Algoritmos Genéticos**

Os Algoritmos Genéticos (GA) são métodos estocásticos de busca cega de soluções quase ótimas, baseados nas leis de evolução natural propostas por Darwin, onde as possíveis respostas geradas para um problema são vistas como indivíduos que competirão entre si pela oportunidade de se reproduzir.

O termo *genético* diz respeito à maneira como as possíveis soluções para o problema são codificadas em genomas, uma estrutura composta por uma cadeia finita de elementos, os genes.

Um Algoritmo Genético compreende um conjunto de elementos individuais (a população de modelos) e um conjunto de operadores (reprodução, seleção, mutação, etc) inspirados na Biologia para produzir um novo conjunto de modelos melhores a cada iteração, obtendo uma maior adaptação da população em geral. Deste modo os GA's emulam a natureza em um de seus mais fortes atributos, a adaptabilidade.

**Parâmetros de Busca**

Em um primeiro estudo da estrutura crustal na região de Goiás com o Método da Tentativa e Erro, avaliamos mais de 1.3 milhão de modelos e chegamos aos resultados apresentados na tabela 3. Devido à lentidão desse processo de modelagem, optamos em utilizar um GA como um método de busca mais eficiente.

Com o GA pesquisamos um espaço com cerca de 2 milhões de modelos possíveis com a variação dos valores da velocidade da onda P na crosta superior (V1), crosta inferior (V2), manto superior (V3), espessura da crosta superior (Z1), espessura crustal (Zc) e razão VP/Vs. O espaço de busca está definido na tabela 2.

Tabela 2: Parâmetros de Busca utilizados no GA

Parâmetro	Limites de Busca	Intervalo de Busca	Unidade
V1	5.80 - 6.50	0.10	km/s
V2	6.60 - 7.30	0.10	km/s
V3	7.80 - 8.50	0.10	km/s
Z1	10.0 - 25.0	1.00	km
Zc	30.0 - 45.0	1.00	km
VP/Vs	1.65 - 1.80	0.01	

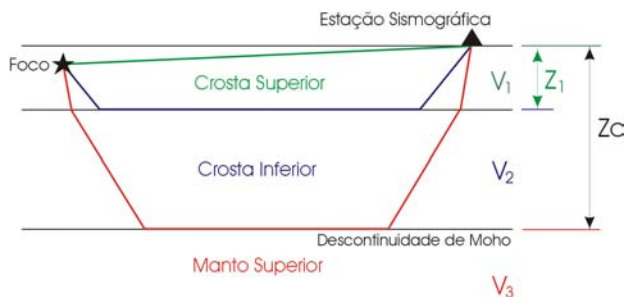


Figura 2: Esquema dos parâmetros a serem modelados

O fator mais importante nesta modelagem com o GA é possibilitar a análise de um conjunto maior de modelos em 2% do tempo usado com o Método da Tentativa e Erro.

**Implementação de um Algoritmo Genético - GALU**

Os Algoritmos Genéticos operam sobre uma população, fazendo com que esta evolua de acordo com uma função de adaptação (chamada de função objetivo). Nossa função objetivo,  $f(V_1, V_2, V_3, Z_1, Z_c, V_p/V_s)$ , é dada pela média quadrática dos resíduos dos tempos de chegada das ondas sísmicas utilizadas nas determinações hipocentrais dos sete eventos (equação 1).

$$f(V_1, V_2, V_3, Z_1, Z_c, V_p/V_s) = \sqrt{\frac{\sum_{i=1}^7 r_{ms}^2 N_i}{\sum_{i=1}^7 N_i}} = RMS_{Geral}$$

(1)

Onde:  $r_{msi}$  e  $N_i$  são respectivamente o resíduo quadrático das leituras e o número de leituras utilizadas na determinação hipocentral do evento  $i$ .

Além da função objetivo, um fator diretamente relacionado ao problema é a definição do genoma a ser utilizado. Usamos 3 bits na representação de cada velocidade e 4 bits na representação de cada um dos parâmetros restantes, perfazendo um total de 21 bits (figura 3). Ou seja, cada parâmetro de velocidade terá  $2^3$  (8) valores possíveis dentro do intervalo de busca.

