Inversão Bayesiana 1D de Dados Geofísicos de Dipolo Magnético Vertical

Roberto Livy da Costa Madeira

Trabalho de Conclusão de Curso
Licenciatura em Matemática

Faculdade de Matemática
Campus Castanhal

2019

Inversão Bayesiana 1D de Dados Geofísicos de Dipolo Magnético Vertical

Graduação
05/05/2016–05/12/2019

Submissão 07/12/2019
Defesa 12/12/2019
Versão Final 18/12/2019

Universidade Federal do Pará
Campus de Castanhal
Faculdade de Matemática

Roberto Livy da Costa Madeira
llivybetica21@gmail.com
Licenciado em Matemática 2019
UFPA-Castanhal

Banca Examinadora

Prof. Valdelírio da Silva e Silva
Membro da Faculdade (Orientador)

Prof. Nildsen Fernando Lisbôa da Silva
Membro da Faculdade

Prof. Romário Silva Duarte
Membro Externo

Agradecimentos

Manifesto meus agradecimentos aos meus amigos, familiares e professores que estiveram comigo nesse período. Ressalto aqui alguns nomes dos quais tive ampla ajuda no processo de graduação. Agradeço ao meu pai Raimundo Roberto Madeira; minha mãe Dulcirene da Costa; minhas irmãs Emily Madeira e Milena Madeira. Uma grande ajuda profissional e financeira me veio das pessoas da empresa IGASAT à qual devo muito. Obtive grande ajuda no meio acadêmico de alunos e professores do campus; agradeço ao meu amigo e parceiro de classe washington; deixo também minha gratidão aos professores Fábio e Valdelírio Silva pelas valiosas orientações. Por fim, ao meu grande amigo Dr. Aaron Joslin, minha gratidão pelas grandes contribuições científicas e pela ajuda na elaboração do abstract deste documento.

Epígrafe

"Um dos paradóxos dolorosos do nosso tempo reside no fato de serem os estúpidos os que têm a certeza, equanto os que possuem imaginação e inteligência se debatem em dúvidas e indecisões."

Berthrand Russell

Resumo

Em métodos geofísicos busca-se estimar propriedades geoelétricas da subsuperfície, tais como as resistividades de camadas na superfície terrestre e suas propriedades geométricas. Usou-se no presente trabalho métodos *Bayesianos*. Foram propostos *Métodos de Monte Carlo* baseados em *Cadeias de Markov* (MCMC), Os quais são os métodos bayesianos mais usuais em inversão de parâmetros. Primeiramente utilizou-se o MCMC Metropolis-Hastings (MH) fixando um número de espessuras e estimando apenas as resistividades de cada uma; depois aplicou-se o *Reversible Jump MCMC* (RJ-MCMC) usando como vaiável a própria dimensão do modelo. O MH retornou valores aceitáveis de resistividade das camadas rasas, gerando imprecisão nas camadas intermediárias e profundas; o RJ-MCMC se mostrou mais adequado por se autoparametrizar, ajustando bem resistividade e espessura da primeira camada para pesquisa profunda, deixando incertezas nas camadas posteriores, e, para pesquisa rasa, ajustou corretamente o modelo de 3 camadas e menor tempo de computação, tendo um leve desvio dos valores reais em função das correlações entre o par de propriedades de cada camada (resistividade e espessura) que é mostrada nos resultados.

Palavras-chaves: RJ-MCMC, Resistividade, Cadeias de Markov, Monte Carlo

Abstract

Various geophysical methods are used to estimate subsurface geolectric properties, such as the resistivity of subsurface layers of the earth and their geometric properties. Bayesian methods were used in this study to estimate resistivity and thickness of subsurface horizons. Specifically, Monte Carlo Methods based in Markov Chains (MCMC) were proposed, which are the most common bayesian methods used in parameter inversions. First, MCMC Metropolis-Hastings (MH) were used to fix different thickness to each layer and then estimated the resistivity of each. Then, Reversible Jump MCMC (RJ-MCMC) was applied using the modelgenerated thicknesses and number of layers as variables. The MH model returned acceptable resistivity values of the shallow layers, generating imprecision in the intermediate and deep layers. The RJ-MCMC model was better suited for self-parameterizing and for faster computing time, generating uncertainty in the deeper layers. Our results show that for studies of shallow layers RJ-MCMC correctly adjusted the three-layer model with low deviation of real values as related to correlations between the pair of properties.

Keywords: RJ-MCMC, Resistivity, Markov chain, Monte Carlo

Lista de Figuras

3.1 Modelo teórico de um meio estratificado com resistividades de 100, 50 e 100 Ω.m respectivamente 19
3.2 Valor esperado para as resistividades, dados fixados, o número de camadas e espessuras . . . 20
3.3 Estimador ML, dados fixados, o número de camadas e espessuras . . . 20
3.4 Histogramas não normalizados das PDFs referentes às camadas com h fixo . . . 21
3.5 Distribuição a posteriori do modelo $p(n|\mathbf{d})$ . . . 21
3.6 Estimadores . . . 22
3.7 Modelo teórico de um meio estratificado com resistividades de 200, 50 e 100 Ω.m respectivamente 22
3.8 Distribuição a posteriori do modelo $p(n|\mathbf{d})$ . . . 22
3.9 Cadeia de Markov Marginal de $log_{10}\rho_1$ . . . 23
3.10 Amostras da 1º Camada . . . 23
3.11 Amostras da 2º Camada . . . 24
3.12 Estimador Valor Esperado . . . 24

Sumário

1 Introdução **8**

2 Metodologia de Inversão **10**
2.1 Regra de Bayes . . . 10
2.2 Problema inverso . . . 11
2.3 Prioris . . . 11
2.4 Função de verossimilhança . . . 11
2.5 O Algoritmo Metropolis-Hastings . . . 12
2.6 MCMC com Saltos Reversiveis . . . 14
2.7 Considerações usadas na implementação . . . 15
2.7.1 Prioris . . . 15
2.7.2 Probabilidade de aceitação e movimento entre os modelos . . . 16
2.7.3 Distribuição proposta e probabilidade de aceitação . . . 16
2.7.3.1 Movimento em dimensão fixa (update) . . . 16
2.7.3.2 Nascimento de camada (birth) . . . 16
2.7.3.3 Morte de camada (death) . . . 17
2.7.4 Considerações na Implementação . . . 17

3 Resultados **19**
3.1 Metropolis-Hastings (MH) na Investigação com o DMV . . . 19
3.2 MCMC com Saltos reversíveis (RJ-MCMC) na Investigação com o DMV . . . 20

4 Considerações finais **25**

Referências **27**

1

Introdução

Em geofísica o maior objetivo é obter informações da superfície investigada, sendo esta a subsuperfície abaixo dos receptores, como na investigação de contaminação de subsolo, ou a região ao redor dos sistema fonte-receptor que é usado por exemplo na perfilagem com ferramentas de indução de poços, que destina-se a saber a existência ou não de água ou hidrocarbonetos. Vários métodos são aplicados pela geofísica e a designação de cada um deles está associada ao tipo de campo físico observado. Se o campo gravitacional, que não é uniforme em toda superfície da Terra, é de interesse em se observar, então se tem os métodos gravimétricos; se ondas sísmicas são de objetivo medir, então se tem os métodos sísmicos; e no caso do campo eletromagnético ser de interesse, se tem então os métodos eletromagnéticos. A aplicação de um método ou outro é determinada mediante o conhecimento de que propriedades físicas na aŕea de investigação são mais relevantes. Por exemplo, se o intuito é de averiguar a existência de minérios, o campo eletromagnético traz sinais mais fortes que o sísmico, e, dependendo da profundidade e dimensão do alvo às vezes um método elétrico (cujo sinal é a diferença de potencial lida entre eletrodos enterrados) ou gravitacional também resgatam bons resultados.

Após a coleta dos dados uma aplicação de filtragem é realizada para minimizar ou reduzir os efeitos de ruídos, que são geológicos, ferramental ou de interferências externas. Depois dessa fase começa então o processo de inversão geofísica. Isso é um processo computacional usando em geral um ou vários métodos matemáticos que tentan inferir sobre as propriedade físicas da área investigada. Dentro dessas técnicas matemáticas existem as chamadas determinísticas e as probabilísticas. As determinísticas baseiam-se em procedimentos de minimização de um funcional de ajuste (diferença entre os dados observados e os retornados pelo processo de inversão computacional iniciada a partir de um modelo interpretativo informado pelo usuário), junto com técnicas de regularização matemática, que são métodos inseridos à minimização a fim de tornar o processo de inversão um problema com solução única e estável.

Os métodos de inversão probabilísticos mais usados são os bayesianos; livres de derivadas, buscam ajustar os dados observados a um conjunto de parâmetros usando *funções densidade de probabilidade*. Estes métodos consideram que todas as variáveis do problema são aleatórias por levarem em conta, na modelagem, que os erros de medição ou ruídos geológicos são variáveis aleatórias aditivas nas amostras coletadas. Os métodos de inversão bayesiana usuais são os *Métodos de Monte Carlo via Cadeia de Markov(MCMC)*.

