

Universidade Federal de Santa Catarina  
Departamento de Informática e Estatística  
Programa de Pós-Graduação em Ciência da Computação

Isaac Leonardo Santos Sacramento

Texto entregue como requisito para defesa do  
Exame de Qualificação de Doutorado, contendo re-  
visão bibliográfica, problemática, proposta e resul-  
tados prévios.

Orientador: Mauro Roisenberg

Florianópolis  
2017

# Resumo

O processo de caracterização de reservatórios de hidrocarbonetos consiste na determinação tridimensional e quantitativa da estrutura e das propriedades petrofísicas das rochas da área de interesse.

Palavras chave: Inversão Sísmica; Modelagem de Incerteza; Inversão Geoestatística; Redes Neurais Convolucionais.

# Abstract

The characterization process of hydrocarbon reservoirs entails in determining the 3D structure and petrophysical properties of the rocks at the area of interest.

Keywords: Seismic Inversion; Uncertainty Modeling; Geostatistical Inversion; Convolutional Neural Networks.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Problema Inverso . . . . .	2
1.2	Inversão Sísmica . . . . .	3
1.3	Simulação Multiponto . . . . .	6
1.4	Redes Neurais Convolucionais . . . . .	6
1.4.1	Convolução . . . . .	7
1.4.2	Pooling . . . . .	8
1.5	Objetivo . . . . .	8
1.6	Organização do Texto . . . . .	9
<b>2</b>	<b>Modelos de Super-resolução</b>	<b>10</b>
2.1	Inversão Sísmica Linear e Não Linear . . . . .	11
2.1.1	Máximo <i>a posteriori</i> . . . . .	12
2.2	Modelagem de Incerteza . . . . .	14
2.2.1	<i>Global Stochastic Inversion</i> . . . . .	17
2.3	Resumo . . . . .	20
<b>3</b>	<b>Super-Resolução de Imagens Pós-Inversão</b>	<b>21</b>
3.1	Resultados Preliminares . . . . .	23
3.2	Proposta e Plano de Trabalho . . . . .	26
	<b>Bibliografia</b>	<b>29</b>

# Capítulo 1

## Introdução

Um aspecto importante nas ciências físicas é poder inferir sobre parâmetros físicos a partir de dados. Em geral, as leis da física disponibilizam os artefatos necessários para calcular valores de dados, a partir de um modelo. Este procedimento é conhecido como problema direto (*forward problem*). A modelagem direta, portanto, inicia com um modelo, sobre o qual um experimento ou processo é simulado matematicamente. Se o modelo estiver correto, a resposta obtida deve parecer com dados reais. O processo de inversão faz exatamente o contrário, consiste em utilizar as medidas efetuadas para inferir os valores de parâmetros que caracterizam o sistema (Tarantola, 2005). Muitas vezes, o problema inverso se caracteriza por ser não determinístico.

Considere o seguinte exemplo: imagine um exame de eletrocardiograma (ECG). Neste exame, a corrente elétrica responsável pelos batimentos cardíacos pode ser medida através da disposição de eletrodos sobre a superfície do corpo, próximos e distantes do coração. Partindo da suposição de que um coração doente seja examinado, é possível para a um médico identificar no ECG os padrões que ratificam a existência de um problema (problema direto). Neste contexto, o objetivo do problema inverso é recuperar a atividade elétrica e fisiológica do coração dado um conjunto de dados de ECG. Como a maioria dos problemas inversos, o problema inverso do exame de ECG recai sobre duas características comuns. Primeiro, a não unicidade de solução, ou seja, o mesmo conjunto de medidas observadas no exame pode resultar de mais de uma configuração do coração doente. Segundo, a natureza mal-posta do problema inverso, isto é, uma pequena mudança arbitrária nos valores observados no ECG pode causar uma mudança grande da solução fonte equivalente.

## 1.1 Problema Inverso

A teoria de inversão é utilizada em diversas áreas para inferir os valores de parâmetros relacionados com processos importantes a partir dos dados medidos, também chamados de dados experimentais. Pode-se descrever o problema inverso como o processo de obter informações de um sistema parametrizado, a partir de dados observáveis, das relações teóricas dos observáveis com os parâmetros não observáveis e do conhecimento *a priori* sobre os não observáveis.

O procedimento científico para o estudo de um sistema físico pode ser dividido em três passos: a parametrização do sistema, a modelagem direta e a modelagem inversa. O problema da modelagem inversa, por sua vez, se caracteriza pelo uso de resultados atuais das medições dos parâmetros físicos observáveis, para inferir os valores atuais dos parâmetros do modelo. A maioria dos problemas inversos podem ser escritos em uma forma discreta como:

$$F(m) \approx d \quad (1.1)$$

onde,  $m = (m_1, m_2, \dots, m_n) \subset R^n$  é o modelo estimado que pertence a um conjunto de modelos  $M$  admissíveis em termos de conhecimento prévio (*à priori*). O termo  $d \in R^s$ , são os dados observados e

$$F(m) = (f_1(m), f_2(m), \dots, f_s(m)) \quad (1.2)$$

representa o direto.

Na prática  $d$  pode ser uma função no domínio do tempo e/ou espaço, ou pode ser uma coleção de observações discretas. Uma questão relevante é a presença de ruído nas observações. Geralmente considera-se que os dados são observações de um experimento perfeito  $d_v$  mais uma componente de ruído  $e$ :

$$\begin{aligned} d &= F(m_v) + e \\ &= d_v + e \end{aligned} \quad (1.3)$$

onde  $d_v$  satisfaz 1.1 para o modelo considerado verdade  $m_v$ ,  $d_v = G(m_v)$ , assumindo o operador  $G$  como exato. Em muitos casos é possível ajustar matematicamente todas as partes do ruído  $e$  utilizado a equação 1.1, deixando o ruído aparecer no modelo encontrado. Situação que na prática é indesejável, pois em muitos casos uma solução de  $m$  que é influenciada por até mesmo um pequeno ruído  $e$  pode ter pouca similaridade com o modelo verdadeiro  $m_v$  (Aster; Borchers e Thurber, 2013, p. 19).

## 1.2 Inversão Sísmica

No contexto da indústria de óleo e gás, o problema de inversão sísmica é importante para o processo de caracterização de reservatórios de hidrocarbonetos, que consiste na determinação tridimensional e quantitativa da estrutura e das propriedades petrofísicas (porosidade, densidade, etc) das rochas da subsuperfície. O método de aquisição sísmica de reflexão utiliza pulsos sísmicos de uma fonte artificial controlada e monitora a resposta em função do tempo. Neste sistema, cada interface entre dois tipos de rochas diferentes gera reflexão e refração do pulso sísmico, como demonstrado na Figura 1.1.

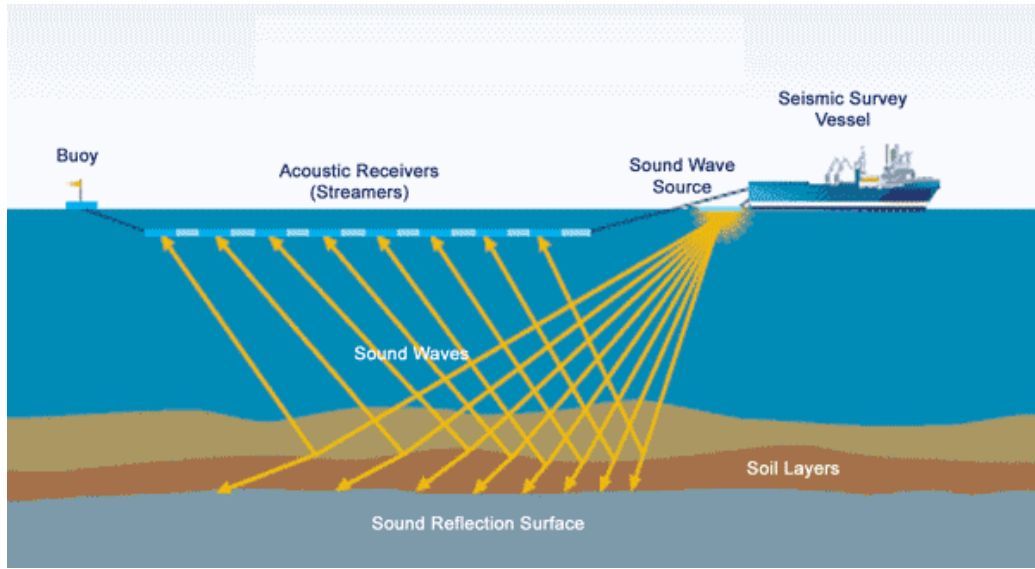


Figura 1.1: Método de sísmica de reflexão (FishSAFE, 2014)

Para inversão acústica, o modelo mais utilizado para representar a resposta do pulso sísmico atravessando as interfaces entre rochas é o convolucional. Nele é assumido um modelo em camadas para a subsuperfície e ângulo de incidência e reflexão de  $90^\circ$ . A medida efetuada é representada pelo resultado da convolução do pulso sísmico, também chamado de *wavelet*, com as refletividades das interfaces. Os coeficientes de reflexão com ângulo de reflexão normal são modelados por (Sen, 2006, p. 69):

$$r(t) = \frac{z(t + \delta t) - z(t)}{z(t + \delta t) + z(t)} \quad (1.4)$$

onde  $z(t)$  é a impedância acústica no tempo  $t$  definida por  $z(t) = \rho(t)v(t)$ , onde  $\rho(t)$  é a densidade da rocha e  $v(t)$  a velocidade de propagação da onda acústica. Utilizando os coeficientes de reflexão, modela-se a resposta do sistema  $d(t)$  aplicando a convolução da *wavelet*  $s$  com os

coeficientes de refletividade:

$$d(t) = \int_{-\infty}^{\infty} s(\tau)r(t - \tau)d\tau + e_d(t) \quad (1.5)$$

onde é assumida a presença de um ruído aleatório  $e_d(t)$  e  $d$  é o chamado dado sísmico. Cada sequência de dados  $d$  representa um ponto no plano  $xy$  e suas posições ( $t$ ) são coordenadas em tempo que se relacionam com a profundidade. Cada  $d_{xy}$  é muitas vezes chamado de traço sísmico, o que no caso discreto pode ser uma coluna de uma matriz 2D, por exemplo. Um conjunto de traços sísmicos também é chamado de uma imagem, seção ou cubo, no caso de um levantamento 3D. A *wavelet* ideal seria um pulso tipo delta contendo todas as frequências, mas gerar tal pulso não é viável. Na prática as *wavelets* são pulsos de banda limitada entre  $6Hz$  e  $65Hz$ , o que limita a frequência da sísmica e sua resolução (Sen, 2006, p. 11). A Figura 1.2 ilustra uma *wavelet* típica extraída de dados reais.

