



**UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE
UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
MESTRADO EM CIÊNCIA DA COMPUTAÇÃO**



CLÁUBIO LANDNEY LIMA BANDEIRA

**MODELAGEM E EXTRAPOLAÇÃO DO MOVIMENTO DE FLUIDOS
EM UM RESERVATÓRIO PETROLÍFERO USANDO MÉTODOS DE
PREVISÃO DE SÉRIES TEMPORAIS BASEADOS EM
CO-REGISTRO DE IMAGENS**

**MOSSORÓ - RN
2012**

CLÁUBIO LANDNEY LIMA BANDEIRA

**MODELAGEM E EXTRAPOLAÇÃO DO MOVIMENTO DE FLUIDOS
EM UM RESERVATÓRIO PETROLÍFERO USANDO MÉTODOS DE
PREVISÃO DE SÉRIES TEMPORAIS BASEADOS EM
CO-REGISTRO DE IMAGENS**

Dissertação apresentada ao Mestrado em Ciência da Computação – associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Marcelino Pereira dos Santos Silva

Co-Orientador: Harold Ivan Ângulo Bustos

**MOSSORÓ - RN
2012**

Catálogo da Publicação na Fonte.

Bandeira, Cláudio Landney Lima.

Modelagem e extrapolação do movimento de fluidos em um reservatório petrolífero usando métodos de previsão de séries temporais baseados em co-registro de imagens . / Cláudio Landney Lima Bandeira. – Mossoró, RN, 2012.

56 f

Orientador(a): Prof. Marcelino Pereira dos Santos Silva

Dissertação (Mestrado em Ciência da Computação). Universidade do Estado do Rio Grande do Norte.

1. Fluidos - Dissertação. 2. Lapso de tempo - Dissertação. 3. Previsão de séries temporais - Dissertação. I. Silva, Marcelino Pereira dos Santos. II. Universidade do Estado do Rio Grande do Norte. III. Título.

UERN/BC

CDD 532

Bibliotecária: Elaine Paiva de Assunção CRB 15 / 492

CLÁUBIO LANDNEY LIMA BANDEIRA

**MODELAGEM E EXTRAPOLAÇÃO DO MOVIMENTO DE FLUIDOS
EM UM RESERVATÓRIO PETROLÍFERO USANDO MÉTODOS DE
PREVISÃO DE SÉRIES TEMPORAIS BASEADOS EM
CO-REGISTRO DE IMAGENS**

Dissertação apresentada ao Mestrado em
Ciência da Computação para a obtenção do
título de Mestre em Ciência da Computação.

Aprovada em: ___/___/_____

BANCA EXAMINADORA

Prof. D. Sc. Marcelino Pereira dos Santos Silva (UERN)
Orientador

Prof. D. Sc. Harold Ivan Ângulo Bustos (UERN)
Co-orientador

Prof. D. Sc. Fátima Nelsizeuma Sombra de Medeiros (UFC)
Examinadora Externa

Prof. D. Sc. João de Deus Lima (UERN)
Examinador Interno

Agradecimentos

Este trabalho não poderia ser terminado sem a ajuda de diversas pessoas às quais expresso meus agradecimentos:

Deus, em primeiro lugar, por me abençoar e me conceder a sabedoria, paciência e força necessárias para superar os desafios da vida.

Aos meus pais, Cenóbio e Liduína, pelo incentivo, orientação e carinho em todos os momentos de minha vida.

Às minhas amigas Aparecida Queiroz e Maria José, bem como à minha madrinha Maria Cecília pelo trabalho que eu sei que lhes dei.

Aos meus orientadores, Marcelino Pereira e Harold Bustos, que me mostraram os caminhos a serem seguidos de forma simples e objetiva, e também por me acompanhar durante todo esse percurso sempre com grande paciência.

À CAPES e ao CNPq pela concessão das bolsas de pesquisa e pelo apoio financeiro para realização da mesma.

Ao meu colega de apartamento Ewerton Ribeiro pelo companheirismo e que foi presença diária durante esse período.

A todos os novos amigos conquistados durante esses dois anos pelos momentos de descontração, que eles fiquem sempre na nossa memória.

Por fim, a minha amada e linda noiva, Priscila, pelo apoio e paciência que me foi dado durante todo esse tempo. Um dia você será recompensada...

Muito obrigado a vocês todos!