Em qualquer que seja o método de inversão é sempre necessário fazer uso da modelagem direta, isto é, das respostas computacionais do método geofísico aplicado na investigação considerando-se as propriedades do modelo. A modelagm direta é calculada em cada iteração para poder suas respostas serem comparadas com as dos dados observados. Mediante a comparação entre esses dois tipos de dados é que são realizadas decisão de tarefas a serem realizadas computacionalmente. No caso da inversão bayesiana utilizada neste trabalho, as propriedades do modelo serão as resistividades e as espessuras das camada estratificadas hori-

zontalmente, e a fonte de campo eletromagnético será de um dipolo magnético vertical (DMV). O programa da modelagem direta usado é o de Hanssens et al. (2019), e sobre ele os dados retornados foram apenas da componentes magnética vertical H_z, normalizados pelo valor dessa componente se o meio fosse infinito, isotrópico e ilimitado, $H_z^o = \frac{-m_z}{4\pi r^3}$. Nesta expressão, m_z é o momento magnetico do dipolo e r a distância do ponto de medida até a posição do dipolo.

No capítulo 2 introduzimos alguns conceitos básicos da modelagem de inversão probabilística, bem como, a descrição dos algoritmos usados, e os detalhes requeridos para implementá-los; O capítulo 3 mostra os resultados das aplicações do código de dimensão fixa e dimensão variável; por fim, nas considerações finais são mostradas as análises finais sobre a pesquisa e os próximos passos para investigações futuras.

2

Metodologia de Inversão

2.1 Regra de Bayes

Um pequeno texto deve ser dado para mencionar a um resultado muito importante da teoria das probabilidades. O teorema (ou regra) de Bayes nada mais é do que uma consequência simples da definição de probabilidade condicional. Sejam um espaço amostral S_e e dois eventos A e B associados a um experimento e; denotamos por $p(B|A)$ à probabilidade condicionada do evento B quando A tiver ocorrido. Quando calculamos $p(B)$, estamos calculando o quão provável será estarmos em B, e ao calcularmos $p(B|A)$ estamos calculando a probabilidade de estarmos em B, sabendo que devemos estar em A; ou seja, na ocorrência de A, ao calcularmos $p(B|A)$ estamos calculando a probabilidade de B com o espaço amostral reduzido de S_e para A.
O modelo de probabilidade condicional tem a seguinte expressão:

$$p(B|A) = \frac{p(A \cap B)}{p(A)}$$

Analogamente:

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Isolando $p(A \cap B)$ da segunda equação e substituindo na primeira, chegamos ao *teorema de Bayes*:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)} \tag{2.1}$$

O teorema de Bayes também pode ser escrito através de *funções Densidade de Probabilidade (PDF)* ou também *Funções de Massa de Probabilidade (PMF)*, as quais são uma classe de funções matemáticas que têm como argumentos *variáveis aleatórias* e que tratam de incertezas relacionadas a um experimento, ou seja, sempre que um experimento produz um resultado aleatório toda vez que é realizado, sob as mesmas condições, tal experimento tem modelagem probabilística. Caso tal experimento citado anteriormente produza valores iguais todas as vezes em que executado nas mesmas condições n vezes, tem variável determinística e pode ser representado por funções não *estocásticas*. Alguns modelos determinísticos frequentemente são modelados probabilisticamente por serem altamente sensíveis às condições iniciais (caóticos), como, por exemplo, o lancamento de uma moeda equilibrada (não viciada).

2.2 Problema inverso

Em um problema inverso estamos diante da tentativa de fazer estimativas de um conjunto de n parâmetros $\mathbf{m} = \{m_1, m_2, \ldots, m_n\}$ não determinados diretamente do conjunto de k dados $\mathbf{d} = \{d_1, d_2, \ldots, d_k\}$ observados, mas que são relacionadas teoricamente por leis físicas por exemplo. Matematicamente podemos escrever:

$$\mathbf{d} = g(\mathbf{m}) + \boldsymbol{\epsilon} \tag{2.2}$$

onde g representa o operador do problema direto e $\boldsymbol{\epsilon}$ representa as incertezas, os ruídos advindos de qualquer sistema de medição e/ou inerentes do ambiente (modelo) investigado. Considerando que o conjunto $\mathbf{d}$ é composto por variáveis aleatórias, ou seja, cada d_i é uma observação de alguma distribuição de probabilidade, pode-se usar técnicas da álgebra linear para combiná-las através de um *vetor aleatório* como *função de verossimilhança* dos parâmetros $\mathbf{m}$ dado $\mathbf{d}$: $p(\mathbf{d}|\mathbf{m})$. Outra PDF importante é distribuição a priori, que representa o conhecimento probabilístico que se tem dos parâmetros antes de se obter algum dado observado, denotada por $p(\mathbf{m})$. Usando a regra de Bayes é possível combinar a função de verossimilança com a distribuição a priori para obter a chamada distribuição (*a posteriori*) dos parâmetros dado o conjunto de dados:

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})} \tag{2.3}$$

O termo $p(\mathbf{d})$ em (2.3) não depende dos parâmetros a serem estimados, desempenhando o papel de constante de normalização, e é chamado distribuição preditiva de $\mathbf{d}$, tendo o papel de garantir que a distribuição a posteriori dos parâmetros integre para a unidade [1]. Para os métodos usados neste trabalho não será requerido esse termo e usaremos a relação de Bayes não normalizada:

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}). \tag{2.4}$$

2.3 Prioris

Basicamente usaremos dois tipos de distribuições a priori na inversão. As distribuições uniformes discretas e contínuas denotadas por $U(a, b)$, onde a e b delimitam o suporte da PDF, garantindo sua normalidade, e serão diferidas no texto na especificação do suporte de suas variáveis aleatórias.

2.4 Função de verossimilhança

A *função de verossimilhança* $p(\mathbf{d}|\mathbf{m})$ em (2.4) é de grande importância na inversão bayesiana. De acordo com Bodin & Sambridge (2009) esta função quantifica a capacidade de um determinado conjunto de valores de parâmetros reproduzir bem os dados observados. Uma vez que definimos prioris uniformes a fim de evitar inicialmente que algum valor de parâmetro seja priorizado, a função de verossimilhança carrega a informação sobre os parâmetros $\mathbf{m}$ na ocorrência de $\mathbf{d}$. Desta maneira, considerando que os erros em (2.2) sejam aditivos, gaussianos, centrados em zero e com desvio padrão constante, ou seja, $\epsilon_i \sim N(0, \Sigma)$, algumas estatísticas do

[1] Caso requerido, $p(\mathbf{d})$ pode ser obtido pela seguinte integral: $\int p(\mathbf{d}|\mathbf{m})p(\mathbf{m})d\mathbf{m}$

conjunto de dados podem ser inferidas. Como $g(\mathbf{m})$ é uma função determinística, temos que:

$$E[\mathbf{d}] = E[g(\mathbf{m})] + E[\boldsymbol{\epsilon}] = g(\mathbf{m})$$

e

$$\mathbf{C}_d = \begin{bmatrix} E[\epsilon_1^2] & E[\epsilon_1\epsilon_2] & \cdots & E[\epsilon_1\epsilon_k] \\ E[\epsilon_2\epsilon_1] & E[\epsilon_2^2] & \cdots & E[\epsilon_2\epsilon_k] \\ \vdots & \vdots & \ddots & \vdots \\ E[\epsilon_k\epsilon_1] & E[\epsilon_k\epsilon_2] & \cdots & E[\epsilon_k^2] \end{bmatrix}$$

onde $E[\mathbf{d}]$ é o valor esperado de $\mathbf{d}$ e $\mathbf{C}_d$ a matriz de covariância dos erros nas medidas de $\mathbf{d}$. Considerando que as amostras são independentes umas das outras, temos que $E[\epsilon_i\epsilon_j] = 0$ para $i \neq j$; logo a matriz $\mathbf{C}_d$ se simplifica para $\Sigma^2\mathbf{I}_k$, onde $\mathbf{I}_k$ é a matriz identidade de ordem k (em que k é o número de amostras). A modelagem sugere que as variáveis aleatórias de $\mathbf{d}$ coponham um *vetor aleatório gaussiano* e, por serem observações fixas de tal vetor, podemos escrever a *função de verossimilhança* da seguinte maneira:

$$p(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^{k/2}[det(\mathbf{C}_d)]^{1/2}} exp\left\{-\frac{[\mathbf{d}-g(\mathbf{m})]^t\ \mathbf{C}_d^{-1}\ [\mathbf{d}-g(\mathbf{m})]}{2}\right\} \tag{2.5}$$

De acordo com a função $p(\mathbf{d}|\mathbf{m})$ notamos que quanto melhor for o ajuste entre os dados e os resultados da modelagem direta, maior será o valor da *função de verossimilhança*, e quanto pior tal ajuste (maiores diferenças em $\mathbf{d}-g(\mathbf{m})$), menor será a verossimilhança; daí estamos diante do problema de encontrar $\hat{\mathbf{m}}_{ML}$, denominado estimador de máxima verossimilhança (ML, de *Maximun Likelihood*), que minimize essa diferença. Para os métodos MCMC aqui utilizados, a fim de demandar menores tempos de processamento, usaremos a *função de verossimilhança* não normalizada:

$$p(\mathbf{d}|\mathbf{m}) \propto exp\left\{-\frac{[\mathbf{d}-g(\mathbf{m})]^t\ \mathbf{C}_d^{-1}\ [\mathbf{d}-g(\mathbf{m})]}{2}\right\} \tag{2.6}$$

o que não acarreta nenhuma perda de informação para o algoritmo, como será explicado mais adiante. Outros estimadores bastante utilizados em inversão bayesiana são o estimador de máxima solução *a posteriori* (MAP) e o valor esperado dos parâmetros, os quais denotaremos, respectivamente por $\hat{\mathbf{m}}_{MAP}$ e $\hat{\mathbf{m}}_{\mu}$. Os estimadores ML e MAP diferem entre si na inclusão da função *a priori*. Enquanto o ML maximiza (2.6), o MAP maximiza a *posteriori* (2.4).