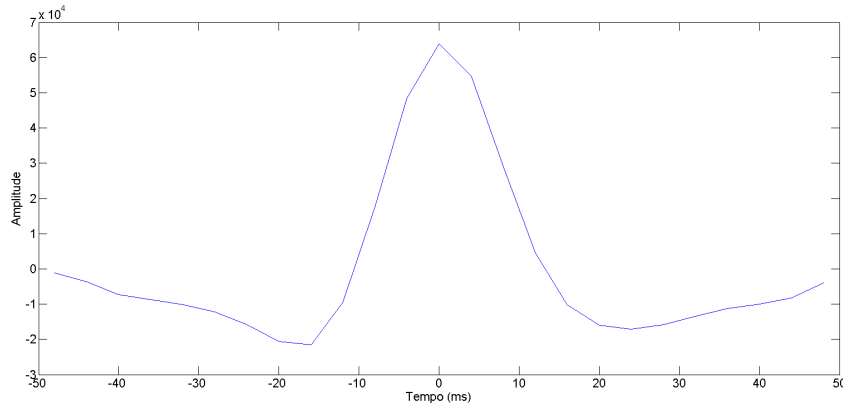


Figura 1.2: *Wavelet* extraída de dados reais

O objetivo da inversão sísmica acústica é determinar os valores de impedância acústica das camadas de rocha. Esse é um problema não linear, pois as equações que determinam o modelo direto são não lineares. Também é considerado mal posto, pois é ambíguo, ou seja, várias combinações de camadas e suas impedâncias podem gerar o mesmo resultado no dado sísmico. Essas características geram a necessidade de alta intervenção de especialistas para restringir o resultado (regularização). Quando poços são perfurados, são utilizadas outras ferramentas para observar dados mais próximos aos reais. Os principais métodos de inversão utilizam os dados de poços para regularização e geração de estatísticas.

No processo de inversão é importante a modelagem da incerteza dos resultados obtidos, ou seja, obter soluções equivalentes respeitando os dados medidos e as informações *a priori*.



Essas soluções devem representar bem a região onde o erro é menor que a tolerância desejada, chamada de região de equivalência (Tompkins; Fernández Martínez e Fernández Muñoz, 2013). Existem várias razões para essa região de equivalência existir, incluindo erros nas medidas, cobertura dos dados sobre as localizações desejadas, limitação de banda do pulso sísmico, suposições sobre o modelo físico (e.g. isotropia, modelo de camadas, homogeneidade, etc.) e aproximações matemáticas do modelo direto. A modelagem de incerteza ajuda na avaliação dos riscos envolvidos nos processos de perfuração, exploração e produção de óleo e gás.

As médias e variâncias locais das soluções não são suficientes para caracterizar a incerteza na inversão, pois os modelos de impedância geralmente são utilizados como entrada de processos como a simulação de fluxo. Esse tipo de simulação é essencialmente uma função não linear, ou seja, a média das simulações de fluxo não é necessariamente igual a simulação da média das impedâncias:  $F(\overline{\mathbf{m}}) \neq \overline{F(\mathbf{m})}$ , onde  $\mathbf{m}$  representa as soluções aceitáveis,  $F(\cdot)$  a função não linear da simulação de fluxo e  $\overline{(\cdot)}$  a operação de média. Amostragem, de uma forma geral, possibilita análises mais sofisticadas da distribuição posterior (Hansen; Cordua e Mosegaard, 2012).

Na literatura recente, diversos métodos foram propostos para tentar solucionar o problema da inversão sísmica com modelagem de incerteza. Modelos envolvendo técnicas de inteligência computacional foram propostos para tentar resolver o problema (Sen e Stoffa, 1991; Mallick, 1995; Ma, 2002; Artun e Mohaghegh, 2011; Fernández Martínez et al., 2012; Sambridge, 2013). Apesar de terem oferecido bons resultados, esses modelos baseados em otimização não capturam de forma ideal o conhecimento do especialista. A incerteza resultante desses métodos é relacionada com o nível em que foi explorada a superfície de erro, ao invés da incerteza relacionada aos dados e conhecimentos *a priori* inseridos.

O uso de ferramental probabilístico baseado no teorema de Bayes pode ser aplicado, em conjunto com o amostrador de Gibbs, para gerar realizações estocásticas da distribuição *a posteriori* (Figueiredo et al., 2013). Assim, pode-se computar estatísticas do conjunto de soluções, obtido à alto custo computacional. Para atacar a maldição da dimensionalidade e diminuir o custo da amostragem de soluções, foram propostos métodos que utilizam redução dimensional (Tompkins et al., 2011). Neste caso, aplicando análise de componentes principais (PCA) é possível reduzir a dimensão do problema e utilizar um esquema de amostragem determinística, ao custo de perda na resolução espacial das amostras.

A área de Geoestatística teve início a partir do uso de ferramentas estatísticas para tratar dados espaciais, modelar suas continuidades e lidar com incertezas presentes nos problemas da Geologia (Isaaks e Srivastava, 1989, p. 3). Antes de sua criação, um atributo de rocha, por

exemplo, era tratado como se fosse uma só variável aleatória. Ou seja, os valores desse atributo nos pontos espaciais de uma área eram tratados como amostras de uma mesma distribuição de probabilidade. A Geoestatística passou a assumir a característica espacial dos dados, onde cada posição da área é tratada como uma variável aleatória possuindo correlações com suas vizinhas.

Soares; Diet e Guerreiro (2007) propõem um método de inversão, chamado *Global Stochastic Inversion (GSI)*, baseado em *Direct Sequential Simulation (DSS)*, uma importante ferramenta da Geoestatística. Na DSS são utilizados modelos de continuidade direcional dos dados (variogramas direcionais, ou funções de correlação) para simular pontos não observados da área, respeitando assim a continuidade espacial, a variabilidade e o histograma observados nos dados de poços. O processo de inversão faz uso de DSS para encontrar um conjunto de soluções que tenha alta correlação com o dado sísmico. Apesar da vantagem de se modelar a continuidade espacial em quaisquer direções desejadas, o método foi concebido para trabalhar em áreas pequenas, pois assume a estacionariedade das médias. Além disso, não converge em problemas mais avançados como a inversão para dados de porosidade (Soares; Diet e Guerreiro, 2007).

## 1.3 Simulação Multiponto

## 1.4 Redes Neurais Convolucionais

Nesta seção a convolução será descrita. Serão apresentados os principais conceitos relacionados às redes neurais convolucionais e sua estrutura e as principais aplicações deste modelo de aprendizagem de máquina. Um ponto não abordado nesta seção é como escolher a arquitetura de uma rede convolucional.

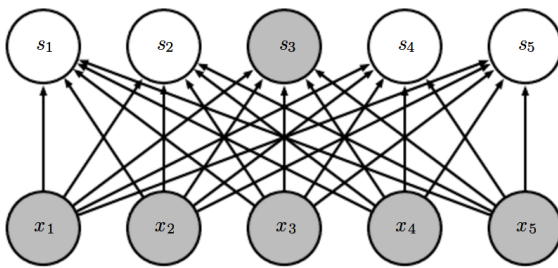
As redes convolucionais, também chamadas de redes neurais convolucionais (CNN), são um tipo de rede neural especializada em processamento de dados que possuam uma topologia conhecida e em forma de grade. Exemplos deste tipo de dado são as séries temporais, que podem ser vistas como uma grade em uma dimensão (1D) com amostras em intervalos de tempo regulares, e dados de imagem, que podem ser pensados como uma grade 2D de *pixels*. As redes convolucionais são um tipo de rede neural que usa a operação de convolução no lugar de multiplicação de matrizes em pelo menos uma de suas camadas.

### 1.4.1 Convolução

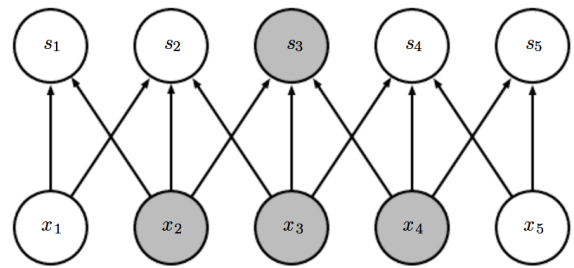
A operação de convolução costuma ser denotada com um asterisco (Eq. 1.6). Na Equação. 1.6,  $x$  refere-se ao conjunto de imagens de entrada, uma sequência multidimensional de dados, e  $w$  é denominado *kernel* ou filtros, uma sequência multidimensional de parâmetros otimizados pelo algoritmo de aprendizagem.

$$s(t) = (x * w)(t) \quad (1.6)$$

A convolução se sustenta sobre três pilares: interações esparsas, compartilhamentos de parâmetros e representações equivalentes. As redes neurais tradicionais utilizam a multiplicação de matriz por uma matriz de parâmetros para descrever a interação entre cada unidade de entrada e cada unidade de saída. Isso significa que toda unidade de saída interage com toda unidade de entrada. As redes convolucionais, por outro lado, tipicamente possui interações esparsas, também chamadas de conectividade esparsa ou pesos esparsos. Para isto, é necessários que os filtros sejam menores que a entrada. De um ponto de vista prático, no processamento de uma imagem, a imagem de entrada pode ter milhares de pixels, entretanto, é possível detectar apenas pequenas regiões de características importantes, por exemplo com significado geológico, com filtros que compreendam apenas algumas dezenas ou centenas de pixels na imagem. Como consequência, menos parâmetros são armazenados e aumenta a eficiência estatística do modelo. As figuras 1.3a e 1.3b ilustram os modelos citados anteriormente. É possível notar que o número de elementos que afetam elemento de saída em destaque ( $s_3$ ) é definido pela convolução com filtro de largura 3 (figura 1.3b), por outro lado  $s_3$  é afetado por todos os elementos da entrada quando formado por multiplicação matricial (figura 1.3a).



(a) Conectividade tradicional.



(b) Conectividade esparsa.