Resumo

Em um reservatório petrolífero em regime de produção, os fluidos estão em constante movimento devido a diversos fatores. Um destes principais fatores é a utilização de métodos de recuperação que visa manter e até mesmo aumentar a recuperação de petróleo e gás do reservatório. Para isso, esta técnica busca preservar a pressão do reservatório e deslocar o óleo para o poço produtor. No entanto, a maior parte do óleo tende a deslocar-se para outras camadas na subsuperfície, e não apenas aos poços. Assim, uma solução capaz de prever posições futuras desses fluidos é um processo muito importante para a indústria petrolífera. Apresentamos algoritmos que executam essa tarefa usando métodos de previsão de séries temporais. Para tanto, foram desenvolvidos algoritmos baseados em regressão linear e polinomial e também foram utilizados algoritmos baseados em aprendizagem de máquina no Waikato Environment for Knowledge Analysis (WEKA). Mapas de diferenças em lapso de tempo de imagem de dados reais do sinal sísmico variação da amplitude com o afastamento (AVO – do inglês *amplitude variation with offset*), antes e depois do co-registro de imagens, foram utilizados para validá-los.

Palavras-chave: Monitoramento de Fluidos, Predição de Fluidos, Lapso de Tempo, Co-registro, Previsão de Séries Temporais.

Abstract

In an oil reservoir in production, the fluids are in constant movement due to several factors. Among these factors, it can be highlighted the use of recovery methods aim of maintaining and even increasing the recovery of oil and gas from the reservoirs. For that, this technique seeks to preserve reservoir pressure and move the oil to the producing well. However, most of the oil tends to move to other subsurface layers and not just to the wells. Thus, a solution capable of predicting future positions of these fluids is a very important process for petroleum industry. We present algorithms that execute this task using time series forecasting methods. For this, algorithms based on linear and polynomial regression were developed and furthermore algorithms based on the machine learning from WEKA were used. Additionally, maps of time-lapse image differences from real AVO signal seismic data, before and after image co-registration, were used to validate them.

Keywords: Fluid Monitoring, Prediction of Fluid, Time-Lapse, Co-registration, Time Series Forecasting.

Lista de Tabelas

1	Média.	41
2	Variância.	41
3	Covariância.	41
4	MAE (%) calculado a partir do registro de 13 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.	46
5	RMSE (%) calculado a partir do registro de 13 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.	47
6	MAE (%) calculado a partir do registro de 19 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.	47
7	RMSE (%) calculado a partir do registro de 19 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.	47
8	MAE (%) calculado a partir do registro de 13 meses depois do co-registro de imagens, e suas extrapoladas.	50
9	RMSE (%) calculado a partir do registro de 13 meses depois do co-registro de imagens, e suas extrapoladas.	50

Lista de Figuras

1	Esquema representando o processo de reflexão sísmica.	10
2	Geração de ondas P e S refletidas e refratadas a partir de uma onda P incidente numa interface plana.	10
3	Modelo representando as aquisições sísmicas.	12
4	(a) Sismograma contendo as amplitudes lidas nos receptores em cada instante de tempo. (b) Representação gráfica do sismograma da esquerda composto por quatro traços sísmicos.	13
5	(a) representação gráfica de um sismograma e (b) representação gráfica do sismograma da esquerda em tons de cinza.	14
6	Sinal AVO.	16
7	Anomalia AVO.	17
8	Injeção de gás em um poço.	18
9	Modelo de regressão linear.	20
10	Modelo de regressão quadrática.	22
11	Co-registro rígido.	25
12	Visão geral do Madagascar.	29
13	Campo de Duri, Indonésia.	31
14	Modelo real do campo de Duri, Indonésia, colorido.	32
15	Corte após o co-registro de imagens.	33
16	Representação colorida do corte após o co-registro de imagens.	33
17	Modelo sintético em (a) um instante de tempo e (b) em um instante posterior.	34
18	Mapa de diferença para os dados sintéticos.	34
19	Diferença antes do co-registro de imagens.	35
20	Mapa de diferença para os dados sintéticos após o co-registro de imagens.	36
21	Resultado após o co-registro de imagens.	37
22	Extrapolador Linear - Parte 1.	38
23	Visualizador.	38
24	Sequências sísmicas de 13 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.	42
25	Sequências sísmicas coloridas correspondendo a 13 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.	43

26	Sequências sísmicas de 13 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.	43
27	Sequências sísmicas coloridas correspondendo a 13 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.	44
28	Sequências sísmicas de 19 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.	44
29	Sequências sísmicas coloridas correspondendo a 19 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.	45
30	Sequências sísmicas de 19 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.	45
31	Sequências sísmicas coloridas correspondendo a 19 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.	46
32	Sequências sísmicas correspondendo a 23 meses após a injeção do vapor extrapoladas por (a) regressão linear e (b) quadrática.	48
33	Sequências sísmicas coloridas correspondendo a 23 meses após a injeção do vapor extrapoladas por (a) regressão linear e (b) quadrática.	49
34	Sequências sísmicas de 13 meses coloridas (a) Original, extrapolada por (b) regressão linear (c) processo linear do WEKA, (d) processo gaussiano e (e) regressão quadrática.	49