A distribuição *a posteriori* contém todas as informações probabilísticas dos parâmetros de interesse na inversão, de modo que medidas estatísticas como a esperança sobre funções de tais parâmetros ou distribuições marginais posterioris são necessárias para estudos de estimações (BROOKS, 1998; EHRLERS, 2014); e, todas essas medidas são obtidas por integração (inclusive *a posteriori*). Porém, dificilmente podemos realizar essas integrações analiticamente e aí procedemos com aproximações numéricas, que dentre elas têm métodos estocásticos cujas estimativas vêm de números gerados aleatoriamente denominados *métodos de Monte Carlo*.

2.5 O Algoritmo Metropolis-Hastings

Para modelos realísticos, como nosso problema é não linear, achamos não ser possível assumir uma densidade de probabilidade *a priori* $p(\mathbf{m})$ de modo que se obtenha uma densidade de probabilidade *a posteriori* $p(\mathbf{m}|\mathbf{d})$ analítica; então um método numérico é usado para gerar amostras de todos parâmetros possíveis, de modo

que a inferência sobre $p(\mathbf{m}|\mathbf{d})$ torne-se inferência sobre as amostras $\mathbf{m}$. Os métodos numéricos iterativos que aproximam a densidade $p(\mathbf{m}|\mathbf{d})$ geram uma grande amostra, e a técnica muito usada é o método de amostragem de Monte Carlo via cadeia de Markov (MCMC) usando os algoritmos de Metropolis-Hastings, (METROPOLIS et al., 1953; HASTINGS, 1970), para problemas de dimensão fixa, e *reversible jump markov chain monte carlo* (RJ-MCMC) (GREEN, 1995), onde a quantidade de parâmetros também é variável.

De acordo com Gomes (2016), o algoritmo MCMC é baseado numa cadeia de Markov que gera um passeio aleatório (*random walk*) pelo espaço amostral dos parâmetros $\mathbf{m}$ e sucessivamente visita soluções com frequências estáveis decorrentes de uma distribuição estacionária. A grosso modo, a partir de um conjunto de parâmetros $\mathbf{m}_i$ do estado atual da cadeia de Markov, gera-se uma nova proposta $\mathbf{m}_j$ usando uma distribuição salto $q(\mathbf{m}_i \rightarrow \mathbf{m}_j)$ geradora de amostras, e o algoritmo faz a aceitação dessa nova estimativa $\mathbf{m}_j$ se satisfaz a condição de probabilidade (MOSEGAARD; SAMBRIDGE, 2002):

$$\alpha = min\left\{1, \frac{\pi(\mathbf{m}_j)}{\pi(\mathbf{m}_i)}\right\}. \tag{2.7}$$

onde $\pi(\mathbf{m})$ é alguma distribuição a qual se deseja amostrar que tenha $\mathbf{m}$ como variável aleatória. Note que em (2.7) toda constante de normalização é automaticamente cancelada, daí a razão de desprezá-la nessas metodologias. Em MULLER (2008) se tem o destaque que tais amostras podem ser geradas por uma cadeia de Markov que tenha como distribuição estacionária e invariante $\pi(\mathbf{m})$, uma vez que ela exista.

Consideremos a sequência de variáveis aleatórias $(\mathbf{m}_1, \mathbf{m}_2, \ldots)$. Esta sequência é definida cadeia de Markov com espaço de estados S, se para todo $s_i, s_j \in S$, com $t \geq 1$, termos (SILVA et al., 2017):

$$Pr(\mathbf{m}_{t+1} = s_j|\mathbf{m}_0 = s_{i_0}, \mathbf{m}_1 = s_{i_1}, \ldots, \mathbf{m}_{t-1} = s_{i_{t-1}}, \mathbf{m}_t = s_{i_t}) = Pr(\mathbf{m}_{t+1} = s_j|\mathbf{m}_t = s_{i_t}) \tag{2.8}$$

Onde $Pr(.|.)$ é uma medida de probabilidade condicionada. Em palavras, uma cadeia de Markov é um *processo estocástico* onde o estado futuro depende unicamente do estado atual, e independe de todos passados. É necessário ainda que a cadeia seja, para os métodos MC, homogênea, irredutível e aperiódica. Respectivamente, a grosso modo, é necessário que: as probabilidades de ir de um estado para outro sejam invariantes com o tempo; qualquer estado possa ser alcançado a partir de outro num número finito de iterações; e ter seu período igual a 1.

Após obter as amostras aleatórias aceitáveis de $\mathbf{m}$ pelo MCMC teremos uma distribuição de frequências que é assintoticamente proporcional à $\pi(\mathbf{m})$ (MOSEGAARD; SAMBRIDGE, 2002).

O método MCMC mais utilizado para gerar amostras de uma densidade de probabilidade intratável analiticamente é o algoritmo Metropolis-Hastings. Ele consiste num método de simulação baseado em uma Cadeia de Markov que permite amostrar aleatoriamente qualquer distribuição de probabilidades $\pi(\mathbf{m})$. A cadeia que compõe o algoritmo tem que estar definida no mesmo suporte de $\pi(\mathbf{m})$, ou seja no espaço amostral de $\mathbf{m}$; deve também ser composto por um critério de aceitação ou rejeição das transições propostas pela Cadeia de Markov (GILKS; RICHARDSON; SPIEGELHALTER, 1995). O comportamento da sequência aleatória é regido por uma distribuição de probabilidade condicionada, também chamada de distribuição de transição, proposta ou de salto, que é responsável por determinar a probabilidade de ocorrência de valores em $\mathbf{m}$, dado o estágio atual da cadeia. Essa distribuição é comumente denotada por $q(\mathbf{m}_t|\mathbf{m}_{t-1})$, a qual representa a probabilidade de transição do estágio $\mathbf{m}_{t-1}$, para $\mathbf{m}_t$. Originalmente o algoritmo de Metropolis foi elaborado para receber apenas distribuições de salto simétricas, apenas depois, Hastings expandiu o método para casos de distribuições de salto não simétricas. Supondo que $\pi(\mathbf{m})$ seja a distribuição no qual se deseja obter amostras aleatórias e supondo que já seja conhecida a função geradora $q(\mathbf{m}_t|\mathbf{m}_{t-1})$, podemos sequenciar o algoritmo

Metropolis-Hastings da seguinte forma:

1. Inicialize a cadeia em $\mathbf{m}_{t=1}$;
2. Repita:
   - Gerar um candidato $\mathbf{m}_t$ usando $q(\mathbf{m}_t|\mathbf{m}_{t-1})$;
   - Avaliar a probabilidade de aceitação do movimento da cadeia através de

$$\alpha = min\left\{1, \frac{\pi(\mathbf{m}_t)q(\mathbf{m}_{t-1}|\mathbf{m}_t)}{\pi(\mathbf{m}_{t-1})q(\mathbf{m}_t|\mathbf{m}_{t-1})}\right\} \tag{2.9}$$

   - Gerar aleatoriamente u da distribuição uniforme $U(0,1)$;
   - se $\alpha \geq u$, aceitar $\mathbf{m}_t$ como amostra de $\pi(\mathbf{m})$ e fazer $\mathbf{m}_{t-1} = \mathbf{m}_t$;
   - Se não, rejeitar a proposta de transição e descartar $\mathbf{m}_t$.

Se uma densidade de salto $q(.)$ simétrica é usada, a probabilidade de aceitação dada pela equação (2.9) simplifica-se para

$$\alpha = min\left\{1, \frac{\pi(\mathbf{m}_t)}{\pi(\mathbf{m}_{t-1})}\right\} \tag{2.10}$$

e o algoritmo MH assume a configuração conhecida como *Random Walk Metropolis*

Em cada estágio, a cadeia propõe um novo candidato de acordo com a função geradora. Se o candidato é aceito, a cadeia se move, caso contrário, permanece no mesmo estado.