O **compartilhamento de parâmetros**, também chamado de **pesos amarrados** em uma rede convolucional se refere ao uso do mesmo parâmetro para mais de uma função no modelo. Nas redes neurais tradicionais, cada elemento da matriz de pesos é usado apenas uma vez quando a saída da camada é calculada, pois é multiplicado por apenas um elemento da entrada. No

compartilhamento de pesos, o valor do peso aplicado a uma entrada está relacionado ao valor de um peso aplicado em algum outro local. Na rede convolucional, cada elemento do filtro é usado em toda posição da entrada, de modo que, ao invés de aprender um conjunto separado de parâmetros toda localização da imagem, apenas um conjunto é aprendido.

### 1.4.2 Pooling

Uma camada em uma rede convolucional consiste de três estágios. No primeiro estágio, a camada realiza diversas convoluções para produzir um conjunto de ativações lineares. O segundo estágio é chamado etapa de detecção, na qual cada ativação é submetida a uma função não-linear. A terceira etapa é chamada de *pooling*, responsável por modificar a saída para o resumo estatístico das saídas em uma determinada vizinhança. A operação de *pooling* permite tornar invariante pequenas translações no conjunto de entrada, ou seja, ainda que haja pequenas translações na entrada, os valores da maioria das saídas após a o *pooling* permanecem iguais. A figura 1.4 ilustra o funcionamento da função de *pooling*.

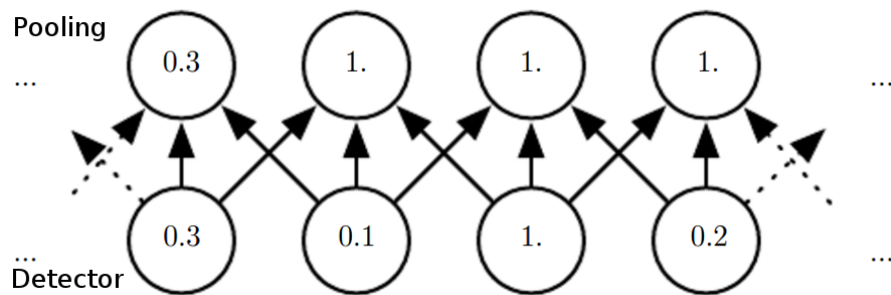


Figura 1.4: Operação de *pooling* com região de tamanho 3. Nesta operação é selecionado o máximo valor de ativação da etapa de detecção.

## 1.5 Objetivo

O objetivo do presente trabalho consiste em propor um modelo capaz de integrar a atualização Bayesiana no processo de inversão GSI. Resultados prévios indicam que utilizar ...

Outro objetivo é ...

## 1.6 Organização do Texto

Este documento está organizado da seguinte forma. Após esta breve introdução, o Capítulo 2 apresenta o estado da arte em modelos de inversão sísmica com modelagem de incerteza. O Capítulo 3 trata da proposta do projeto e resultados preliminares da integração de atualização Bayesiana em inversão Geoestatística, a ser desenvolvido em parceria com o CERENA - Centro de Recursos Naturais e Ambiente do Instituto Técnico da Universidade de Lisboa, sob orientação do professor Dr. Amílcar Soares, especialista em Geoestatística e inversão sísmica. O estágio será iniciado em julho de 2015 e terá 12 meses de duração. Após o retorno ao Brasil, estão planejados mais 8 meses de trabalho para finalizar a escrita da tese e defesa.

## Capítulo 2

# Modelos de Super-resolução

Numa configuração de experimentos físicos se tem o espaço dos parâmetros do modelo e o espaço das medidas. No contexto da sismica de reflexão, os dados sísmicos  $d$  são representados no espaço das medidas e a propriedade de impedância acústica  $m$  das rochas é representada no espaço do modelo.

A princípio parece descomplicada a utilização de equações físicas que descrevem o sistema de forma inversa, fazendo o mapeamento do espaço das medidas ao espaço do modelo. No entanto, esses métodos de inversão direta sofrem de instabilidades devido à ruído e características do problema (Sen, 2006, p. 50). Outra opção é utilizar tentativa e erro para ajustar os parâmetros até conseguir uma resposta semelhante aos dados experimentais. Formalmente isto é automatizado utilizando métodos de otimização. Para tanto, é preciso definir uma função de custo, ou função objetivo, que mede o ajuste dos dados produzidos pelos parâmetros do modelo (dado sintético) ao dado medido.

O objetivo da inversão, no entanto, vai além de encontrar os parâmetros que melhor se ajustam aos dados. Quando os dados são ruidosos, o modelo direto não é exato e não existem dados suficientes, a inversão tem solução não única, ou seja, vários modelos ajustam aos dados de forma equivalente. Consequentemente é importante modelar a incerteza envolvida no processo, indicando qual a variabilidade dos modelos que se ajustam bem aos dados.

A relação entre o modelo e os dados (modelo direto) é dada por:

$$d = G(m_v) + e \quad (2.1)$$

onde  $G(\cdot)$  é uma função não linear, e assume-se que um ruído  $e$  está presente. Em teoria o ruído é uma interferência aleatória que não se tem controle, na prática se considera ruído tudo que

não é explicado pela função  $G$ , e.g. imprecisões no modelo físico e problemas com filtragem e processamento dos dados.

## 2.1 Inversão Sísmica Linear e Não Linear

O modelo mais utilizado para aproximar a função  $G$ , no caso da inversão sísmica, é o modelo convolucional. No caso discreto a convolução é dada pelo produto acumulado de cada amostra do vetor de refletividades por todas as amostras da *wavelet*. Portanto pode ser representado por uma operação matricial:

$$\mathbf{d} = \mathbf{G}\mathbf{r} \quad (2.2)$$

onde  $\mathbf{G}$  é uma matriz convolucional construída utilizando uma *wavelet* e  $\mathbf{r}$  o vetor das refletividades definido por:

$$r(t) = \frac{z(t + \delta t) - z(t)}{z(t + \delta t) + z(t)} \quad (2.3)$$

o que torna não linear a relação entre impedância  $\mathbf{z}$  e o dado sísmico. Uma aproximação válida quando valores de refletividades não ultrapassam 0.3 é:

$$r(t) = \frac{1}{2} \Delta \ln(z(t)) \quad (2.4)$$

Utilizando estas aproximações para o modelo direto, a alternativa mais objetiva é incorporar as aproximações na matriz  $\mathbf{G}$  e invertê-la para obter  $\ln(\mathbf{z})$ , por fim aplicar o exponencial para obter  $\mathbf{z}$ . Neste caso temos os seguintes problemas: existência; unicidade; estabilidade; e robustez (Sen, 2006, p. 56-57). Utilizando a formulação de mínimos quadrados também é possível resolver sistemas sobredeterminados, solucionando problemas com a melhor estimativa possível no sentido de minimizar o erro quadrático (norma  $L_2$ ). Apesar de ser uma solução mais geral, ainda é utilizado somente um critério de ajuste aos dados, o que não possibilita a inserção de conhecimento *a priori*. É possível regularizar o método de mínimos quadrados, mas ainda não se tem muita liberdade para inserir conhecimentos *a priori* e de outras fontes (Clapp; Clapp e Biondi, 2005).

Quando não é possível o uso da aproximação da Equação 2.4, o problema deve ser abordado utilizando métodos de otimização não linear. Com isso os erros devido às aproximações do modelo *forward* diminuem, mas a otimização se torna mais custosa. Como a relação entre os

dados e os parâmetros é não linear, a função objetivo a ser minimizada irá possuir mínimos locais, tornando necessário o uso de métodos de otimização global. Esta prática está bem documentada na literatura de inversão, como o uso de *simulated annealing* (Ma, 2002), de algoritmos genéticos (Mallick, 1995) e enxame de partículas (Zhe e Hanming, 2013).

Outra forma de inversão presente na literatura é a elástica (Azevedo et al., 2013; Buland e Omre, 2003 a). Nesse tipo de inversão os dados sísmicos estão em um nível diferente de processamento onde os traços sísmicos são empilhados em intervalos de ângulos de incidência. Com isso é possível inverter para Vp (velocidade primária/compressional), Vs (velocidade secundária/cisalhante) e  $\rho$  (densidade), ao invés de somente impedância acústica. A Equação 2.5 modela a relação da refletividade  $c_{pp}$  com Vp ( $\alpha$ ), Vs ( $\beta$ ) e  $\rho$  para cada ângulo disponível.

$$c_{pp}(\theta) = a_{\alpha}(\theta) \frac{\Delta\alpha}{\bar{\alpha}} + a_{\beta}(\theta) \frac{\Delta\beta}{\bar{\beta}} + a_{\rho}(\theta) \frac{\Delta\rho}{\bar{\rho}} \quad (2.5)$$

onde:

$$a_{\alpha}(\theta) = \frac{1}{2}(1 + \tan^2 \theta), \quad (2.6)$$

$$a_{\beta}(\theta) = -4 \frac{\bar{\beta}^2}{\bar{\alpha}^2} \sin^2 \theta, \quad (2.7)$$

$$a_{\rho}(\theta) = \frac{1}{2} \left( 1 - 4 \frac{\bar{\beta}^2}{\bar{\alpha}^2} \sin^2 \theta \right). \quad (2.8)$$

Adicionalmente  $\bar{\alpha}$ ,  $\bar{\beta}$  e  $\bar{\rho}$  são as respectivas médias sobre a interface;  $\Delta\alpha$ ,  $\Delta\beta$  e  $\Delta\rho$  são os contrastes e  $\theta$  o ângulo médio de reflexão. Essas propriedades são importantes, pois a velocidade secundária é indicador de hidrocarbonetos por não se propagar em meio líquido, desta forma áreas com presença destes compostos se destacam numa imagem de Vs, podendo indicar interfaces rocha/óleo e água/óleo. As metodologias para resolução do problema são semelhantes, mas neste caso aumentando a dimensão dos dados e números de parâmetros a serem estimados (Buland e Omre, 2003 a).