Lista de Siglas e Abreviaturas

WEKA - Waikato Environment for Knowledge Analysis

AVO – do inglês *amplitude variation with offset* - variação da amplitude com o afastamento

CO₂ - dióxido de carbono

RNA - Redes Neurais Artificiais

AR - Auto-Regressivo

GPL, do inglês *General Public License* - Licença Pública Geral

BEG - *Bureau of Economic Geology*

SLIM - *Seismic Laboratory for Imaging and Modeling*

UBC - *Uniform Building Code*

SCons - *Software Construction*

PDF – do inglês *Portable Document Format* - Formato de Documento Portátil

HTML – do inglês *HyperText Markup Language* - Linguagem de Marcação de Hipertexto

RSF - *Regularly Sampled Format*

SEPLib - *Stanford Exploration Project Library*

PES - Projeto de Exploração Stanford

SU - *Seismic Unix*

PGM - Portable Gray Map

ASCII, do inglês American Standard Code for Information Interchange - Código Padrão Americano para o Intercâmbio de Informação

CSV - Comma-Separated Values

PNG - Portable Network Graphics

PAIV - Processamento e Análise de Imagens e Vídeos

MAE – do inglês *Mean Absolute Error* - Erro Médio Absoluto

RMSE – do inglês *Root Mean Squared Error* - Erro Médio Quadrático

Sumário

1	Introdução	5
1.1	Contextualização	5
1.2	Objetivos	6
1.3	Conteúdo	7
2	Princípios Básicos Associados à Sísmica	9
2.1	Método Sísmico de Reflexão	9
2.1.1	Aquisição	11
2.1.2	Processamento	12
2.1.3	Interpretação	12
2.2	Sismograma	12
2.3	Ponto Médio Comum	13
2.4	Velocidades Sísmicas	15
2.5	Inversão Sísmica	15
2.6	Sinal AVO	16
3	Predição de Fluidos	18
3.1	Métodos de Previsão de Séries Temporais	19
3.1.1	Regressão Linear	19
3.1.2	Processo Gaussiano	21
3.1.3	Regressão Quadrática	21
4	Co-Registro de Imagens	24
4.1	Atributo de Similaridade Local	25
5	Metodologia	28
5.1	Tecnologia de Software	28
5.1.1	Madagascar	28
5.1.2	WEKA	29
5.2	Dados Utilizados	30
5.2.1	Dados Utilizados no WEKA	31
5.3	Metodologia Proposta	33

5.3.1	Procedimento de Mapas de Diferenças	33
5.3.2	Procedimento de Co-registro	35
5.3.3	Procedimento para Extrapolação	36
5.3.4	Determinação do erro	39
5.3.5	Verificando a Estacionariedade	40
6	Resultados	42
6.1	Resultados Com o WEKA	48
7	Conclusões	51
	Referências	53

1 Introdução

1.1 Contextualização

Em um reservatório em regime de produção, é comum a migração dos fluidos presentes nele para outras camadas ao seu redor. A quantidade e a direção tomada pelo fluxo destes fluidos vão depender das propriedades destas ditas camadas. Além disso, outros fatores contribuem substancialmente para este deslocamento como, por exemplo, a utilização de métodos de recuperação.

Estes métodos são empregados quando há declínio na produção desses reservatórios, algo considerado comum e natural. Isso ocorre, principalmente, devido à redução da pressão interna e a permeabilidade do reservatório. Este efeito, por sua vez, afeta significativamente o fluxo dos fluidos presentes nele e, conseqüentemente, a produtividade do poço (GIRALDO *et al.*, 2000).

Com esse declínio, são empregadas técnicas que objetivam otimizar a recuperação dos hidrocarbonetos (petróleo e gás) contidos no interior do reservatório, visando manter e até mesmo aumentar sua produção. Dentre elas, podemos ressaltar o uso dos já citados métodos de recuperação, que por sua vez, têm como finalidade tentar manter a pressão do reservatório e deslocar o óleo para o poço produtor. Isso pode ser feito, por exemplo, injetando-se água ou gás dentro do reservatório. Todavia, grande parte do óleo tende a deslocar-se para outras camadas e não só para os poços.