2.6 MCMC com Saltos Reversiveis (RJ-MCMC)

O MH é eficiente uma vez que o número de parâmetros seja fixo; entretanto, o MCMC com Saltos Reversiveis (RJ-MCMC) permite, além de estimar parâmetros, fazer inferência sobre incertezas no próprio modelo, com a versatilidade de poder variar a dimensão paramétrica. Por esse motivo, tem sido chamado de "algoritmo transdimensional auto parametrizável" (RAY; KEY, 2012). Segundo Salazar-Gonzales (2004), o MCMC com saltos reversíveis pode ser interpretado como uma maneira mais geral do MH por introduzir saltos entre parâmetros de diferentes dimensionalidades. Um possível sequenciamento do RJ-MCMC pode ser dado da seguinte maneira descrita abaixo (EHLERS, 2011). Para ela, suponha que a cadeia está em um estado $(n, \mathbf{m})$, ou seja, n especifica o modelo e $\mathbf{m}$ é o vetor atual com parâmetros, e um novo modelo n' com parâmetros $\mathbf{m}$' é proposto com alguma probabilidade de salto $j_{n\to n'}$ que, de maneira geral, pode incluir ou retirar parâmetros. Supondo ainda que o modelo proposto tem dimensão maior que o atual ($dim(n') > dim(n)$) e que $\mathbf{m}' = G(\mathbf{m}, \mathbf{u})$ tal que G é função determinística e $\mathbf{u} \sim q(\mathbf{u})$ algum vetor aleatório com dimensão $dim(n') - dim(n)$. Logo pode-se usar a seguinte sequência de passos:

- propor saltar de $(n, \mathbf{m})$ para $(n', \mathbf{m}')$ com probabilidade $j_{n\to n'}$
- gerar aleatóriamente $\mathbf{u} \sim q(\mathbf{u})$ com dimensão $dim(n') - dim(n)$
- aceitar $(n', \mathbf{m}')$ como amostra da cadeia com probabilidade

$$\alpha_{(\mathbf{m}'|\mathbf{m})} = min\left\{1, \frac{\pi(\mathbf{m}')}{\pi(\mathbf{m})} \times \frac{j_{n'\to n}}{j_{n\to n'}q(\mathbf{u})}\left|\frac{\partial G(\mathbf{m}, \mathbf{u})}{\partial(\mathbf{m}, \mathbf{u})}\right|\right\} \tag{2.11}$$

Como requisito, função $\mathbf{m}' = G(\mathbf{m}, \mathbf{u})$ necessita ser uma bijeção para garantir a reversibilidade da cadeia (AVILA, 2008).

2.7 Considerações usadas na implementação

Adotaremos para este trabalho a configuração *birth-death* onde em cada iteração uma das três ocasiões acontecem: Nasce uma camada no modelo (*birth*); desaparece uma camada do modelo (*death*); ou o modelo permanece com dimenção fixa e uma camada é atualizada (*update*). Esse modo de utilização do RJ-MCMC simplifica o jacobiano em (2.11) para a unidade como é dito por Ray & Key (2012) e mostrado por Bodin & Sambridge (2009).

De maneira geral, como foi visto no capítulo 1, temos que a componente radial de campo no dipolo magnético $H_z = g(n, \boldsymbol{\rho}, \mathbf{h})$ onde $\boldsymbol{\rho}^t = [\rho_1, \rho_2, \ldots, \rho_n]$ é o vetor de resistividades das camadas e $\mathbf{h}^t = [h_1, h_2, \ldots, h_{n-1}]$ as espessuras. Procederemos nos movimentos entre os modelos de maneira similar a Ray & Key (2012) que aplicaram o RJ-MCMC para estimativas de resistividades em meios anisotrópicos invertendo em células; e Madeira (2018) que usou o algoritmo MH para fazer inferência de parâmetros da geofísica que possibilitam a estimativa de resistividade aparente em um meio homogêneo. Nas subseções seguintes estão as descrições dos elementos envolvidos na modelagem bayesiana, cuja aplicação resultaram nos resultados apresentados no próximo capítulo.

2.7.1 Prioris

A informação *a priori* carrega a informação sobre os parâmetros independentemente das observações. Uma vez que um modelo $\mathbf{m}$ é escolhido, sua dimenção é fixada e a priori deste modelo e dada da seguinte forma:

$$p(\mathbf{m}) = p(\mathbf{m}|n)p(n). \tag{2.12}$$

Definimos a priori uniforme discreta para n como sendo:

$$p(n) = \begin{cases} \dfrac{1}{n_{max} - n_{min} + 1}, & \text{se} \quad n_{min} \leq n \leq n_{max} \\ 0, & \text{caso contrário} \end{cases} \tag{2.13}$$

Considerando as prioris dos parâmetros independentes entre si, podemos escrever o seguinte produto de prioris contínuas uniformes:

$$p(\mathbf{m}|n) = p(\boldsymbol{\rho}|n)p(\mathbf{h}|n) \tag{2.14}$$

onde

$$p(\boldsymbol{\rho}|n)p(\mathbf{h}|n) = \begin{cases} \dfrac{1}{(\rho_{max} - \rho_{min})^n} \times \dfrac{1}{(h_{max} - h_{min})^n}, & \text{se} \quad \rho_{min} \leq \rho \leq \rho_{max} \text{ e } h_{min} \leq h \leq h_{max} \\ 0, & \text{caso contrário} \end{cases} \tag{2.15}$$

Fazendo $(n_{max} - n_{min} + 1) = \Delta n$, $\rho_{max} - \rho_{min} = \Delta\rho$ e $h_{max} - h_{min} = \Delta h$, a priori do modelo toma a seguinte forma:

$$p(\mathbf{m}) = \begin{cases} (\Delta\rho\Delta h)^{-n}(\Delta n)^{-1}, & \text{se} \quad \rho \in [\rho_{min}, \rho_{max}] \text{ e } h \in [h_{min},\ h_{max}], \forall n \in [n_{min}, n_{max}] \\ 0, & \text{caso contrário} \end{cases} \tag{2.16}$$

2.7.2 Probabilidade de aceitação e movimento entre os modelos

A probabilidade de aceitação $\alpha_{(\mathbf{m}'|\mathbf{m})}$ de (2.11) pode ser reescrita para um problema inverso de maneira geral como fizeram Ray & Key (2012, p. 1139):

$$\alpha_{(\mathbf{m}'|\mathbf{m})} = min\left[1, \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \times \frac{p(\mathbf{m}')}{p(\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} \times |\mathbf{J}|\right] \tag{2.17}$$

em que $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$ é a razão de verossimilhança entre os modelos; $\frac{p(\mathbf{m}')}{p(\mathbf{m})}$ é a razão a priori; e $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}$ é a razão de salto entre modelos. O termo $|\mathbf{J}|$ é o último fator de (2.11) e é considerado sendo 1, como já foi dito.

De acordo com Ray & Key (2012), o algoritmo terá uma oposição ao movimento *birth*; pelo fato de que no nascimento de uma camada a razão a priori sempre vai ser menor que 1, pelo fato de que $p(\mathbf{m}')$ tem que integrar pra a unidade em um número maior de parâmetros que seu denominador. Isso evita que o programa tente ajustar os dados apenas acrescentando camadas no modelo. Propomos os saltos *birth*, *death* e *update* com as seguintes probabilidades $p[b, d, u] = [1/3, 1/3, 1/3]$. A função de salto que propõe novos valores para ρ e h, como será justificado posteriormente nos esquemas de movimentos, é um vetor aleatório gaussiano $N(\boldsymbol{\mu}_{\rho h}, \mathbf{C}_{bd})$ centrado nos valores de parâmetros na camada amostrada na última vez que o modelo decorrente foi visitado; e

$$\mathbf{C}_{bd} = \begin{bmatrix} \Sigma_\rho^2 & 0 \\ 0 & \Sigma_h^2 \end{bmatrix}$$

A escolha da função de salto é de crucial importância para a velocidade de convergência da cadeia (AVILA, 2008).