### 2.1.1 Máximo *a posteriori*

A inversão por Máximo *a posteriori* (MAP) (Buland e Omre, 2003 a; Figueiredo et al., 2014) é realizada para cada traço individualmente. Baseado no modelo convolucional e assumindo que o ruído presente nos dados é Gaussiano, o vetor das sísmicas experimentais  $\mathbf{d}$ , é modelado pela distribuição de probabilidade:



$$p(\mathbf{d}|\boldsymbol{\mu}_d, \boldsymbol{\Sigma}_d) = N(\boldsymbol{\mu}_d, \boldsymbol{\Sigma}_d), \quad (2.9)$$

onde  $\boldsymbol{\mu}_d = \mathbf{G}\mathbf{m}$  é o vetor com a s smica s ntetica e  $\boldsymbol{\Sigma}_d$    a matriz de covari ncia do ru do da s smica, a qual   definida conforme a confiabilidade que o especialista tem no dado s smico ou seu n vel de ru do. Geralmente se utiliza uma matriz diagonal com mesma vari ncia para todos os elementos.

Para o vetor modelo  $\mathbf{m}$  com o logaritmo natural da imped ncia ac stica, considerou-se tamb m uma distribui   normal:

$$p(\mathbf{m}|\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = N(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m), \quad (2.10)$$

no qual  $\boldsymbol{\mu}_m$    um vetor contendo a baixa frequ ncia do logaritmo natural da imped ncia. Este dado   outra informa  o adicional que   fornecida pelo especialista via an lise de velocidades ou interpolando dados de po os por Krigagem. Os componentes da matriz de covari ncia  $\boldsymbol{\Sigma}_m$  foram definidos conforme (Figueiredo et al., 2014):

$$\nu_{t,t'} = \sigma_m^2 \exp\left(-\frac{(t-t')^2}{L^2}\right), \quad (2.11)$$

que define a correla  o entre as componentes de  $\mathbf{m}$  no tempo  $t$  e  $t'$ , na qual  $\sigma_m^2$    a vari ncia da imped ncia ac stica calculada nos dados de po os sem a baixa frequ ncia, e  $L$    a dist ncia de correla  o vertical a ser imposta ao resultado.

Neste arcabou o a m dia e vari ncia posterior para cada tra o podem ser calculadas analiticamente via (Figueiredo et al., 2014):

$$\boldsymbol{\mu}_{m|} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} (\mathbf{d}_o - \mathbf{G} \boldsymbol{\mu}_m), \quad (2.12)$$

$$\boldsymbol{\Sigma}_{m|} = \boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} \mathbf{G} \boldsymbol{\Sigma}_m. \quad (2.13)$$

onde o c lculo da matriz inversa acima pode ser aproveitado para v rios tra os de uma regi o de interesse em certos casos, ou seja, quando as matrizes de covari ncia possam ser assumidas iguais para todos os tra os da s smica da regi o. Desta forma alteram-se a s smica  $\mathbf{d}_0$  e a baixa frequ ncia  $\boldsymbol{\mu}_m$  obtendo-se a m dia posterior para o tra o desejado.

Utilizando esta metodologia a solu  o   representada por uma distribui  o posterior Gaussiana com express es expl citas para o valor esperado e para a covari ncia. N o s o necess rias itera  es para ajuste do modelo, tornando o m todo eficiente e  til em casos de uso reais, pos-

sibilitando o especialista alterar parâmetros e avaliar o resultado em tempo real numa pequena área. Satisfeito com a parametrização, o método é aplicado a todo o campo. A matriz de covariância posterior indica a incerteza presente no resultado, não é necessário definir a tolerância de ajuste aos dados explicitamente, mas é preciso definir a matriz de covariância *a priori* do resultado esperado, ou seja, é preciso ter conhecimento, mesmo que de forma grosseira, das correlações espaciais e variâncias que se espera do resultado. Ao final, a covariância posterior é calculada.

Uma desvantagem destacada na literatura é a dificuldade em inserir modelos de continuidade mais abrangentes, pois é necessário incluir as covariâncias entre traços vizinhos, aumentando as matrizes de covariância e a matriz a ser invertida. Inserindo as correlações horizontais também inviabiliza aproveitar o cálculo da matriz inversa, pois desta forma o resultado da inversão precisa ser calculado para todos os pontos a serem invertidos ao mesmo tempo. Atualmente é inserida somente covariância entre amostras no mesmo traço, ou seja, somente na direção vertical.

## 2.2 Modelagem de Incerteza

Novas propostas surgiram recentemente para tentar encontrar modelos  $\mathbf{m}_f \in \mathbf{M} \subset \mathbf{R}^N$ , onde  $\mathbf{M}$  é o conjunto de modelos admissíveis que atendem o conhecimento *a priori* e são consistentes com o domínio do problema (Tompkins et al., 2011). Para o problema da incerteza, o objetivo é encontrar o conjunto de modelos pertencentes a  $\mathbf{M}$  que se ajustam aos dados observados com uma certa tolerância (*tol*):

$$\|G(\mathbf{m}) - \mathbf{d}\|_p \leq tol \quad (2.14)$$

onde  $\|\cdot\|_p$  é uma norma  $L_p$  de erro escolhida. Na prática esse parâmetro cria um vale plano na função de erro (Fernández Martínez; Fernández Muñiz e Tompkins, 2012), como ilustrado em uma dimensão na Figura 2.1.

A justificativa para o uso desse nível de tolerância é evitar o ajuste dos dados ao ruído. No caso do erro atingir o mínimo, o modelo está tentando explicar o ruído presente nos dados medidos. Apesar da Figura 2.1 ser uma simplificação, consegue-se perceber que ao adicionar a tolerância um conjunto de modelos passa a ser aceitável. Sem a tolerância um só modelo seria eleito como melhor e não teríamos a estimativa da incerteza, apesar de se ajustar melhor aos dados. Este parâmetro é definido empiricamente e deve refletir a incerteza e o ruído presente

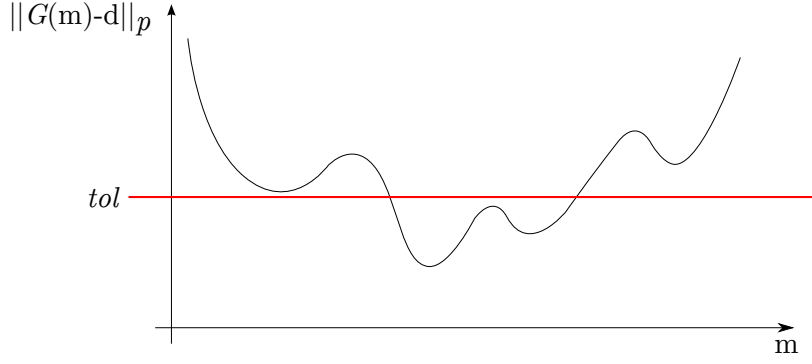


Figura 2.1: Efeito da tolerância na função de erro

nos dados.

Na técnica MAP, citada anteriormente na Seção 2.1.1, além de retornar a matriz de covariância indicando a incerteza presente no resultado, não é necessário definir a tolerância explicitamente, mas é preciso definir a matriz de covariância *a priori* do resultado esperado, ou seja, é preciso ter conhecimento, mesmo que de forma grosseira, das correlações espaciais e variâncias que se espera do resultado. Ao final, a covariância posterior é calculada e o conhecimento *a priori* é atualizado com os dados. Uma desvantagem importante é a dificuldade em inserir modelos de continuidade mais abrangentes, pois para tanto é necessário incluir as covariâncias entre os pontos a serem estimados. Atualmente é inserida somente covariância entre amostras no mesmo traço, ou seja, somente na direção vertical. Inserir covariância em outras direções aumenta o número de elementos da matriz de covariância de forma quadrática, tornando o processo custoso.

Utilizando uma abordagem estocástica Bayesiana, Figueiredo et al. (2013) leva em consideração a incerteza na *wavelet* gerando amostras das distribuições posteriores da impedância acústica e da *wavelet*. O problema continua linearizado e com suposição Gaussiana, mas é preciso utilizar um esquema de amostragem baseado em simulação *Markov-chain Monte Carlo* (MCMC) via algoritmo de Gibbs, pois não é possível obter as distribuições posteriores analiticamente. Com realizações da distribuição posterior, é possível calcular as estatísticas de interesse e modelar a incerteza envolvida. A desvantagem é o alto custo computacional para gerar amostras dessas distribuições, principalmente com o aumento da dimensão do problema. Um estudo recente utiliza MCMC para realizar amostragem da propriedades de porosidade e permeabilidade das rochas (Zunino et al., 2014). Estas propriedades são geralmente estimadas em uma etapa posterior à estimativa de propriedades primárias, como a impedância. Portanto o método gera interesse ao realizar inversão de forma mais avançada no fluxo de trabalho, além de

gerar estimativas da incerteza. Apesar de sofrer com o alto custo computacional da amostragem o autor relata bons resultados.

A estimativa de *wavelets* também é objeto de estudo na literatura. O problema consiste em determinar o pulso sísmico que gerou o sinal medido, utilizando os dados de poços como referência. *Wavelets* podem ser consideradas o elo entre os dados sísmicos e as propriedades de rocha. Como podem haver erros de conversão da escala de profundidade dos poços para a escala de tempo dos dados sísmicos, é necessário avaliar a incerteza na estimativa da *wavelet*. Por isso, um esquema de amostragem semelhante, baseado em MCMC, foi utilizado em Buland e Omre (2003 b).

Esforços na área de inteligência computacional foram inicialmente empregados utilizando métodos que possuem componentes estocásticas, mas foram desenvolvidos originalmente para exploração, ou seja, encontrar rapidamente o mínimo global da função objetivo (Sen e Stoffa, 1995). Otimização por enxame de partículas é um exemplo. O método foi modificado para permanecer em sua fase de exploração, favorecendo soluções que ficam abaixo de certa tolerância. Realizando, portanto, o que é chamado de *importance sampling* (Martínez et al., 2010). Os autores dessa metodologia defendem que a função de erro é um bom *proxy* para a distribuição posterior.

Trabalhos anteriores tentaram sem êxito utilizar várias execuções de algoritmos genéticos (GA) para estimar a incerteza (Sen e Stoffa, 1995, p. 152). A hipótese de trabalho era a que ao executar o algoritmo várias vezes, seriam encontrados vários mínimos locais que representariam a incerteza. Mas como demonstrado em Fernández Martínez; Fernández Muñiz e Tompkins (2012), a topografia da função de erro em inversão sísmica possui grandes vales alongados de mínimo global. Com isso os métodos devem ser modificados para capturar essa característica e portanto, a simples execução múltipla não gera boa estimativa da incerteza.