Dessa forma, com este constante movimento, tarefas tais como monitorar esse deslocamento e mesmo prever posições futuras desses fluidos em rochas reservatórios é um processo bastante relevante. Para tanto, estudos de previsão poderão ser feitos permitindo assim, estimar possíveis posições futuras dos fluidos em reservatórios petrolíferos. A partir deles, é possível, por exemplo, identificar os melhores locais para a perfuração de novos poços (OSTRANDER, 1984) e até mesmo, guiar inteligentemente o uso dos dispendiosos processos de prospecção.

Dentre as técnicas utilizadas para este fim, está a sísmica 4D, que se baseia na aquisição de dados em lapso de tempo (do inglês, *time-lapse*) de sucessivas imagens 3D sobre uma mesma área (NUR, 1989). Por se tratar de predição no tempo, esta torna-se uma importante ferramenta, uma vez que ela permite fornecer os dados para realizar estes objetivos.

Segundo Lumley (2010), a sísmica 4D pode ser extremamente útil para monitorar o dióxido de carbono (CO₂) injetado e armazenado em reservatórios geológicos de subsuperfície. A sísmica 4D e mapas de diferenças podem identificar o fluxo dos fluidos no reservatório. Neste ponto, o perfeito alinhamento destas sequências é extremamente necessário.

Para tanto, existem metodologias que permitem realizar esta tarefa de forma mais precisa como, por exemplo, aplicando o co-registro de imagens, também conhecido como registro de imagens ou alinhamento geométrico. Esta técnica é amplamente estudada em disciplinas como Visão Computacional e bastante aplicada no imageamento médico. Em aplicações sísmicas, o co-registro de imagens pode ser utilizado para rastrear o vapor d'água injetado nos poços entre as sequências sísmicas, possibilitando o mapeamento do deslocamento de fluidos no reservatório. A partir deste ponto, é possível criar um modelo numérico de previsão mais confiável.

Mediante o monitoramento e baseado nas medidas de localizações passadas dos fluidos, estes dados podem ser processados por algoritmos específicos de forma a poder realizar sua predição futura no tempo. Os resultados obtidos por esses algoritmos têm um grande impacto

na indústria petrolífera. Para tanto, isso pode ser realizado utilizando diversas abordagens, tais como redes neurais, meta-heurística, dentre outros. Porém, outros tipos de metodologias permitem realizar tal tarefa, como as utilizadas neste trabalho, onde foram gerados modelos numéricos de reservatórios utilizando métodos de previsão de séries temporais.

Em (RISSO; SCHIOZER, 2008) os autores propuseram um método integrando o histórico de produção e de pressão com os mapas das frentes de saturação obtidos através da sísmica 4D, permitindo ajustá-los de forma que estes modelos possam ser usados na previsão de produção com maior confiabilidade.

Um experimento similar ao apresentado neste trabalho, em que apresenta predição temporal sobre imagens, pode ser visto em (ANDRADE *et al.*, 2004), e usa Redes Neurais Artificiais (RNA) na implementação e análise de desempenho na estimação de padrões espaço-temporais da coroa solar. Para isso, foram utilizados os algoritmos de treinamento Backpropagation, Levenberg-Marquardt (HAGAN; MENHAJ, 1994) e Filtragem de Kalman (NETO, 1995).

Uma série temporal pode ser definida como um conjunto de observações geradas sequencialmente e obtidas ao longo do tempo e que, além disso, apresentam uma dependência entre elas. Dentre seus objetivos, está o de realizar previsões de valores futuros da série e que podem ser obtidos mediante alguns métodos como, por exemplo, os baseados em regressão. Estes, por sua vez, necessitam de valores passados, isto é, um histórico, e são conhecidos por métodos de regressão.

Dentre os métodos simples de regressão podemos destacar o linear, o polinomial, o exponencial, o potencial e o hiperbólico. Todos eles constituem métodos simples de previsão de séries temporais. Neste trabalho foram desenvolvidos os algoritmos de regressão linear e polinomial de ordem 2, ou seja, quadrático.

Além disso, outros métodos de previsão de séries temporais serão utilizados. Estes, por sua vez, baseados em algoritmos de aprendizagem de máquina e correspondem ao processo linear – onde pretende-se compará-lo com o implementado neste trabalho – e ao processo gaussiano. Para tanto, será adotado o programa WEKA (HALL *et al.*, 2009).

1.2 **Objetivos**

Diante deste cenário, esse trabalho possui três objetivos fundamentais. O primeiro, é o desenvolvimento de algoritmos para realizar a predição de fluidos em reservatórios em lapso de tempo (4D). Para isto, modelos numéricos foram gerados usando métodos simples de previsão de séries temporais baseados em modelos de regressão linear e polinomial do sinal de interesse a ser monitorado.