2.7.3 Distribuição proposta e probabilidade de aceitação

2.7.3.1 Movimento em dimensão fixa (update)

Uma vez que é escolhido o movimento de atualização, uma camada é escolida aleatoriamente e tem sua resistividade é espessura atualizadados. Como, para esse tipo de movimento, $\frac{p(\mathbf{m}')}{p(\mathbf{m})}$ e $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}$ são ambos iguais a um, a razão de aceitação se resume no mínimo entre um e a razão de verossimilhança, ou seja,

$$\alpha_u = \begin{cases} \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right], & \text{se} \quad \rho \in [\rho_{min}, \rho_{max}] \text{ e } h \in [h_{min},\ h_{max}], \forall n \in [n_{min}, n_{max}] \\ 0, & \text{caso contrário} \end{cases} \tag{2.18}$$

2.7.3.2 Nascimento de camada (birth)

Supondo que estamos no estado com n camadas, um estado futuro terá $n+1$. Logo, se estivéssemos no estado futuro, teríamos $n + 1 - 1$ camadas na densidade de salto. Considerando que cada componente do vetor

aleatório é independente das demais ficaremos com:

$$q(\mathbf{m}|\mathbf{m}') = \frac{1}{3} \times \prod_{i=1}^{n} \frac{1}{2\pi\Sigma_\rho\Sigma_h} exp\left\{ - \left[\left(\frac{\rho_i' - \rho_i}{2\Sigma_\rho} \right)^2 + \left(\frac{h_i' - h_i}{2\Sigma_h} \right)^2 \right] \right\} \tag{2.19}$$

e uma densidade de salto de n para $n+1$:

$$q(\mathbf{m}'|\mathbf{m}) = \frac{1}{3} \times \prod_{i=1}^{n+1} \frac{1}{2\pi\Sigma_\rho\Sigma_h} exp\left\{ - \left[\left(\frac{\rho_i' - \rho_i}{2\Sigma_\rho} \right)^2 + \left(\frac{h_i' - h_i}{2\Sigma_h} \right)^2 \right] \right\} \tag{2.20}$$

onde ρ_i' e h_i' são os parâmetros novos e ρ_i e h_i os antigos. Como os produtórios das densidades acima são simétricos de 1 até n, a razão de transição toma a forma:

$$\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} \right]_b = 2\pi\Sigma_\rho\Sigma_h exp\left\{ \left(\frac{\rho_i' - \rho_i}{2\Sigma_\rho} \right)^2 + \left(\frac{h_i' - h_i}{2\Sigma_h} \right)^2 \right\} \tag{2.21}$$

Logo, de (2.16), (2.17) e (2.21), considerando que $n \in [n_{min}, n_{máx}]$, temos qua a probabilidade de aceitação para o nascimento de camada é dada por:

$$\alpha_b = \begin{cases} \min\left[1, \dfrac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \dfrac{2\pi\Sigma_\rho\Sigma_h}{\Delta\rho\Delta h} exp\left\{ \left(\dfrac{\rho_i' - \rho_i}{2\Sigma_\rho} \right)^2 + \left(\dfrac{h_i' - h_i}{2\Sigma_h} \right)^2 \right\} \right], & \text{se } \rho \in [\rho_{min}, \rho_{max}] \text{ e } h \in [h_{min},\ h_{max}] \\ 0, & \text{caso contrário} \end{cases} \tag{2.22}$$

2.7.3.3 Morte de camada (death)

Usando um processo dedutivo similar ao citado no movimento de nascimento de camada, chegamos à seguinte probabilidade de aceitação:

$$\alpha_d = \begin{cases} \min\left[1, \dfrac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \dfrac{\Delta\rho\Delta h}{2\pi\Sigma_\rho\Sigma_h} exp\left\{ - \left[\left(\dfrac{\rho_i' - \rho_i}{2\Sigma_\rho} \right)^2 + \left(\dfrac{h_i' - h_i}{2\Sigma_h} \right)^2 \right] \right\} \right] & \text{se } \rho \in [\rho_{min}, \rho_{max}] \text{ e } h \in [h_{min},\ h_{max}] \\ 0, & \text{caso contrário} \end{cases} \tag{2.23}$$

2.7.4 Considerações na Implementação

Os códigos para os MCMCs foram feitos em ambiente MATLAB, e para aplicação dos algoritmos MH e RJ-MCMC seguimos alguns cuidados para assegurar a convergência, uma vez que aumentar a quantidade de parâmetros acarreta consequentemente na nessecidade de maior número de iterações para percorrer o espaço amostral dos mesmos. Um período de aquecimento das amostras é necessário para que a cadeia não armazene valores que não estejam no suporte paramétrico. Para N iterações, tal período chamado *burn-in*, descarta as N_b primeiras amostras onde N_b especifica o número máximo de amostras que não farão parte da cadeia, considerando que a mesma tenha convergido para $N \geq N_b$. Neste trabalho usou-se $N = 50000$, das quais 10% serviram de aquecimento para a cadeia de Markov. O critério de convergência usado foi a avaliação da *posteriori* 2.4; tomando uma amostra $\mathbf{m}_i$, consideramos que a cadeia atingiu a área de convergência se $p(\mathbf{m}_i|\mathbf{d}) > 0$. Os cálculos foram executados em um CPU de 2 núcleos (@2,40 GHz) com, em média, uma hora

e meia para o experimento com RJ-MCMC e sete horas para o MH, em cadeia única. Para se trabalhar com múltiplas cadeias em paralelo, como se faz comumente, é necessário bem mais alto desempenho computacional que o disponível para esta pesquisa.

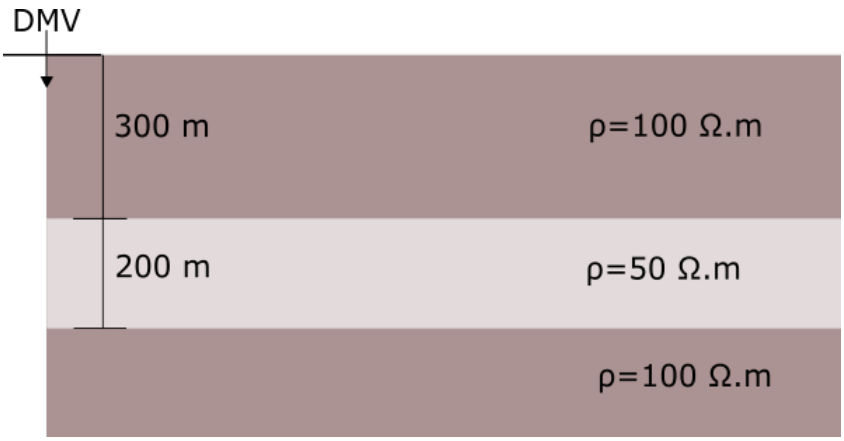
Assumiu-se, para o código de saltos reversíveis, que o meio investigado pudesse variar, em número de camadas, entre $n_{min} = 1$ e $n_{máx} = 6$. Os intervalos assumidos a *priori* para suporte de busca para ρ e h foram, respectivamente, $[\rho_{min} = 10^{-1}, \rho_{máx} = 10^{3}]$ $\Omega.m$ e $[h_{min} = 1, h_{máx} = 500]$ m, para a pesquisa em camadas profundas e $[h_{min} = 0.1, h_{máx} = 10]$ m para pesquisa mais rasa. Uma vez que tanto ρ quanto h podem assumir diferentes ordens de grandeza, optamos por inverter $log_{10}(\boldsymbol{\rho}, \mathbf{h})$ e posteriormente aplicou-se a transformação inversa para exibir os estimadores em suas contrapartes lineares.

3

Resultados

Para a aplicação dos códigos MH e RJ-MCMC simulamos, primeiramente para uma investigação mais profunda, os dados de um meio de $n = 3$ camadas como ilustra a figura (3.1). Usou-se para o MH e RJ-MCMC quatro posições de medidas. As características das frequências e posições de medidas serão discutidas posteriormente para cada experimento. Após gerar os dados, foi adicionado um ruído gaussiano nas observações simuladas.

Figura 3.1: Modelo teórico de um meio estratificado com resistividades de 100, 50 e 100 Ω.m respectivamente



3.1 Metropolis-Hastings (MH) na Investigação com o DMV

Na aplicação do MH, foi fixado um número de camadas para inversão de suas resistividades. Dividiu-se o meio em 6 camadas horizontais, como modelo interpretativo para o código, onde as 5 primeiras tiveram suas espessuas fixadas constantes igual a 100 m. O intervalo de busca para ρ foi [0.1,1000] $\Omega.m$. As posições dos receptores foram 100, 150, 200 e 250; 4 frequências logaritmicamente espaçadas entre 0.1 e 1×10^3 Hz; e o ruído aditivo usado Σ foi 1×10^{-3}. Para função densidade de salto foi usado um vetor aleatório gaussiano centrado, a cada iteração, no vetor de valores anteriores da cadeia, ou seja, $q(\mathbf{m}_t|\mathbf{m}_{t-1}) \sim N(\mathbf{m}_{t-1}, \mathbf{C}v)$, onde $\mathbf{C}v = Diag(3, 5, 5, 9, 5, 5)$ é a matriz diagonal de covariância usada para os saltos. O tempo requerido para esta aplicação foi de aproximadamente 3 horas. A figura (3.2) mostra a estimativa, em azul, atravéz do valor esperado $E[\mathbf{m}|\mathbf{d}] = \hat{\mathbf{m}}_\mu$ e (3.3) mostra o estimador ML $\hat{\mathbf{m}}_{ML}$ em vermelho, comparando com o modelo real destacado em preto. Note que que o ML se aproximou mais do modelo real que o valor esperado. Esse efeito se dá pela incerteza advinda dos erros aditivos das amostras. Pode-se analisar a figura (3.4), a qual

representa os histogramas não normalizados das PDFs marginais dos parâmetros. É visível que a PDF para estimar a resistividade dessas região geométricas, 3º e 5º camadas, é multimodal. Porém, é notório que a região com maior probabilidade está próxima do valor real para este modelo, 50 $\Omega.m$, apenas na 3º camada. Os demais histogramas trazem uma única moda. Notou-se nesta pesquisa que quanto maior a quantidade de parâmetros a se inverter, maior a incerteza sobre os mesmos se levarmos em conta a mesma estrutura de dados coletados. Na próxima seção descreve-se a aplicação do RJ-MCMC numa tentativa de minimizar esse problema.