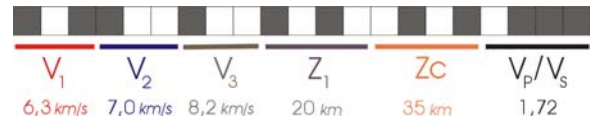


Figura 3: Representação de um modelo crustal na forma de um genoma. Enquanto os genes de cor escura possuem valor 1, os genes não preenchidos possuem valor 0 na representação binária. Como exemplo, a informação codificada neste conjunto de genes está expressa abaixo dos respectivos genes.

Com exceção da função objetivo e da estrutura do genoma, os GA's possuem um fluxograma bem definido para a evolução da população (figura 4). Tal roteiro evolucionário começa com a inicialização da primeira população, criada com auxílio de funções aleatórias que garantem uma maior *biodiversidade*, fundamental para garantir uma boa abrangência do espaço de pesquisa. Optamos por criar nossa primeira população com uma Inicialização Randômica Uniforme (IRU), onde cada gene do indivíduo recebe como valor um elemento do conjunto de alelos, sorteado de forma aleatoriamente uniforme.

O primeiro passo da seleção de modelos ocorre na avaliação, onde cada modelo é testado com a função objetivo, visando avaliar o seu grau de adaptação.

O grau de adaptação possui uma relação direta com a probabilidade de perpetuação das características do modelo e sua chance de sobrevivência, e é dado por uma espécie de normalização dos dados obtidos com a função objetivo, chamada de *Scaling*. A metodologia adotada no cálculo do *Scaling* (*Linear Scaling*) consiste em uma normalização dos resíduos entre 0 e 100 de modo linear (o menor resíduo vale 100 e o maior vale 0).

Na seleção propriamente dita, os indivíduos (modelos) são escolhidos para posterior cruzamento. Neste ponto, fazendo uso do grau de adequação de cada um, é

efetuado um sorteio onde os mais aptos possuem maior probabilidade de se reproduzirem. Optamos em fazer uma seleção por *ranking* (*Rank Selection*) devido a fácil implementação desta. Na seleção por *ranking* os modelos são ordenados pelo grau de adequação e em seguida há um sorteio de índices, os quais representam os casais a serem cruzados (pai e mãe). Uma vez selecionados os indivíduos, estes passam com uma probabilidade pré-estabelecida (taxa de *crossover*) pelo processo de cruzamento (*crossover*), onde partes dos genes dos pais são combinadas para geração de filhos.



Figura 4: Fluxograma do programa GALU

Como na maior parte das rotinas de um GA, a reprodução possui um caráter aleatório controlado, isso porque as posições de *crossover* são delimitadas por um operador de cruzamento. O funcionamento dos operadores de cruzamento pode ser representado como uma seleção por máscara, a qual é representada por um vetor cujos elementos possam assumir valores binários e que possuam um comprimento igual ao dos cromossomos a serem combinados. Seu uso pelo operador se deu segundo o seguinte algoritmo:

```

se máscara=0
  então
    irmã←mãe;
    irmão←pai;
  senão
    irmã←pai;
    irmão←mãe;
    
```

O GALU foi implementado com um cruzamento segmentado (SX) que sorteia o número um *n* de pontos de corte a cada execução e em seguida sorteia a posição de cada

um dos cortes. Sua máscara é representada por *n* mudanças em seqüência de zeros e uns (figura 5).

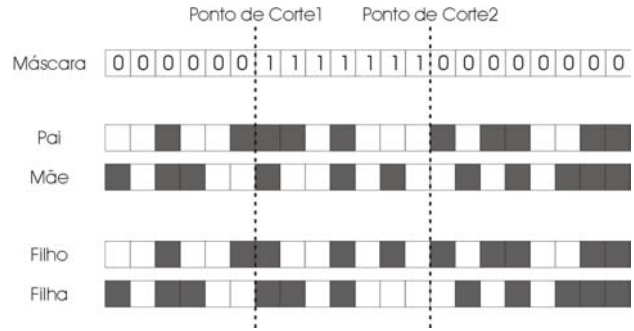


Figura 5: Exemplificação do uso de máscara para um operador de cruzamento SX com *n* =2.