O segundo objetivo é identificar se o co-registro de imagens permite obter resultados mais exatos no processo de predição. Assim, para os trabalhos desenvolvidos foram utilizados dados reais de reservatórios obtidos da base de dados do pacote Madagascar (MADAGASCAR, 2011), antes e depois do co-registro de imagens, que por sua vez, utiliza a métrica de similaridade baseada no atributo de similaridade local (FOMEL, 2007a). Com estes dados, busca-se validar os algoritmos desenvolvidos.

Por fim, o terceiro objetivo é avaliar se o método de regressão linear desenvolvido é melhor ou tão bom quanto o implementado no WEKA. A partir dos resultados, pode-se identificar qual dos métodos de regressão mostrou-se superior.

O processo de co-registro de imagens foi desenvolvido dentro do pacote Madagascar. Já a implementação dos algoritmos para a predição foi realizada na linguagem de programação JAVA. Também, modelos numéricos do reservatório foram gerados baseados nos processos linear e gaussiano utilizando algoritmos de aprendizado de máquina do WEKA.

1.3 Conteúdo

Os capítulos desse trabalho estão organizadas da seguinte forma: o Capítulo 2 discorre sobre alguns princípios básicos que estão associados à sísmica; O Capítulo 3 aborda a predição de fluidos e os métodos de previsão de séries temporais utilizados para este fim; O Capítulo 4 introduz uma técnica importante que é utilizada nesta pesquisa: co-registro de imagens; o Capítulo 5 aborda os materiais e a metodologia empregada neste trabalho; O Capítulo 6 apresenta os resultados obtidos; E o Capítulo 7 conclui esta dissertação.

No intuito de esclarecer quais são os princípios básicos associados à sísmica, realizou-se no Capítulo 2, uma revisão de conceitos. Geralmente, o subsolo é constituído por camadas de sedimentos que por sua vez, possuem propriedades físicas distintas, tais como a elasticidade, o que permite a propagação de ondas com velocidade também diferente. Essas ondas, ao encontrarem uma interface entre os meios com essas propriedades, têm parte de sua energia refletida e outra parte refratada. Este princípio, é fundamental para o processo de aquisição dos dados.

Ainda no Capítulo 2, são esclarecidos detalhes sobre a técnica de prospecção sísmica que se baseia na medição do tempo de percurso, denominado tempo de trânsito, de ondas elásticas introduzidas artificialmente em pontos próximo à superfície do terreno ou do mar. Também, pode ser visto neste capítulo, o principal método usado na prospecção de hidrocarbonetos: o método sísmico de reflexão. Além disso, será abordado desde o processo de aquisição de dados até a sua interpretação, bem como conceitos básicos sobre sismograma que pode ser convertido para imagens, sendo o principal objeto de estudo deste trabalho.

No Capítulo 3 é abordada a predição de fluidos e sua importância para o setor petrolífero. É discutido o conceito de séries temporais e de métodos de previsão de séries temporais. Também, são descritos os métodos que foram utilizados neste trabalho, regressão linear e quadrática, para gerar as estimativas do comportamento futuro do reservatório.

O Capítulo 4 define uma das principais técnicas que utilizadas nesta pesquisa: o co-registro de imagens. Neste capítulo pode-se ver a divisão desta técnica, em rígido e não-rígido, e o conceito da métrica de similaridade denominada atributo de similaridade local, que foi aplicada nos dados utilizados neste trabalho.

O Capítulo 5 exhibe os materiais, ferramentas e dados empregados. Nele, é introduzido o software de processamento sísmico e realizada uma breve discussão sobre sua utilização. Veremos que estes softwares constituem uma importante ferramenta na exploração de petróleo e gás, mas que seus custos são elevados. Assim, surgiram softwares poderosos distribuídos livremente, como o Madagascar. Ele possui um formato de dados simples e flexível que podem facilmente serem lidos, pois apresentam-se em formatos de arquivos de textos. Também, será apresentado o programa WEKA, que é implementado em Java e de distribuição livre e código aberto. Ele possui várias técnicas de aprendizagem de máquina.

Ainda no Capítulo 5 são abordadas: a metodologia proposta para a criação dos mapas de diferenças, a partir de sequências sísmicas em lapso de tempo; o procedimento para a realização

do co-registro de imagens através do software Madagascar; e a criação de um programa com interface gráfica na linguagem de programação Java para a realização da extrapolação linear; o programa para determinar o erro entre as sequências sísmicas reais e os modelos numéricos de previsão obtidos pelas técnicas de regressão; e por fim, é verificada a estacionariedade da série.