Figura 3.2: Valor esperado para as resistividades, dados fixados, o número de camadas e espessuras

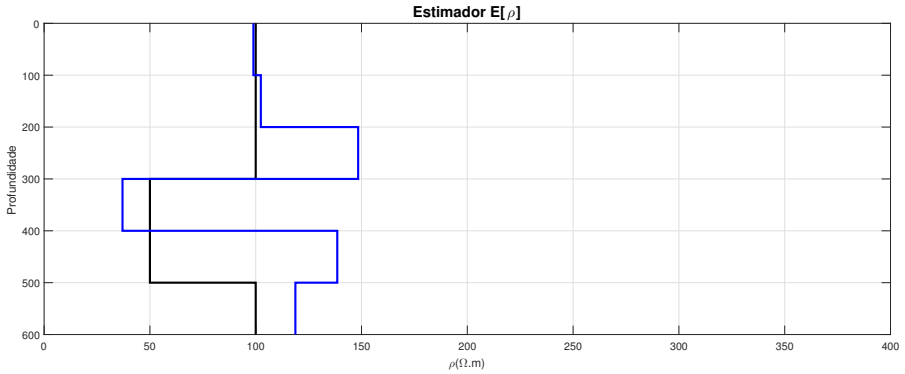
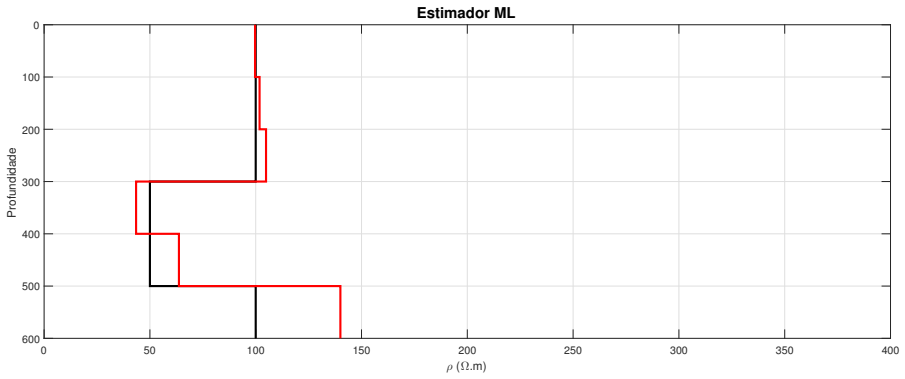


Figura 3.3: Estimador ML, dados fixados, o número de camadas e espessuras



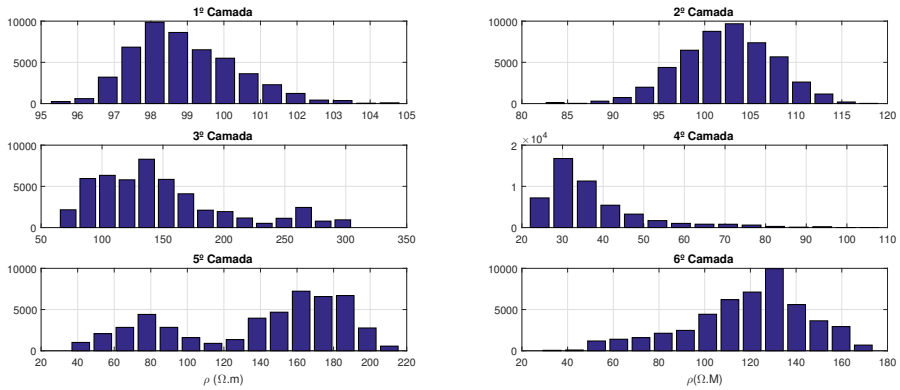
3.2 MCMC com Saltos reversíveis (RJ-MCMC) na Investigação com o DMV

Ao configurar o RJ-MCMC para aplicação no modelo ilustrado na figura (3.1), usamos as posições de medidas 100, 150, 200 e 250 e 4 frequências logaritmicamente espaçadas entre 0.1 e 1×10^3 Hz. Para o ruído aditivo Σ foi usado 1×10^{-3}. Para a matriz de covariância de salto usamos os desvios padões $\Sigma_\rho = 0,1$ e $\Sigma_h = 0,1$. O código foi inicializado com o modelo mínimo de camadas, $n = 1$, com chutes iniciais para ρ e h iguais a $log_{10}50$ e $log_{10}150$ respectivamente, uma vez que a busca foi feita em escala logarítmica.

A marginalização da PMF[1] do modelo executada pelo programa é mostrada na figura (3.5). A figura

[1]Para o caso da função de probabilidade do modelo, é uma *probability mass function (PMF)* por se tratar de um conjunto

Figura 3.4: Histogramas não normalizados das PDFs referentes às camadas com h fixo



representa o histograma normalizado de $p(n|\mathbf{d})$ e foi obtido pelo quociente $f_{n_i}/(N - N_b)$ (frequência absoluta de cada n_i/total de iterações pós *burn-in*).

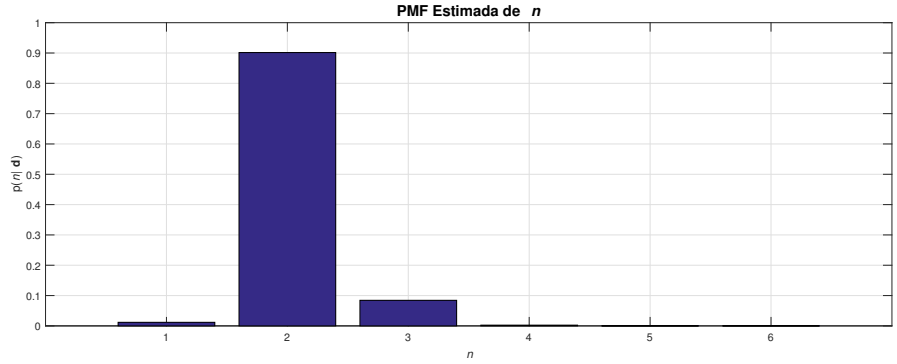


Figura 3.5: Distribuição a posteriori do modelo $p(n|\mathbf{d})$

Note que o modelo mais visitado foi $n = 2$, o que nos dá como solução os vetores estimados $\boldsymbol{\rho}^t = [\rho_1, \rho_2]$ e $\mathbf{h} = [h_1]$.

Notou-se que a *função de verossomilhança* em (2.6) é altamente sensível às características das frequências e posições de medidas usadas. Outro fator importante a ser levado é que se trata de uma inversão estatística. A variância de um sistema o torna probabilístico, ou seja, um modelo aleatório.

A figura (3.6) mostra as estimativas retornadas; em preto temos o modelo real, em vermelho o valor esperado $\hat{\mathbf{m}}_\mu$ e em azul o estimador $\hat{\mathbf{m}}_{MAP}$.

Ainda a respeito da figura (3.6) temos eixo horizontal representando a resistividade em escala linear, enquanto o vertical representa a espessura e profundidade. Note que o programa conseguiu reconstruir a resistividade da primeira camada com boa aproximação, tendo um pequeno aumento na espessura da mesma. Já a segunda camada teve seu valor mais aproximado da segunda camada do modelo real, enquanto uma terceira não foi inferida pelo código.

Aplicou-se também o código de dimensão variável a investigações mais rasas usando o modelo teórico representado pela figura (3.7). Usou-se quatro posições de medidas, sendo elas, 10, 15, 20 e 25 metros do DMV. As frequências foram 8000, 8500, 8750 e 9000 Hz e $\Sigma = 1 \times 10^{-3}$. Para a densidade de salto usou-se

discreto $N_{\mathbf{m}}$ com $n \in N_{\mathbf{m}}$, tal que, $n = 1, 2, \ldots, 6$