O trabalho de Tompkins et al. (2011) trata de inversão de dados eletromagnéticos, outro tipo de metodologia utilizada na indústria para imageamento de subsuperfície. Neste problema o modelo direto é mais complexo e não linear. Apesar de não utilizar dados sísmicos, a metodologia proposta tem objetivos pertinentes a inversão sísmica. A hipótese é que não é necessário amostrar exaustivamente a distribuição posterior para modelar a incerteza, e sim utilizar amostras representativas. A proposta utiliza redução dimensional com análise de componentes principais (PCA) sobre a matriz de covariância. Esta matriz, por sua vez, é calculada a partir de um conjunto de soluções ou aproximada utilizando informação do gradiente da função objetivo após a execução de um método de inversão não linear. Restrições *a priori* são mapeadas para o espaço

reduzido e a técnica de Smolyak (1963) é utilizada para amostrar de forma determinística e hierárquica este espaço. Retornando ao espaço original, é utilizado um critério de rejeição para eliminar soluções acima de certa tolerância. Os passos de amostragem e rejeição são repetidos, refinando o *grid* de Smolyak, até que haja convergência de estatísticas das amostras aceitas.

Ao reduzir a dimensão do problema se perde resolução espacial das amostras a serem geradas, pois serão considerados somente parte dos primeiros autovetores, ordenados de forma decrescente em relação aos seus autovalores, da matriz de covariância. Ou seja, menores variabilidades são desconsideradas (Jolliffe, 2002, p. 3). Essa perda é justificada por serem desprezíveis em certos casos e pela diminuição do custo computacional ao trabalhar no espaço reduzido. Além do custo de computar os autovetores e autovalores, são adicionados os custos de computar o *grid* de Smolyak e do mapeamento de restrições do espaço original ao espaço reduzido. Esses custos extras se tornam proibitivos quando a dimensão do problema cresce e não é mais possível reduzir a dimensão à centenas de variáveis. Com isso alternativas surgiram para tentar expandir o limite do tamanho do problema, em detrimento da resolução das amostras utilizando decomposição em valores singulares (Tompkins et al., 2011) e utilizando outro esquema de amostragem determinística (Tompkins, 2012). Apesar dos avanços não se sabe se é possível aplicar a metodologia aos problemas de larga escala ( $> 10^5$  variáveis) (Tompkins; Fernández Martínez e Fernández Muñiz, 2013).

A metodologia de Scheidt e Caers (2009) também utiliza redução dimensional, mas baseada em *multi-dimensional scaling* (MDS). Técnica que visa preservar distâncias do espaço original no espaço reduzido. Característica que possibilita o uso de diferentes métricas para modelar distâncias entre modelos. Análise de agrupamento de dados (*cluster analysis*) é utilizada para definir modelos equivalentes no espaço reduzido, selecionando somente poucos modelos para simulação de fluxo de petróleo. Esse fluxo de trabalho é possível pois os autores modelam a incerteza já com o objetivo de verificar sua influência na simulação de fluxo. Essa metodologia permite realizar análise de sensibilidade e integrar outras fontes de incerteza ao processo de inversão, como diferentes *wavelets* ou parâmetros de continuidade espacial. Desta forma, é possível verificar os pontos que mais influenciam na incerteza do modelo final visualizando os modelos em um espaço reduzido e selecionando representativos via agrupamento de dados (Caers, 2011, p. 188).

### **2.2.1 Global Stochastic Inversion**

Soares (2001) propõe o uso de Simulação Sequencial Direta (SSD) utilizando dados de poços

e seus variogramas para realizar amostragem sequencial da distribuição de probabilidades da impedância baseada na média e variância da Krigagem. O autor avança no sentido de amostrar da distribuição global sem precisar fazer transformações para variáveis Gaussianas. Inversão sísmica utilizando SSD, chamada de *Global Stochastic Inversion* (GSI), é proposta em Soares; Diet e Guerreiro (2007), onde a ideia de combinação presente em algoritmos genéticos é utilizada. Regiões locais de simulações estocásticas são combinadas afim de encontrar uma amostra global que maximize a correlação da sísmica sintética com a medida. Como o método utiliza SSD, a continuidade espacial é modelada em qualquer direção definida pelos variogramas.

Pode ser utilizado qualquer tipo de Krigagem para a SSD, incluindo a co-Krigagem que utiliza uma variável secundária  $Z_1$ . Esta variável representa um atributo ou outra propriedade que possui correlação com a variável primária  $Z_2$  a ser simulada, por exemplo, o resultado de um método determinístico. A SSD utilizando co-Krigagem é denominada co-SSD. A estimativa de co-Krigagem simples para a variável primária  $Z_2$  é dada por (Soares, 2001):

$$[Z_2(x_u)^*]_{CKS} = \sum_{\alpha=1}^N \lambda_{\alpha} z_2(x_{\alpha}) + \lambda_{\beta} z_1(x_u)^* \quad (2.15)$$

onde  $x_u$  são as coordenadas dos pontos onde se quer estimar e  $x_a$  são as coordenadas dos pontos onde possuem dados amostrados ou que já foram simulados,  $\lambda$  são os pesos calculados pela Krigagem (Goovaerts, 1997, p. 169). Para calcular os pesos da co-Krigagem é necessário definir uma correlação da variável primária com a secundária, informação retirada dos dados de poço ou informada por um especialista e que pode ser global, todos os pontos de  $Z_1$  tem a mesma correlação com  $Z_2$ , ou local onde cada ponto possui sua correlação entre as variáveis primária e secundária. A variância da co-Krigagem simples é dada por:

$$\sigma_{CKS}^2(x_u) = Var\{Z_2(x_u)^* - Z_2(x_u)\} \quad (2.16)$$

A variância e média da Krigagem são então utilizadas para amostrar a distribuição global  $F_{Z_2}$  via transformação local Gaussiana como demonstrado na Figura 2.2.

Com esta transformação o autor demonstra que é possível amostrar a distribuição global utilizando amostras de Gaussianas equivalentes a  $G(y(x_u)^*, \sigma_{SK}^2(x_u))$ , ou seja, centrada nas médias e com as variâncias locais da Krigagem. Esta abordagem se diferencia da *Sequential Gaussian Simulation* (SGS) por não realizar a transformação da distribuição global para uma variável Gaussiana, o que pode acarretar problemas em certos casos (Soares, 2001).

O processo de inversão GSI inicia-se sem imagem secundária, efetuando uma SSD somente

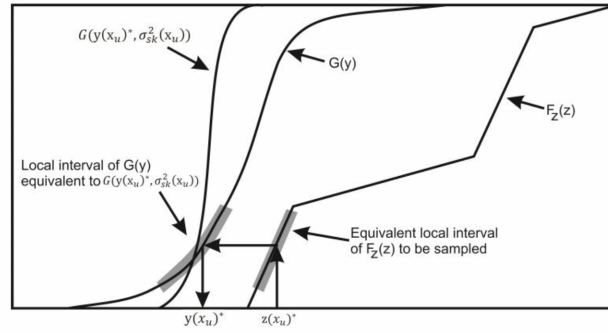


Figura 2.2: Amostragem global via transformação local feita pela SSD (Soares, 2001)

utilizando os variogramas e dados de poço. As melhores regiões dos dados simulados são selecionadas baseado na correlação da sísmica sintética com a original. Uma imagem auxiliar é então criada com essas melhores regiões e utilizada como variável secundária na próxima iteração. Os passos da inversão GSI estão resumidos a seguir:

1. Gerar um conjunto de inicial de imagens de impedância acústica utilizando SSD
2. Calcular a sísmica sintética efetuando a convolução das refletividades, calculadas a partir das impedâncias, com uma *wavelet* conhecida
3. Avaliar o casamento das sísmicas sintéticas com a real utilizando correlação local
4. Ordenar as impedâncias por melhor casamento de sua sísmica sintética e selecionar delas os melhores traços de impedância. Com os melhores traços, compor uma imagem auxiliar para o próximo passo
5. Gerar um novo conjunto de imagens utilizando co-SSD e retornar ao passo 2 até que um critério de convergência seja atingido

Uma limitação dos métodos geoestatísticos é a necessidade de transformar o espaço físico para um espaço deposicional, ou seja, levar em consideração que ao longo do tempo a estrutura em que foram depositados os sedimentos foi deformada por falhas e dobramentos. No espaço físico são modelados os processos de propagação de ondas e fluxo, por exemplo. No espaço deposicional são modeladas as propriedades das rochas. Para contornar essa limitação é feita uma transformação onde são utilizados dados chamados de horizontes, criados pelo especialista para acompanhar as deformações sofridas pelas camadas utilizando uma superfície interpretada na sísmica. Desta forma assume-se que quando as camadas de sedimentos foram depositadas

elas eram planas e os dados são transformados horizontalizando as superfícies (Caers, 2011, p. 140). O processo é ilustrado na Figura 2.3. A transformação deve ser aplicada aos dados de poços e à variável secundária antes de modelar os variogramas e efetuar a SSD, e revertida para visualizar os resultados e calcular a símica sintética (Equações 2.3 e 2.2). Esta transformação também é chamada de conversão para grade estratigráfica.

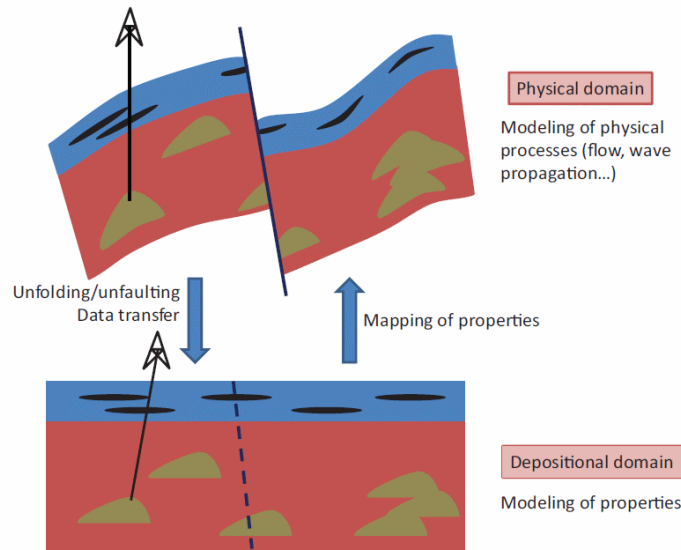


Figura 2.3: Processo de horizontalização (Caers, 2011, p. 143)

## 2.3 Resumo

Neste capítulo foi revisado o estado da arte em inversão sísmica acústica com modelagem de incerteza. Pontos críticos dos métodos foram considerados e identificados para pesquisa futura. O próximo capítulo irá definir a proposta de pesquisa, apresentar o plano de trabalho e concluir com as perspectivas de contribuição.