O Capítulo 6 apresenta os resultados realizados sobre dados de campo reais. Visando validar os algoritmos propostos e os utilizados no Weka, veremos neste capítulo, que foram geradas previsões para sequências sísmicas conhecidas. Para os algoritmos desenvolvidos, estas sequências correspondem a 13 e 19 meses após a injeção do vapor no reservatório, antes e depois da aplicação do co-registro de imagens. Estes resultados foram objetos de publicações no V e-Science Workshop 2011 realizado no XXXI Congresso da Sociedade Brasileira de Computação, podendo ser visto em (BUSTOS *et al.*, 2011b) e no IADIS Applied Computing 2011, disponível para consulta em (BUSTOS *et al.*, 2011a). Já para os algoritmos do Weka, foram gerados previsões para a sequência de 13 meses após a aplicação do co-registro de imagens.

O Capítulo 7 exhibe as conclusões dos experimentos realizados, bem como sugestões para trabalhos futuros.

2 Princípios Básicos Associados à Sísmica

Neste capítulo, são apresentadas definições pertencentes às áreas distintas de conhecimento, especialmente conceitos básicos relacionados à área de exploração sísmica.

2.1 Método Sísmico de Reflexão

Neste texto pretende-se introduzir o método sísmico de reflexão visando a compreensão do processo para a geração dos dados, uma vez que estes constituem uma das partes fundamentais deste trabalho.

O subsolo é geralmente constituído por diferentes camadas de sedimentos. Essas camadas possuem algumas propriedades físicas distintas, dentre elas a elasticidade, o que faz com que a propagação das ondas também apresente velocidades diferentes. Ao encontrarem meios com essas propriedades, estas ondas têm parte de sua energia refletida e outra parte refratada, similarmente como o que ocorre em ótica, na Física. Com isso, é possível determinar as velocidades de propagação das ondas e a posição das interfaces que separam os meios com valores de elasticidades diferentes. Para tanto, isto pode ser realizado medindo o tempo de percurso das ondas em diferentes pontos, bem como a distância entre esses pontos.

Desta forma, tem-se a prospecção sísmica que é especialmente útil para determinar a espessura de pacotes sedimentares, pois a velocidade das ondas elásticas é maior no embasamento cristalino e difere consideravelmente entre alguns tipos de sedimentos (LUIZ; SILVA, 1995).

A prospecção sísmica baseia-se em medir, em diferentes pontos, o tempo de percurso de ondas elásticas introduzidas artificialmente geradas num ponto próximo à superfície do terreno ou no mar. Essas ondas podem ser geradas de diversas formas como, por exemplo, mediante o uso de explosivos e por motivos ecológicos, pode-se utilizar outros métodos, tais como ar comprimido, queda de pesos ou vibradores. Devido às descontinuidades elásticas das rochas, parte dessas ondas retornarão a outros pontos da superfície e outra parte continuará a penetrar no interior do subsolo, dividindo, assim, a sísmica em dois: sísmica de reflexão e sísmica de refração.

A sísmica da reflexão é uma técnica que faz uso das ondas refletidas e é aplicada na determinação das profundidades onde se encontram as superfícies refletoras. Além disso, ela é útil para determinar as velocidades sísmicas das rochas que compõem as várias camadas do subsolo. Este modelo possui inúmeras vantagens em relação a outros modelos pelo fato de utilizar receptores que se estendem por vários metros e até mesmo, quilômetros. Dentre essas vantagens, pode-se citar a possibilidade de produzir imagens sísmicas de alta definição das feições geológicas da área explorada, determinadas pela distância entre esses sensores: quanto mais próximo um sensor do outro, maior a resolução das imagens geradas. Uma outra vantagem deste método é a cobertura de uma grande área de aquisição, definida pelo número de sensores posicionados sobre o terreno.

Devido à isso, a sísmica da reflexão é a mais utilizada na prospecção de hidrocarbonetos e visa gerar artificialmente com uma fonte, por exemplo, um explosivo, ondas elásticas em pontos próximos à superfície. O tempo percorrido pela onda dentro da camada desde a fonte até um conjunto de receptores, bem como suas amplitudes na chegada são armazenados e formam o principal ponto de informação para estudos posteriores. Esse tempo de viagem que a onda gasta

é conhecido como tempo de trânsito. A Figura 1 ilustra o funcionamento desse método.