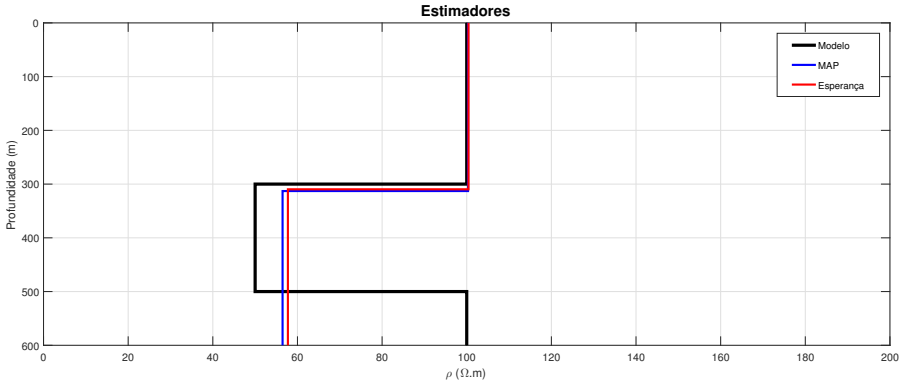
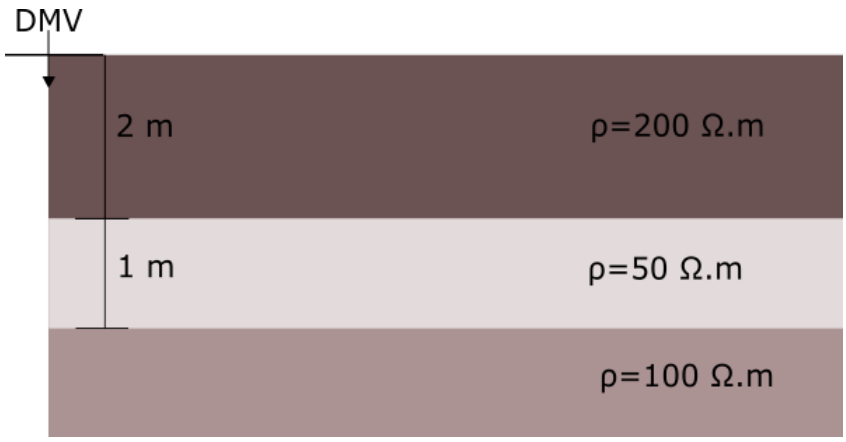


Figura 3.6: Estimadores

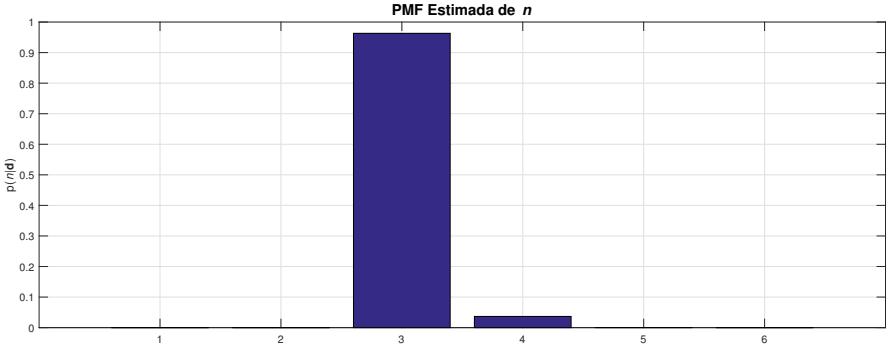
$\Sigma_\rho = 0,2$ e $\Sigma_h = 0,05$ e os chutes iniciais para ρ e h foram, respectivamente $log_{10}250$ e $log_{10}1$

Figura 3.7: Modelo teórico de um meio estratificado com resistividades de 200, 50 e 100 Ω.m respectivamente



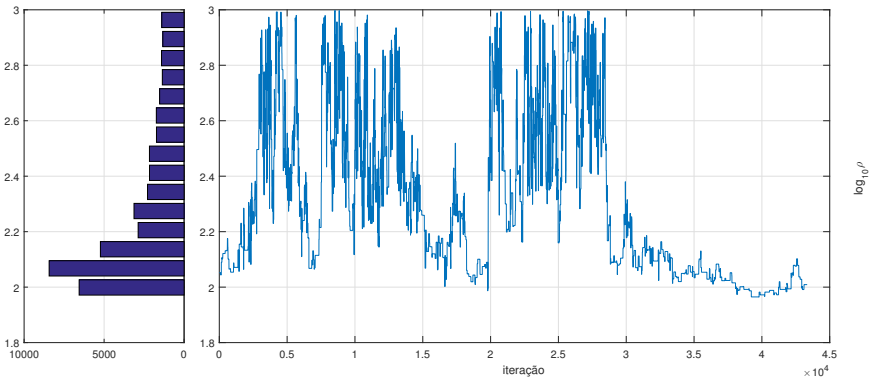
O modelo foi inferido corretamente e é estimado através do histograma normalizado mostrado na figura (3.8). O modelo mais visitado foi o de 3 camadas. Perceba que o modelo verdadeiro foi visitado em mais de 90% das iterações. Após o processo de amostragem, montamos alguns gráficos para tentar entender melhor como as variáveis aleatórias do problema se correlacionavam.

Figura 3.8: Distribuição a posteriori do modelo $p(n|\mathbf{d})$



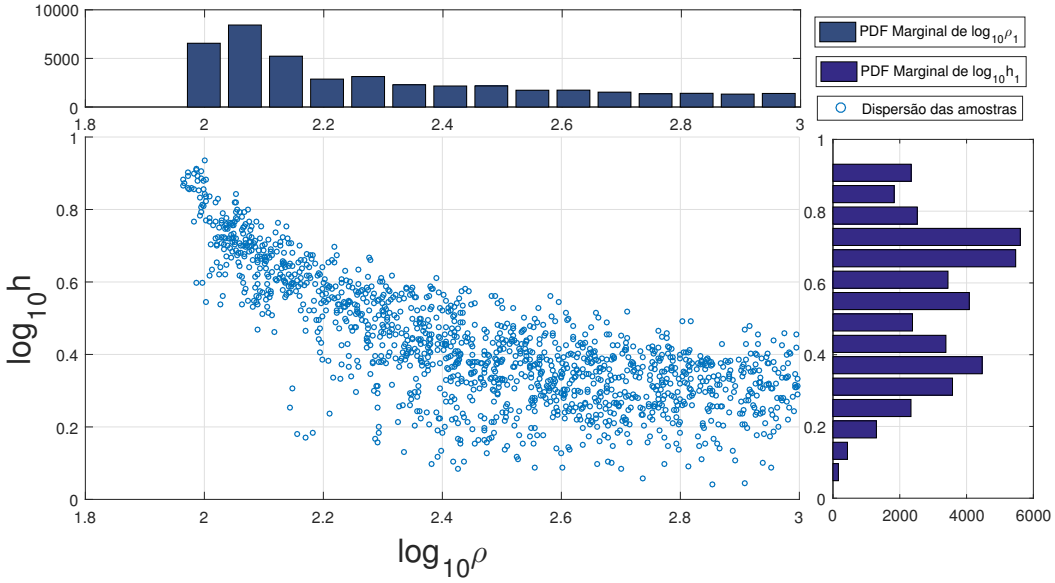
A título de ilustração, note a figura (3.9). O histograma à esquerda do gráfico representa a PDF estimada, não normalizada, de $log_{10}\rho_1$ (referente à resistividade da 1º camada), e o gráfico mostra o estado desta variável aleatória nas iterações pós aquecimento.

Figura 3.9: Cadeia de Markov Marginal de $log_{10}\rho_1$



Os gráficos (3.10) e (3.11) são representações das camadas 1 e 2. O histograma superior representa a PDF marginal para resistividade, o histograma à direita é a PDF referente a espessura, e no centro, o diagrama de dispersão das amostras. Note que há uma visível correlação entre o par de propriedades de cada camada. Na primeira camada tem-se uma correlação exponencial negativa, já na segunda camada, aparentemente, trata-se de uma correlação linear

Figura 3.10: Amostras da 1º Camada



O estimador valor esperado é representado no gráfico (3.12) em azul, comparado com o modelo real em preto. A primeira camada teve sua resistividade aproximada com boa precisão, porém teve a espessura estimada maior que o modelo real; a segunda camada teve um leve aumento na espessura e sua resistividade

Figura 3.11: Amostras da 2º Camada

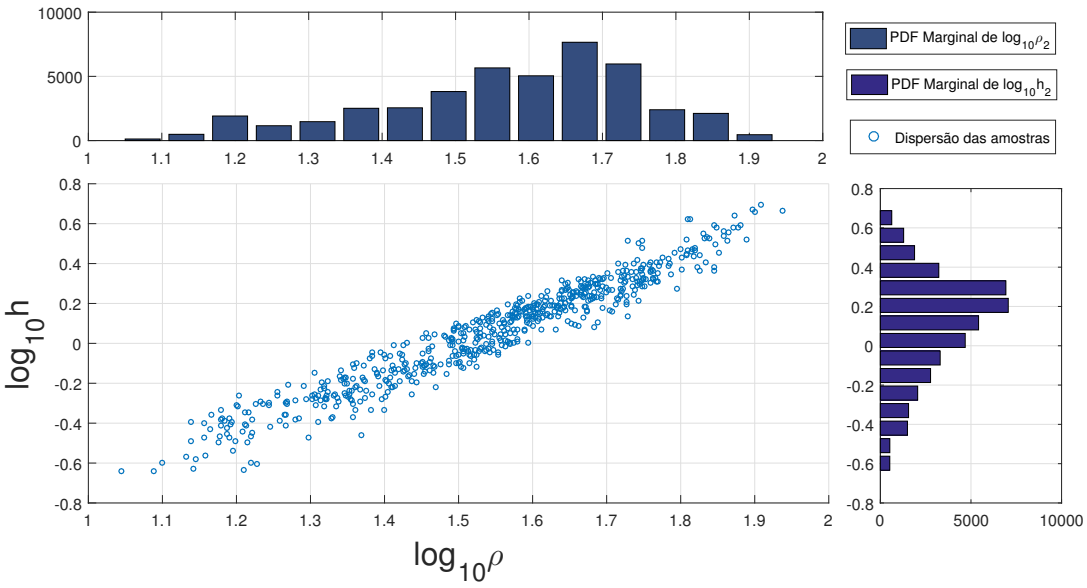
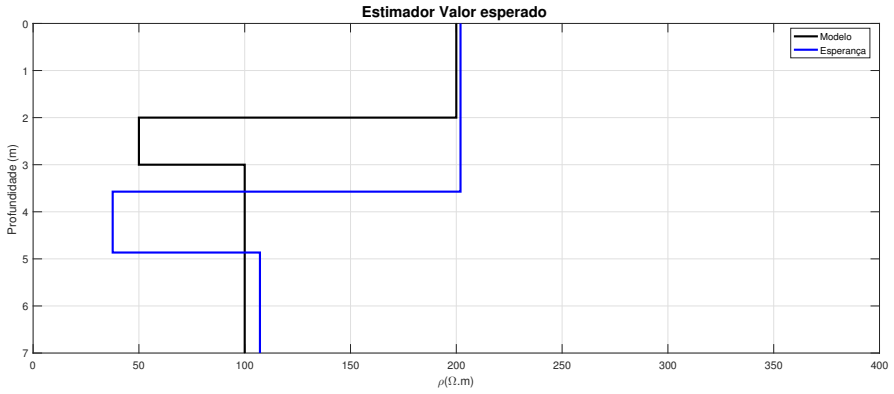