A mutação opera sobre os indivíduos resultantes do processo de cruzamento e com uma probabilidade pré-determinada efetua algum tipo de alteração em sua estrutura. A implementação desse tipo de mecanismo é extremamente simples e consiste no sorteio do número de genes e dos genes que sofrerão mutação. Após o cruzamento e a mutação, o tamanho da nova população é o dobro do tamanho da população inicial. Neste ponto do algoritmo, uma análise dos *scalings* de cada indivíduo diz qual a probabilidade do mesmo sobreviver e ser inserido na próxima geração de modelos. A probabilidade (em %) do modelo *i* ser selecionado para a geração seguinte é dada por:

$$P_i = 100 \frac{Scaling_i}{\sum_{j=1}^{nind} Scaling_j} \tag{2}$$

Todas as rotinas comentadas são executadas pelo programa GALU até atingir o número de gerações pré-determinado.

**Resultados da Inversão com o GA**

Como aconselhado na literatura, adequamos nosso GA a nossos dados com auxílio de alguns testes e chegamos aos seguintes valores de entrada:

- Número de Gerações: 50
- Número de Indivíduos: 50
- Taxa de Mutação: 5%
- Taxa de Crossover: 95%

Um número pequeno de indivíduos impossibilita uma boa convergência do GA (não-convergência ou convergência prematura), enquanto que um número muito grande de indivíduos exige um número elevado de gerações para que a população evolua. Ponderando isso com auxílio de testes, chegamos a uma população de modelos com 50 indivíduos. Para esse tamanho de população, o programa GALU obtém convergência antes da vigésima geração, o que possibilitou fixarmos o número de gerações em 50. Em problemas altamente combinatorios, um operador cego de cruzamento pode facilmente gerar filhos menos

adequados a partir dos cromossomos pais, no fenômeno conhecido como epístase. Por isso, optamos em delimitar os ponto de corte pelos bits que delimitam informações distintas.

Os operadores genéticos de mutação e cruzamento são responsáveis por todas as transformações sofridas pela população, mas possuem funções bastante distintas no que diz respeito a seu impacto na evolução. A mutação é fator fundamental para garantir a biodiversidade, assegurando que o espaço de busca será explorado em uma parte significativa de sua extensão e é fundamental para evitar a convergência prematura, que ocorre quando a população se estabiliza com uma média de adaptação pouco adequada por causa da pressão evolutiva e baixa biodiversidade. Isto geralmente ocorre com o surgimento de um super-indivíduo que domina o processo seletivo e, uma vez incapaz de gerar filhos melhores, transmite suas características por toda população. Relevando a importância da mutação e lembrando que a ela insere um caráter altamente aleatório, executamos alguns testes e constatamos que a combinação de 5% para a taxa de mutação com 95% para a taxa de crossover é ideal para nosso conjunto de dados.