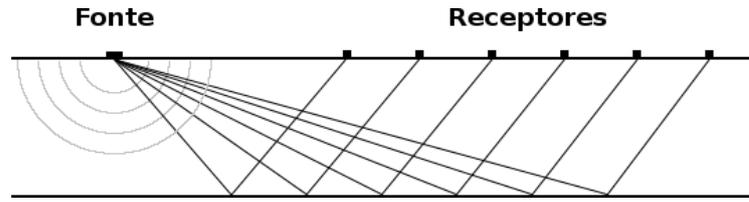


Figura 1: Esquema representando o processo de reflexão sísmica.

A razão entre a energia refletida e a energia incidente é chamada de coeficiente de reflexão. Seja ρ_1 e ρ_2 as densidades da primeira e da segunda camada, respectivamente, e V_{ρ_1} e V_{ρ_2} as velocidades da primeira e da segunda camada (FONSECA, 2002), o coeficiente de reflexão R é definido pela razão entre as densidades e as velocidades sísmicas entre dois materiais, relação dada por

$$R = \frac{\rho_2 V_{\rho_2} - \rho_1 V_{\rho_1}}{\rho_2 V_{\rho_2} + \rho_1 V_{\rho_1}}.$$

A Figura 2 ilustra a propagação de uma onda. Ao mudar de uma camada de rochas para outra, esta com propriedades físicas diferentes, uma onda tem parte de sua energia refletida, enquanto que uma outra parte prossegue e se propaga ainda mais para as camadas inferiores da terra. Nesta transição, a energia da onda P (ondas compressoriais, P vem do fato de que são as primeiras a chegar) refletida e da onda P' refratada (ou transmitida) pode gerar uma onda S (ondas cisalhantes, S vem do fato de que são as segundas a chegar) - também conhecida como onda convertida P-S. O ângulo de incidência θ_1 é igual ao ângulo de reflexão e o ângulo θ_2 é o ângulo de transmissão da onda no meio 2. S' corresponde à onda S refratada.

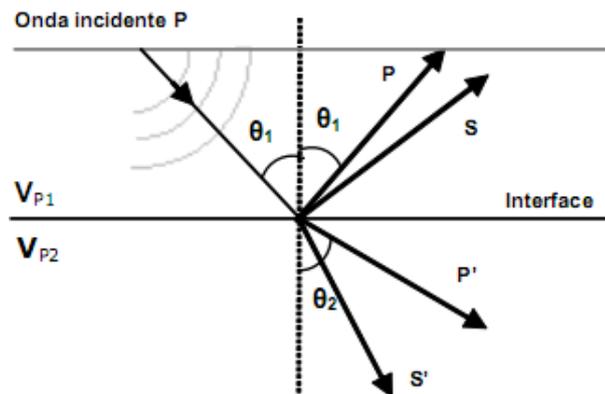


Figura 2: Geração de ondas P e S refletidas e refratadas a partir de uma onda P incidente numa interface plana (MEDEIROS, 2005).

Sendo ρ uma constante conhecida por parâmetro do raio (SHERIFF; GELDART, 1995), e V_1 e V_2 as velocidades nos meios 1 e 2, respectivamente, a lei de Snell mostra que a relação entre as velocidades das camadas e os ângulos de refração e reflexão é dado por:

$$\frac{\text{sen } \theta_1}{V_{\rho_1}} = \frac{\text{sen } \theta_2}{V_{\rho_2}} = \rho.$$

Os meios em que a onda percorre podem ser classificados em homogêneos ou heterogêneos e isotrópicos ou anisotrópicos. No meio homogêneo, a propagação da onda possui a mesma velocidade na mesma direção, enquanto que em um meio heterogêneo, esta velocidade varia. Já um meio dito isotrópico, considera-se que as propriedades físicas do meio são iguais em todas as direções. Entretanto, em um meio anisotrópico, as características variam conforme a direção observada.

De acordo com Robinson e Treitel (1980), a exploração de hidrocarbonetos (óleo e gás) baseada em sísmica pode ser dividida em três etapas principais: aquisição, processamento e interpretação, as quais serão apresentadas a seguir.

2.1.1 Aquisição

A fase de aquisição compreende o levantamento (registro das medidas) dos dados sísmicos. Na aquisição dos dados, assume-se que a propagação de ondas sísmicas (vibrações) através das rochas pode ser compreendida utilizando-se as mesmas leis físicas que controlam a óptica geométrica (THOMAS, 2001).

Para adquirir os dados, ondas elásticas artificiais (impulsos) são geradas em pontos específicos da área na qual se pretende mapear. Essas ondas possuem tempo de duração muito pequeno (aproximadamente 200 milissegundos) e são geradas, principalmente, utilizando-se dinamite (quando em terra) ou canhões de ar comprimido (no caso de regiões marítimas).