Figura 3.12: Estimador Valor Esperado



menor que a real; por fim, a última camada teve um ligeiro aumento de sua resistividade.

4

Considerações finais

No presente trabalho utilizamos uma abordagem bayesiana para inversão de parâmetros geofísicos através de dados que, eventualmente, podem estar contaminados com ruídos. Tal inversão leva em conta os erros nas medidas em sua própria formulação, tornando-a coerente para os meios investigados. Inicialmente descrevemos de maneira breve a modelagem direta do problema e algumas características importantes das propriedades físicas da subsuperfície. É comum usar técnicas de inversão baseadas em gradiente que buscam minimizar a diferença entre os dados observados e o operador direto aplicado a um conjunto de parâmetros. O conjunto de parâmetros que tornar mínima esta diferença pode ser considerado um estimador. Para os modelos bayesianos, um método livre de derivadas, busca-se maximizar a *função de verossimilhança* e, em alguns casos, a função a *posteriori*; analogamente aos métodos de gradiente, o conjunto de parâmetros que maximiza tais funçõ es são estimadores, bem como, também as medidas de tentência central como, por exemplo, esperanças dos parâmetros. Os códigos usados foram o Metroplolis-Hastings e o Método de Monte-Carlo com Saltos Reversíveis. Este primeiro, por lidar com uma quantidade maior de parâmetros, teve grande variância nas PDFs, tornando o estimador, "valor esperado", distante da realidade para as camadas abaixo de 200 metros. Fato que, em um trabalho posterior, pode ser contornado usando um estimador modal, visto que uma algumas PDFs tiveram mais de uma moda, e, ao menos uma delas, estava consideravelmente próxima do valor real. Já o estimador ML se mostrou mais acurado que a esperança. Na aplicação do RJ-MCMC, para camadas profundas, uma boa aproximação da resistividade e espessura da primeira camada foi retornada, enquanto que apenas uma resistividade posterior a primeira foi explicitada; tal fato se deve à característica da fonte. Em pesquisas mais rasas teve-se um desempenho relativamente bom do RJ-MCMC com a identificação do quão correlacionadas são o par de propriedades de cada camada (resistividade e espessura). Acreditamos que nossos códigos podem ser mais acurados em uma configurações de obtenção de dados mais específicas, bem como, dispor de uma quantidade de dados maior e um melhor desempenho computacional, os quais são fatores primordiais para se reduzir incerteza em modelos probabilísticos.

Nossos resultados mostram a aplicabilidade de modelos estocásticos às investigações geofísicas e deixam, ao leitor interessado, alguns dos primeiros passos a se tomar ao se escolher uma análise estatística probabilística de um problema inverso. Os próximos passos são: estudar sobre análise de sensibilidade dos parâmetros de maneira a se ter posições de medidas e frequências que sejam mais informativas para os métodos bayesianos; estudar mais detalhalhadamente a função de salto $q(.)$, aplicando a técnica evolutiva de Haario et al. (2001) ao RJ-MCMC que aprimora a densidade de salto usando propriedades estatísticas das amostras já aceitas; e, em conjunto disso, visto que o meio investigado tem características geométricas que podem eventualmente trazer inúmeras soluções para o problema, aplicar a técnica conhecida por Campo Aleatório de Markov (MILSZTAJN, 2003), onde a propriedade de uma área depende unicamente das propriedades das

áreas adjacentes à mesma. De maneira mais clara, tal método se assemelha a uma cadeia de Markov comum que, ao invés de domínio no tempo, possui domínio no espaço. Modelo este que é bastante usado para estimar áreas de riscos relativos em modelagens epidemiológicas. Acreditamos que as técnicas citadas podem até mesmo trazer bons resultados para uma inversão 2D.

Referências

AVILA, F. R. *Algoritmos Baseados em Modelos Bayesianos para Restauraçao de Sinais de Audio*. Tese (Doutorado) — Tese de mestrado, COPPE/UFRJ Laboratório de Processamento de Sinais, Rio de . . . , 2008. 15, 16

BODIN, T.; SAMBRIDGE, M. Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, Blackwell Publishing Ltd Oxford, UK, v. 178, n. 3, p. 1411–1436, 2009. 11, 15

BROOKS, S. Markov chain monte carlo method and its application. *Journal of the Royal Statistical Society: Series D (The Statistician)*, v. 47, n. 1, p. 69–100, 1998. Disponível em: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9884.00117>. 12

EHLERS, R. S. Inferência bayesiana. *Departamento de Matemática Aplicada e Estatıstica, ICMC-USP*, p. 64, 2011. 14

EHRLERS, R. *Inferência Bayesiana (Notas de Aula)*. [s.n.], 2014. Disponível em: <http://conteudo.icmc.usp.br/pessoas/ehlers/bayes/>. 12

GILKS, W.; RICHARDSON, S.; SPIEGELHALTER, D. *Markov Chain Monte Carlo in Practice*. Taylor & Francis, 1995. (Chapman & Hall/CRC Interdisciplinary Statistics). ISBN 9780412055515. Disponível em: <https://books.google.com.br/books?id=TRXrMWY_i2IC>. 13

GOMES, G. J. C. *Modelagem da interface solo-rocha utilizando inferência bayesiana*. Tese (Tese de Doutorado) — Pontifícia Universidade Católica, 2016. Disponível em: <https://doi.org/10.17771/PUCRio.acad.28488>. 13

GREEN, P. J. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, v. 82, p. 711–732, 1995. 13

HAARIO, H. et al. An adaptive metropolis algorithm. *Bernoulli*, Bernoulli Society for Mathematical Statistics and Probability, v. 7, n. 2, p. 223–242, 2001. 25

HANSSENS, D. et al. Frequency-domain electromagnetic forward and sensitivity modeling: Practical aspects of modeling a magnetic dipole in a multilayered half-space. *IEEE Geoscience and Remote Sensing Magazine*, v. 7, n. 1, p. 74–85, March 2019. ISSN 2373-7468. 9

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. Oxford University Press, 1970. 13

MADEIRA, R. L. C. Inversão bayesiana de dados do método eletromagnético a multi-frequência. 2018. 15

METROPOLIS, N. et al. Equation of state calculations by fast computing machines. *The journal of chemical physics*, AIP, v. 21, n. 6, p. 1087–1092, 1953. 13

MILSZTAJN, F. *Segmentaçao de tecidos cerebrais em imagens de ressonância magnética utilizando campos aleatórios de Markov*. Tese (Doutorado) — Dissertaçao de Mestrado, Universidade Federal do Paraná, Curitiba, 2003. 25

MOSEGAARD, K.; SAMBRIDGE, M. Monte carlo analysis of inverse problems. *Inverse problems*, IOP Publishing, v. 18, n. 3, p. R29, 2002. 13

MULLER, A. Simulação estocástica: o método de monte carlo. *Curitiba, PR*, p. 2–10, 2008. 13

RAY, A.; KEY, K. Bayesian inversion of marine csem data with a trans-dimensional self parametrizing algorithm. *Geophysical Journal International*, v. 191, n. 3, p. 1135–1151, 2012. Disponível em: <+http://dx.doi.org/10.1111/j.1365-246X.2012.05677.x>. 14, 15, 16

SALAZAR-GONZALES, E. *INFERÊNCIA BAYESIANA EM MODELOS AUTO-REGRESSIVOS COM TRANSICAO SUAVE PARA MEDIA E VARI^ANCIA*. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2004. 14

SILVA, L. M. d. et al. Cadeias de markov e aplicações. Volta Redonda, 2017. 13