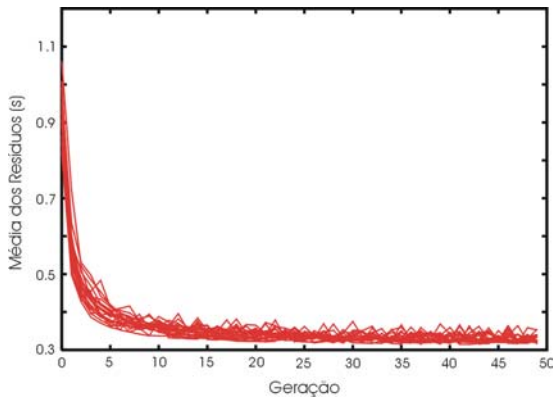


Figura 6: Média dos resíduos de cada geração

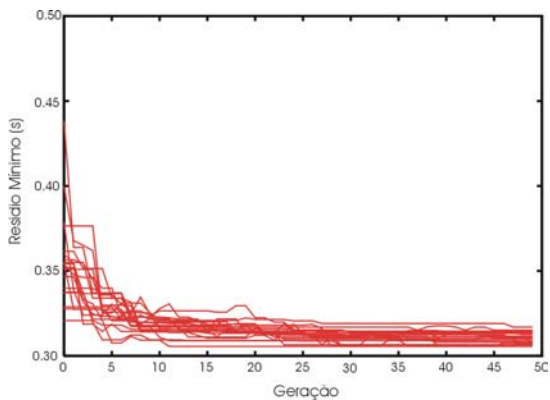


Figura 7: Menor resíduo encontrado em cada geração

Como a cada execução do programa a população evolui de maneira diferente devido ao caráter aleatório dos GA's, executamos o programa 20 vezes e analisamos os resultados desses vinte processamentos. Nas figuras 6 e 7 apresentamos o comportamento geral da população a cada geração, para cada execução do programa GALU.

A figura 7 mostra que o melhor modelo de cada geração não passa necessariamente à geração seguinte. Ou seja, não estamos usando "elitismo". O elitismo pode acelerar a convergência mas com risco maior de se cair em um mínimo local.

A figura 8 mostra os melhores modelos obtidos na inversão com o programa GALU, sendo seus resultados compatíveis com os resultados do Método da Tentativa e Erro (tabela 3), com diferença apenas no tempo de processamento. Enquanto que o Método da Tentativa e Erro testa 1.3 milhão de modelos para caracterização do conjunto de modelos possíveis, o GA utiliza apenas 50 mil modelos. Em termos de tempo de execução, o GA foi 60 vezes mais eficiente.

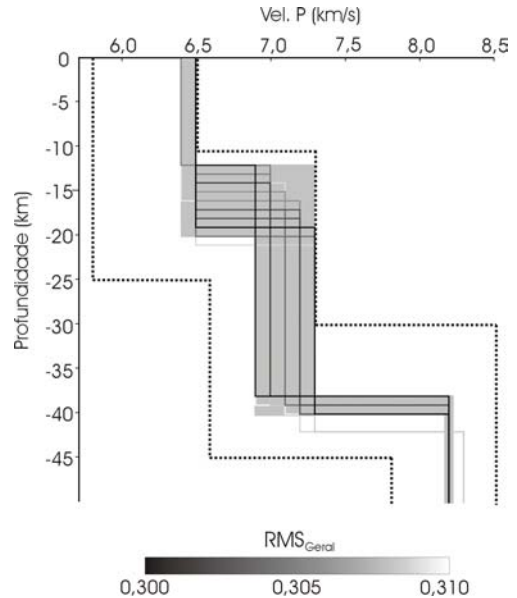


Figura 8: As linhas de diferentes tons de cinza, representam os melhores modelos crustais segundo seus valores de  $f$ , a área selecionada pela cor cinza claro delimita a região com o maior número de modelos e mais representativos da região (modelos mais prováveis) e a linha pontilhada preta representa os limites do intervalo de busca.

**Conclusões**

Os parâmetros mais bem definidos pela inversão com o GA (espessura crustal entre 38 e 40km, velocidade da onda Pn igual a 8.2km/s e razão  $V_P/V_S=1.75$ ) são equivalentes aos obtidos com o Método da Tentativa e Erro (tabela 3). Encontramos valores altos para a velocidade da onda Pg ( $V_1=6.5km/s$ ) com as duas metodologias, mas um estudo preliminar com ondas de superfície nos mostrou valores mais baixos para esse parâmetro (próximo de 6.1km/s).