## Capítulo 3

# Super-Resolução de Imagens Pós-Inversão

Este trabalho propõe desenvolver um novo modelo para inversão sísmica acústica e elástica com modelagem de incerteza. No capítulo anterior foi apresentado um método que trata o problema utilizando amostragem da distribuição posterior via simulação sequencial direta (Soares; Diet e Guerreiro, 2007). Outro método apresentado utiliza o cálculo do máximo *a posteriori* (MAP) para determinar o resultado mais provável (Buland e Omre, 2003 a).

A presente proposta tem como objetivo aumentar a eficiência do método GSI (Soares; Diet e Guerreiro, 2007) afim de evitar amostrar soluções que não estejam dentro da distribuição posterior. Neste caso o termo distribuição posterior é utilizado para denotar o conjunto de soluções que é consistente com os conhecimentos *a priori* e com os dados sísmicos. Para melhorar a eficiência do método, mantendo sua capacidade de modelagem da continuidade espacial direcional, serão utilizadas resultados de Buland e Omre (2003 a) para guiar a amostragem via imagem secundária.

Antes de efetuar a SSD, os dados de poços e resultado do MAP são submetidos à um filtro passa alta com frequência de corte de  $8Hz$ . Desta forma a simulação é efetuada somente para as frequências onde a sísmica tem amplitude significativa, assumindo o mesmo modelo para as baixas frequências ( $< 8Hz$ ) que foi utilizado para o cálculo do MAP. Após a simulação, o modelo de baixa frequência é adicionado aos resultados para cálculo das sísmicas sintéticas e correlações.

O método proposto necessita dos seguintes dados de entrada: impedância acústica medida nos poços na escala da sísmica, *wavelet*, modelo de baixa frequência para impedância acústica ( $< 8Hz$ ), matriz de covariância da sísmica, distância de correlação vertical  $L$  (Equação 2.11),

modelo de horizontes e variogramas verticais e horizontais modelados nos dados filtrados ( $> 8Hz$ ) para simulação sequencial. Resumindo, os principais passos são:

1. Preparar dados de entrada e modelar variogramas
2. Inversão pelo método MAP utilizando as matrizes de covariância da sísmica e do modelo, *wavelet* e modelo de baixa (Equação 2.12).
3. Filtragem dos dados de poços e resultado do MAP com filtro passa alta ( $> 8Hz$ )
4. Horizontalização dos dados de poços e resultado do MAP filtrados utilizando o modelo de horizontes fornecido
5. SSD utilizando dados filtrados e horizontalizados com resultado do MAP como imagem secundária
6. Reverter horizontalização das amostras simuladas pelo SSD
7. Adicionar o modelo de baixa frequência às amostras simuladas, calcular refletividades e sísmica sintética. Avaliar os resultados via correlações

O fluxo do método é ilustrado na Figura 3.1:

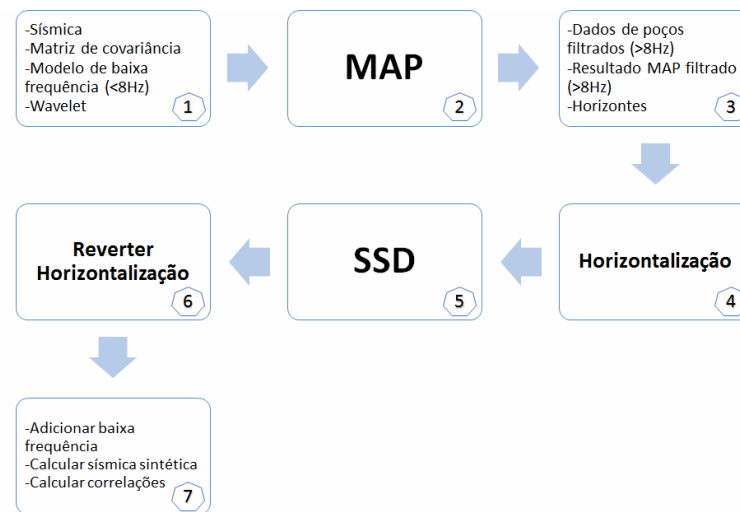


Figura 3.1: Fluxograma do método de inversão e simulação

Utilizando esta metodologia é possível gerar várias realizações da inversão para impedância à um custo menor do que apresentado em (Soares; Diet e Guerreiro, 2007), pois não é mais necessário selecionar melhores amostras e efetuar iterações para realizar a inversão, dado que o resultado do MAP já fornece uma estimativa próxima à média desejada.

### 3.1 Resultados Preliminares

Experimentos realizados com o método para cálculo do MAP obtiveram resultados e tempo de execução comparáveis com métodos implementados na indústria (Figueiredo et al., 2014). Apesar de resultar somente nas médias e variâncias, a parametrização e regularização utilizada na metodologia pode ajudar a restringir a amostragem, de forma a inserir informações da posterior e tornar a amostragem do SSD mais eficiente.

Como a inversão GSI é fundamentada na Simulação Sequencial Direta, os resultados da primeira iteração da GSI são iguais aos resultados de uma SSD utilizando os mesmos parâmetros e dados de entrada. Desta forma é possível verificar a melhora na correlação da sísmica sintética com a sísmica original quando se utiliza o resultado do MAP como imagem secundária para SSD. Este experimento foi realizado e na primeira iteração do GSI a correlação das sísmicas sintética e original foi 0.45 para a média de impedância, com a SSD utilizando o resultado do MAP como imagem secundária foi obtida uma correlação de 0.97 para a média, foram utilizadas populações de 35 amostras para cada método. Mesmo quando são executadas as iterações da GSI, a máxima correlação da média encontrada é 0.9, tomando no mínimo 10 iterações para atingir este nível de qualidade.

Foram utilizados dados fornecidos pela PETROBRAS de um campo real para efetuar os testes. Os dados consistem de 4 poços que se encontram na posição dos traços 17, 210, 409 e 698, modelo de baixa frequência interpolado dos poços, horizontes e uma *wavelet* extraída por um especialista da empresa. O parâmetro de distância de correlação da inversão por MAP foi utilizado  $L = 1.4$ , a variância da sísmica foi estipulada como 0.4 multiplicado pela variância média dos traços. Para a SSD foi utilizado um modelo de variograma horizontal omnidirecional esférico com alcance de 32 traços, ou aproximadamente 800 metros. Para o variograma vertical foi utilizando o mesmo modelo com o alcance de 2 índices de tempo, correspondente a 8ms.

O resultado da inversão MAP é demonstrado na Figura 3.2. A média das realizações amostradas pelo método proposto está na Figura 3.4. Comparando com a Figura 3.3, a qual demonstra a média das realizações da primeira iteração da GSI, verifica-se que utilizar o resultado do MAP como imagem secundária traz a informação da sísmica nos pontos entre dos poços. Comparando a média da GSI após 10 iterações, na Figura 3.5, observa-se a semelhança dos resultados sem a necessidade de realizar as iterações. O método proposto demorou 52s para executar contra 487s para realizar as 10 iterações da GSI.

A metodologia descrita na literatura para a GSI não aplica a filtragem dos dados para

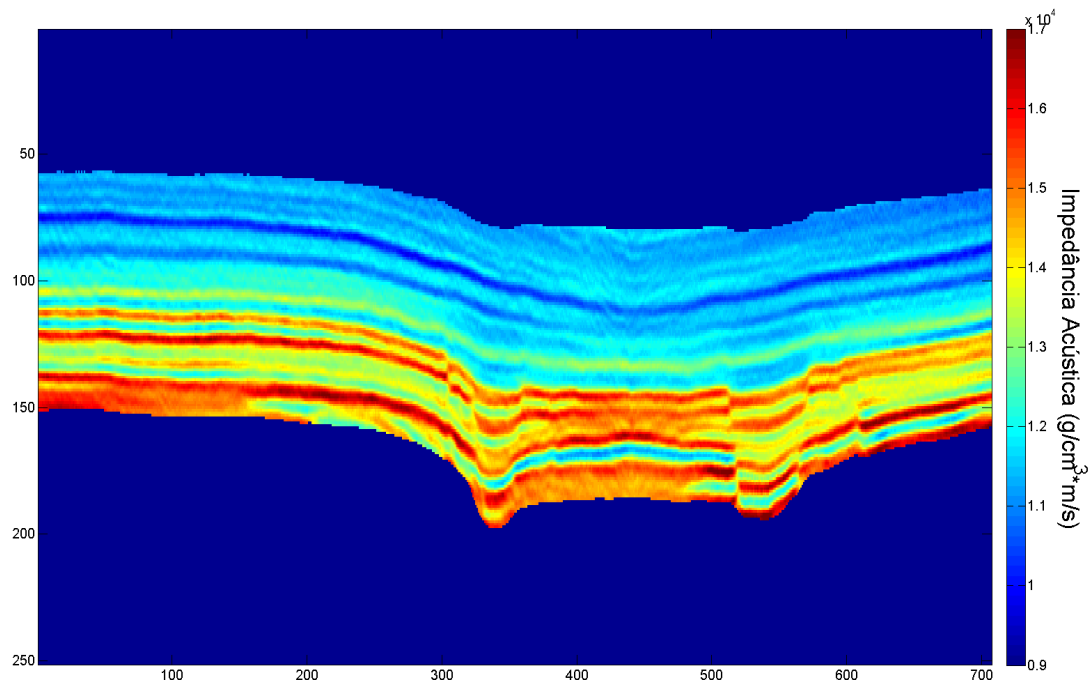


Figura 3.2: Resultado MAP utilizado como imagem secundária

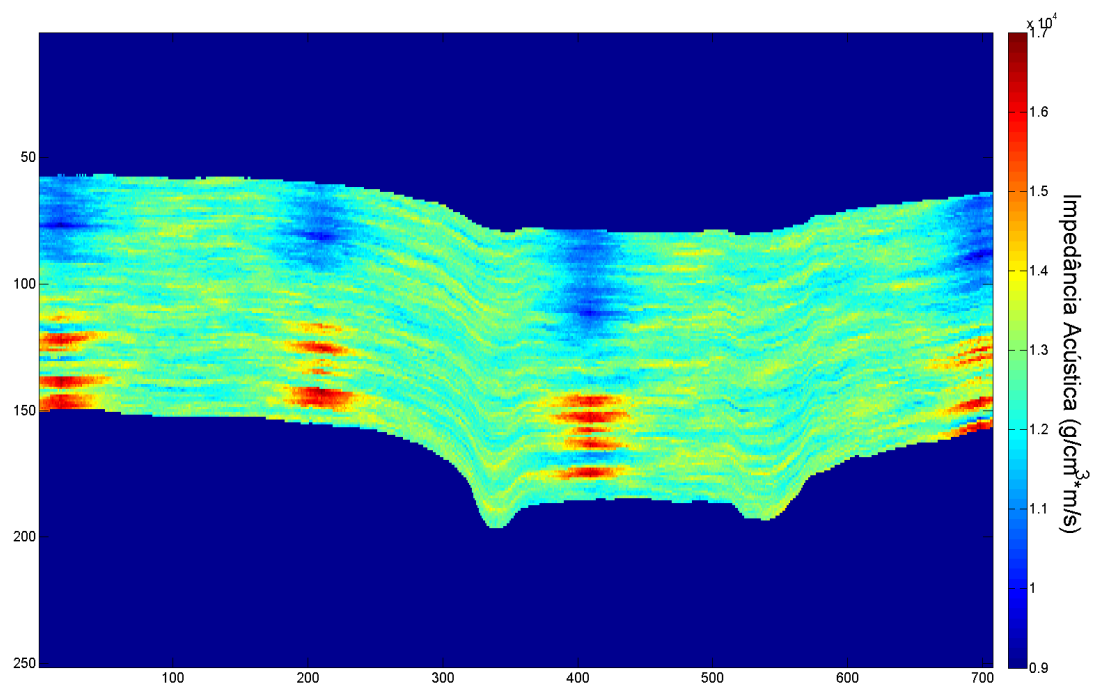


Figura 3.3: Média das amostras da primeira iteração da GSI

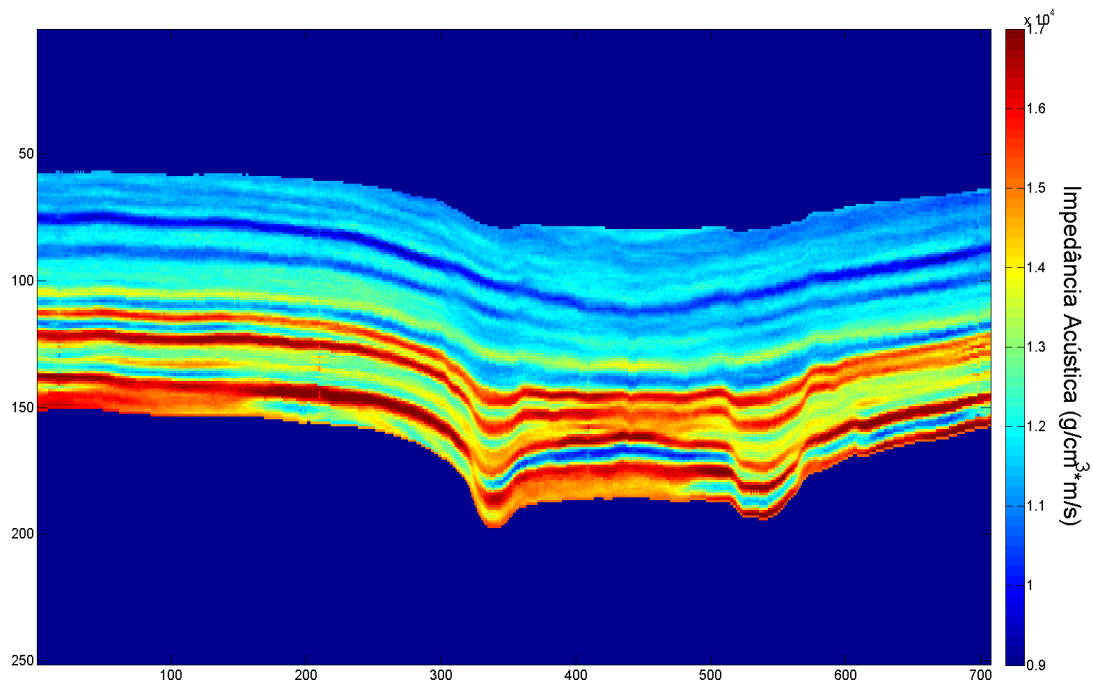


Figura 3.4: Média das amostras da SSD utilizando MAP como imagem secundária

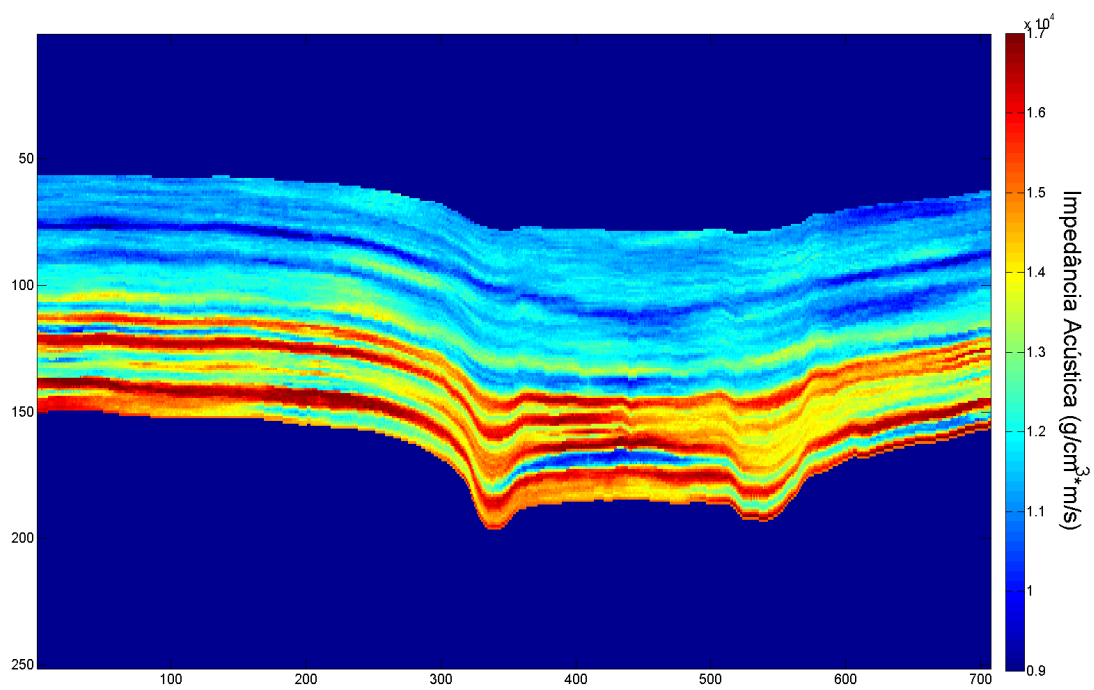
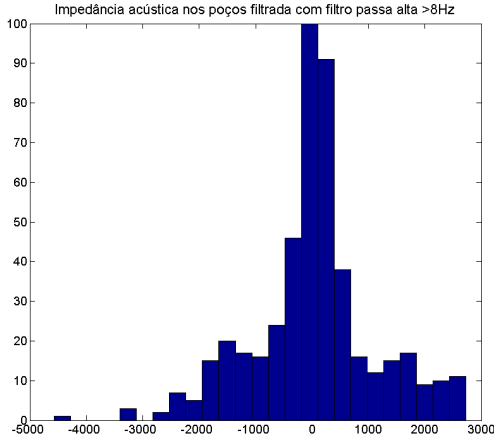
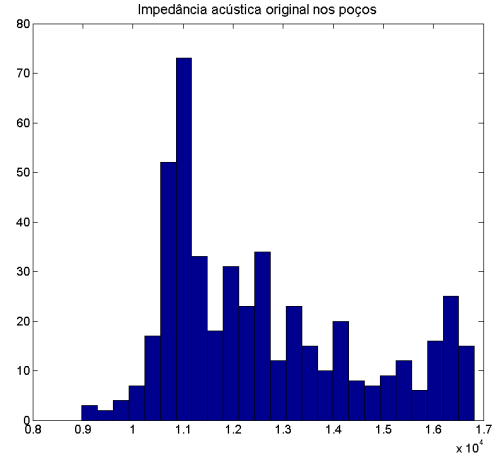


Figura 3.5: Média das amostras após 10 iterações da GSI



(a) Dados filtrados



(b) Dados originais

Figura 3.6: Comparação dos dados de poços filtrados

amostrar somente os resíduos, ou alta frequência. Utilizar o modelo de baixa frequência como informação *a priori* foi identificado como uma regularização importante para agilizar a inversão e amostragem. Principalmente para a SSD, filtrar os dados resulta em uma distribuição global com menor variância, pois assumindo um modelo de baixa frequência são definidas médias locais e as amostras são geradas em torno desta média, respeitando a distribuição do poço filtrado. Sem a filtragem, a informação da média local não é utilizada, amostrando-se toda a distribuição original do poço em todos os pontos da região de interesse. A Figura 3.6 mostra a comparação dos histogramas dos dados filtrados e não filtrados. Portanto, utilizar o modelo de baixa frequência insere um viés no resultado e diminui a incerteza dada pela variância, mas a incerteza referente a escolha do modelo de baixa precisa ser levada em consideração.

## 3.2 Proposta e Plano de Trabalho

A proposta para o restante do projeto é implementar a inversão elástica, que utiliza sísmica pré empilhada para obter as velocidades primária e secundária e densidade na linha de Azevedo et al. (2013) dentro da nova metodologia proposta. Os resultados obtidos serão comparados com os resultados da GSI em termos de variância e qualidade do resultado baseado na metodologia de Coléou (2012), amplamente utilizada para avaliação dos resultados da inversão sísmica na indústria.