A boa definição de  $Z_c$  e  $V_3$  ocorre devido à grande amostragem da interface crosta-manto pelos percursos das ondas Pn, a baixa incerteza do valor de  $V_P/V_S$  se dá devido ao grande número de percursos de ondas P e S, e a delimitação mediana da descontinuidade de Conrad é dada pela baixa densidade de percurso de ondas P\* (figura 9). O valor anômalo encontrado para a velocidade



da onda Pg se dá pelo baixo número de percursos, sua pequena influência nos resíduos das determinações hipocentrais, e erro nos hipocentros utilizados (um erro de alguns quilômetros no hipocentro pode gerar erros na ordem de décimos de  $km/s$  no valor de  $V_1$ ).

Tabela 3: Modelo de crosta calculado com o Método da Tentativa e Erro (MTE) e com o GA

Parâmetro	MTE	GA	Unidade
$V_1$	$6.39 \pm 0.03$	6.4 a 6.5	$km/s$
$V_2$	$6.92 \pm 0.10$	6.9 a 7.3	$km/s$
$V_3$	$8.18 \pm 0.03$	8.2	$km/s$
$Z_1$	$11.5 \pm 1.5$	12 a 20	km
$Z_c$	$37.0 \pm 1.6$	38 a 40	km
$V_P/V_S$	$1.75 \pm 0.01$	1.75	Adimensional

Os resultados obtidos são consistentes com um Escudo Pré-Cambriano, com exceção da velocidade de propagação da onda P na crosta superior ( $V_1=6.4km/s$ ).

A espessura crustal obtida é condizente com a média das espessuras calculadas com Função do Receptor por França (média igual a 37.5km) e os valores de  $V_3$  e  $V_P/V_S$  são bem definidos pelos dois métodos, sendo que o valor de  $V_P/V_S$  modelado é igual ao obtido com o Diagrama de Wadati ( $V_P/V_S=1.750 \pm 0.009$ ).

**Agradecimentos**

Ao Dr. George S. França pela grande participação neste trabalho, tanto na parte técnica (auxílio com textos em LaTeX), como no compartilhamento de dados de sua pesquisa com Função do Receptor na região de Goiás, e ao Prof. Eder Cassola Molina pelos dados de gravimetria.

Ao Observatório Sismológico da Universidade de Brasília (OBSIS) pelo compartilhamento dos dados de suas estações sismográficas do reservatório de Serra da Mesa-GO (SM), de Cana Brava-GO (CAN1), de Brasília (BDFB) e do arranjo Sismográfico de Brasília. A Daniel Caixeta, por toda a atenção, envio dos sismogramas e leituras das estações sismográficas da UnB.

Desejamos expressar aqui nosso agradecimento a FAPESP pelo financiamento de nossas pesquisas e desenvolvimento da ciência no país.