A metodologia de Scheidt e Caers (2009) será integrada na proposta para possibilitar a análise da incerteza envolvida com os parâmetros e dados de entrada definidos pelo especialista,

como o modelo de baixa, *wavelets* e horizontes, pois a essas incertezas geológicas são apontadas como maiores influências na incerteza do resultado da inversão (Caers, 2011, p. 133). Para tanto, será proposto um esquema de amostragem com Monte Carlo para selecionar os dados de entrada aleatoriamente a cada amostragem do resultado da inversão. Após obter várias amostras com combinações diferentes de dados de entrada, uma análise baseada em *Multi-Dimensional Scaling* será aplicada para avaliar a sensibilidade ao variar cada parâmetro de entrada na incerteza final do resultado da inversão. Será utilizada como métrica as medidas de *Quality-Anomaly* (Coléou, 2012).

O caráter de originalidade do trabalho é garantido pelas contribuições para a área de inversão com modelagem de incerteza. A primeira contribuição diminui o tempo de execução por eliminar a necessidade de um método iterativo para realizar a inversão geoestatística, ao mesmo tempo modelando a continuidade lateral. A segunda contribuição será modelar as incertezas presentes nos dados de entrada (*wavelets*, horizontes, modelos de baixa, variogramas, etc...) após a integração da metodologia de Scheidt e Caers (2009).

A Figura 3.7 mostra o cronograma mensal das atividades a serem desenvolvidas durante o período de 12 meses do estágio de doutorado sanduíche a ser realizado no Centro de Recursos Naturais e Ambiente (CERENA) do Instituto Superior Técnico (IST) ligado a Universidade de Lisboa sob orientação do Dr. Amílcar Soares. Também estão planejados os meses após o retorno para finalização da tese e defesa.

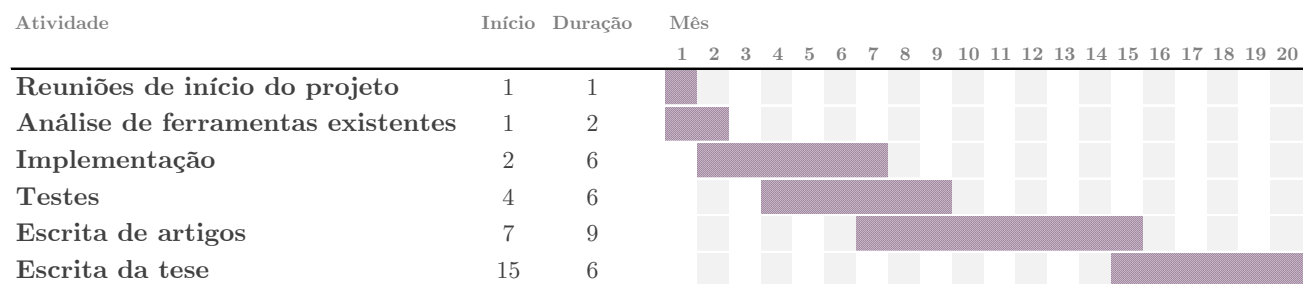


Figura 3.7: Cronograma mensal do projeto a partir de jul/2014

Dentre os periódicos que apresentam trabalhos relacionados com esta área, se destacam os seguintes com suas classificações no Qualis para Ciências da Computação:

- IEEE Trans. on Geoscience and Remote Sensing
- ISSN: 0196-2892 - IEEE - Qualis A1
- Journal of Computational Physics

- ISSN: 0021-9991 - Elsevier - Qualis A1
- Computers & Geosciences
- ISSN: 0098-3004 - Elsevier - Qualis A2
- Computational Geosciences
- ISSN: 1420-0597 - Springer - Qualis B1
- Geophysics
- ISSN: 0016-8033 - Society of Exploration Geophysicists - Qualis A2 Interdisciplinar
- Mathematical Geosciences
- ISSN: 1874-8961 - Springer - Qualis A2 Interdisciplinar

A colaboração com o grupo de pesquisas do Prof. Amilcar Soares é justificada pois o autor é a uma das principais referências na literatura sobre Simulação Sequencial Direta e inversão Geoestatística (Azevedo et al., 2013; Soares; Diet e Guerreiro, 2007; Soares, 2001). A pesquisa nos dois primeiros anos do projeto de Doutorado foi feita sob orientação do Prof. Mauro Roisenberg, coautor de referências relevantes para inversão Bayesiana (Figueiredo et al., 2013, 2014). Desta forma, espera-se cumprir a proposta de integração sob orientação de especialistas em ambos os métodos. A proposta também faz parte de um Termo de Cooperação Petrobras/UFSC/FEESC.



# Bibliografia

- Artun, E. e Mohaghegh, S. (2011), Intelligent seismic inversion workflow for high-resolution reservoir characterization, *Computers & Geosciences* **37**(2), 143 – 157.
- Aster, R.; Borchers, B. e Thurber, C. (2013), *Parameter Estimation and Inverse Problems*, Academic Press, Academic Press.
- Azevedo, L.; Nunes, R.; Soares, A. e Neto, G. (2013), Stochastic seismic AVO inversion, 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013, EAGE, Londres, Reino Unido.
- Buland, A. e Omre, H. (2003 a), Bayesian linearized avo inversion, *Geophysics* **68**(1), 185–198.
- Buland, A. e Omre, H. (2003 b), Bayesian wavelet estimation from seismic and well data, *Geophysics* **68**(6), 2000–2009.
- Caers, J. (2011), *Modeling Uncertainty in the Earth Sciences*, Wiley.
- Clapp, M. L.; Clapp, R. G. e Biondi, B. L. (2005), Regularized least-squares inversion for 3-D subsalt imaging, SEG Annual Meeting, Society of Exploration Geophysicists, Houston, Texas, pp. 1814–1817.
- Coléou, T. (2012), Use of quality and anomaly indicators in the QC of seismic inversion, 74th EAGE Conference and Exhibition - Workshops, EAGE, Copenhagen, Dinamarca.
- Fernández Martínez, J. L.; Fernández Muñiz, M. Z. e Tompkins, M. J. (2012), On the topography of the cost functional in linear and nonlinear inverse problems, *Geophysics* **77**(1), W1–W15.
- Fernández Martínez, J. L.; Mukerji, T.; García Gonzalo, E. e Suman, A. (2012), Reservoir characterization and inversion uncertainty via a family of particle swarm optimizers, *Geophysics* **77**(1), M1–M16.

- Figueiredo, L. P.; Santos, M.; Roisenberg, M. e Neto, G. (2013), Stochastic Bayesian algorithm to a jointly acoustic inversion and wavelet estimation, Society of Exploration Geophysicists, chapter 635, pp. 3273–3277.
- Figueiredo, L. P.; Santos, M.; Roisenberg, M.; Neto, G. e Figueiredo, W. (2014), Bayesian framework to wavelet estimation and linearized acoustic inversion, Geoscience and Remote Sensing Letters, IEEE **PP**(99), 1–5.
- FishSAFE (2014), Seismic surveys, <http://fishsafe.eu/en/offshore-structures/seismic-surveys.aspx>.
- Goovaerts, P. (1997), Geostatistics for Natural Resources Evaluation, Applied geostatistics series, Oxford University Press.
- Hansen, T.; Cordua, K. e Mosegaard, K. (2012), Inverse problems with non-trivial priors: efficient solution through sequential gibbs sampling, Computational Geosciences **16**(3), 593–611.
- Isaaks, E. H. e Srivastava, R. M. (1989), Applied geostatistics, Oxford University Press, New York.
- Jolliffe, I. T. (2002), Principal component analysis, Springer, New York.
- Ma, X. (2002), Simultaneous inversion of prestack seismic data for rock properties using simulated annealing, GEOPHYSICS **67**(6), 1877–1885.
- Mallick, S. (1995), Model-based inversion of amplitude-variations-with-offset data using a genetic algorithm, Geophysics **60**(4), 939–954.
- Martínez, J. L. F.; Gonzalo, E. G.; Muñoz, Z. F.; Mariethoz, G. e Mukerji, T. (2010), Posterior sampling using particle swarm optimizers and model reduction techniques, Int. J. Appl. Evol. Comput. **1**(3), 27–48.
- Sambridge, M. (2013), A parallel tempering algorithm for probabilistic sampling and multimodal optimization, Geophysical Journal International .
- Scheidt, C. e Caers, J. (2009), Representing spatial uncertainty using distances and kernels, Mathematical Geosciences **41**(4), 397–419.

- Sen, M. K. (2006), Seismic Inversion, Society of Petroleum Engineers, Richardson, TX, USA.
- Sen, M. K. e Stoffa, P. L. (1991), Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics* **56**(10), 1624–1638.
- Sen, M. K. e Stoffa, P. L. (1995), Global Optimization Methods in Geophysical Inversion, Vol. 4 of *Advances in Exploration Geophysics*, Elsevier.
- Smolyak, S. (1963), Quadrature and interpolation formulas for tensor products of certain classes of functions, *Soviet Mathematics, Doklady* **4**, 240–243.
- Soares, A. (2001), Direct sequential simulation and cosimulation, *Mathematical Geology* **33**(8), 911–926.
- Soares, A.; Diet, J. e Guerreiro, L. (2007), Stochastic inversion with a global perturbation method, *Petroleum Geostatistics 2007*, EAGE.
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.
- Tompkins, M. J. (2012), Efficient estimation of nonlinear posterior model covariances using maximally sparse cubature rules, *Geophysics* **77**(5), ID1–ID8.
- Tompkins, M. J.; Fernández Martínez, J. L.; Alumbaugh, D. L. e Mukerji, T. (2011), Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: Marine controlled-source electromagnetic examples, *Geophysics* **76**(4), F263–F281.
- Tompkins, M. J.; Fernández Martínez, J. L. e Fernández Muñiz, Z. (2013), Comparison of sparse-grid geometric and random sampling methods in nonlinear inverse solution uncertainty estimation, *Geophysical Prospecting* **61**(1), 28–41.
- Zhe, Y. e Hanming, G. (2013), Non-linear prestack seismic inversion with global optimization using an edge-preserving smoothing filter, *Geophysical Prospecting* **61**(4), 747–760.
- Zunino, A.; Lange, K.; Melnikova, Y.; Hansen, T. e Mosegaard, K. (2014), Reservoir modeling combining geostatistics with markov chain monte carlo inversion, E. Pardo-Igúzquiza; C. Guardiola-Albert; J. Heredia; L. Moreno-Merino; J. J. Durán e J. A. Vargas-Guzmán,

(Eds.), *Mathematics of Planet Earth*, Lecture Notes in Earth System Sciences, Springer Berlin Heidelberg, pp. 683–687.