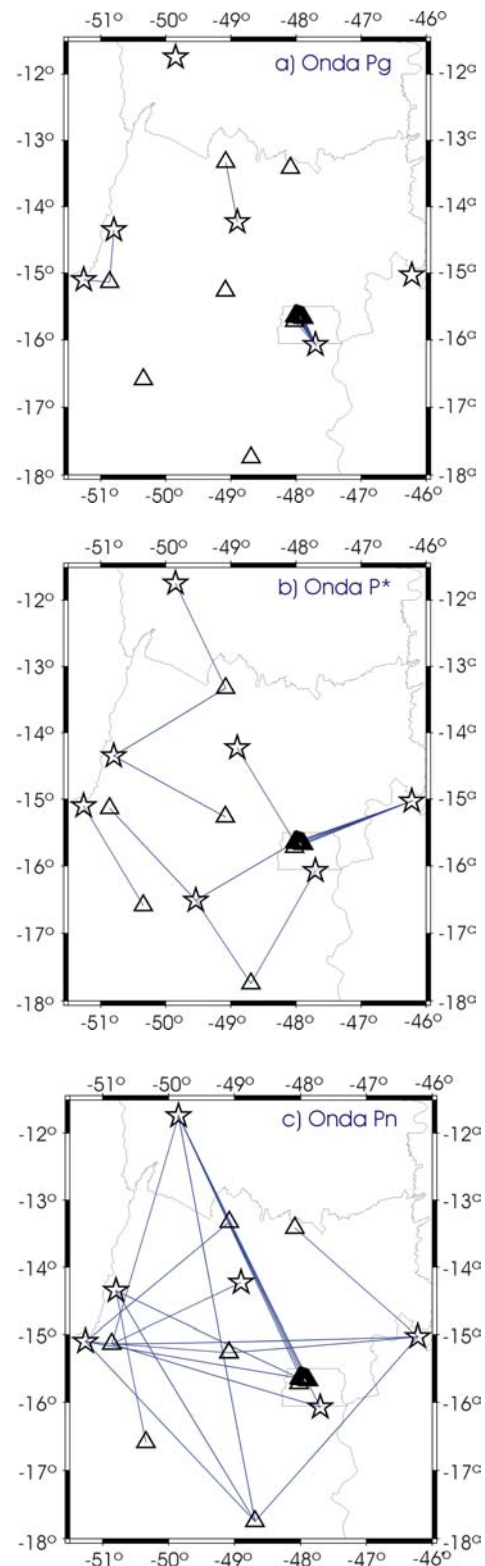


Figura 9: Projeção dos percursos das ondas a) Pg, b) P\* e c) Pn. A boa distribuição e quantidade de percursos das ondas Pn são os grandes responsáveis pela boa definição de  $Z_c$  e  $V_3$ .

**Referencias**

- Almeida, F.F.M. (1977) *O Cráton de São Francisco*, Revista Brasileira de Geociências, volume 7, São Paulo, pp.349-364
- Assumpção, M.S.; Lima T.M. & Tomás L.A.R. (1986) *O Sismo de Araguapaz de 14.01.1986 e o Lineamento Transbrasiliano*, 34º Cong. Bras. Geol., pp. 237-240
- Berrocal, J.; Assumpção, M; Anteza, R. et al. (1984) *Sismicidade do Brasil*, IAG/USP - CNEN, São Paulo, 320 pp.
- França, G.S (2003) *Estudo da Estrutura Crustal no Sudeste e Centro-Oeste do Brasil usando a Função do Receptor*, Tese de Doutorado, IAG-USP, 200pp.
- Fuck, R.A. (1994) *A Faixa Brasília e a Compartimentação Tectônica na Província Tocantins*, 4º Simpósio de Geologia do Centro-Oeste, Anais, Brasília, SBG, pp.184-187
- Lee, W.H.K. & Lahr, J.C. (1978) *HYP071 (Revised): A Computer Program for Determining Hypocenter, Magnitude and First Motion Pattern of Local Earthquakes*, USGS Open File Report 75-311, 448pp.
- Lopes, A.E.V & Assumpção, M.S. (2002) *Sismicidade e Estrutura Crustal em Goiás*, X Simpósio Internacional de Iniciação Científica da USP, XIV Salão de Iniciação Científica da UFRGS e VII Simpósio de Iniciação Científica do IAG-USP, São Paulo, CDROM
- Marangoni, Y.R. (1999) *Modelo Tectônico Compressivo para a Província Tocantins (Central) a partir de um Perfil Gravimétrico*, 7º simpósio Nacional de Estudos Tectônicos, Lençóis, BH, pp.107-109
- Marangoni, Y.R. (1994) *Modelo Crustal para o Norte de Goiás a partir de Dados Gravimétricos*, Tese de Doutorado, IAG-USP
- Marza, V; Barros, L; Chimpliganond, C; Caixeta,D. (2001) *Brasília (DF) Earthquake of Nov. 20, 2000: In Quest of Unravelling*, Congresso Internacional da Sociedade Brasileira de Geofísica, Salvados, CD-ROM
- Veloso, J.A.V.; Marza, V.I.; Carvalho, J.M. & Barros, L.V. (1997) *Recent Seismic Activity in the S-W Edge of the Mato Grosso-Goiás-Tocantins (MGT) Belt (Central Brazil)*, 5º Congresso Int. da Sociedade Brasileira de Geofísica, São Paulo, vol II, pp.964-967