Uma vez gerada, a onda sísmica se propaga estendendo-se através da terra e quando alcança uma interface entre duas rochas de características físicas diferentes, parte da energia incidente da onda é refratada para o meio inferior e uma outra parte é refletida e retorna à superfície - onde poderá ser captada por sensores.

Esses sensores são colocados alinhados em pontos pré-determinados da superfície e captam as ondas que foram refletidas pelas interfaces entre as rochas. De acordo com o meio em que serão empregados, tais receptores podem ser de dois tipos: eletromagnéticos, para captação em terra, denominados geofones; ou de pressão, chamados de hidrofones, para aquisição de dados em regiões oceânicas (FIGUEIREDO, 2007).

Após serem captadas pelos sensores, as informações são gravadas em um dispositivo chamado sismógrafo - que armazena as amplitudes das ondas em intervalos de tempo regular. Tanto o instante de tempo em que a informação chegou, quanto sua intensidade da onda medida nesse momento também são armazenados.

O número de geofones, ou de hidrofones, utilizados para recepção depende do número de canais disponíveis no sismógrafo. A Figura 3 mostra os processos de aquisição em ambos os meios: terrestre e marinho, representados pela Figura 3(a) e pela Figura 3(b), respectivamente.

As rochas possuem características particulares e que mudam dependendo do seu tipo. Dentre essas características, encontram-se a porosidade, a temperatura, a pressão, entre outros. Estes atributos são responsáveis por uma série de fatores diferentes, tais como densidade e constantes elásticas, que fazem com que a velocidade de propagação mude dependendo do meio.

Neste processo de aquisição, o intervalo entre os sensores na superfície assume um fator importante na aquisição dos dados, pois quanto menor o espaçamento lateral entre os sensores na superfície, melhor será a resolução da imagem sísmica captada.

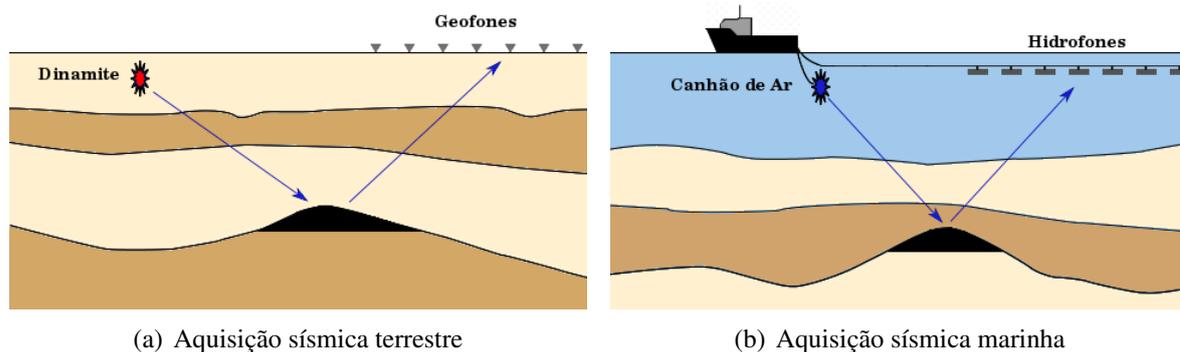


Figura 3: Modelo representando as aquisições sísmicas.

2.1.2 Processamento

A fase de processamento visa a produção de imagens da parte interna da terra que sejam mais exatas e aproximadas das camadas geológicas reais. No entanto, durante a aquisição dos dados, distorções são geradas e é nesta fase que se procura atenuá-las e corrigi-las.

Além disso, as reflexões múltiplas causam um efeito físico indesejado gerado pelas ondas sísmicas, conhecido como reverberações. Contudo, é necessário que seus efeitos sejam atenuados ou mesmo eliminados e, assim, os dados são submetidos a um processo denominado deconvolução, onde se procura eliminar esse efeito. A deconvolução pode também ser usada para aumentar a resolução temporal dos dados sísmicos, comprimindo o traço sísmico de forma a restar somente a função reflexividade (resposta impulsiva da terra) (FIGUEIREDO, 2007).

2.1.3 Interpretação

Na etapa de interpretação, faz-se a interpretação das imagens obtidas para efetivamente localizar feições de interesse presentes nelas. Neste ponto, é altamente recomendado o acompanhamento de um profissional, geralmente um geólogo ou um geofísico para que ele possa discutir os resultados obtidos.

2.2 Sismograma

No processo de aquisição sísmica, as ondas, ao retornarem da subsuperfície, são capturadas pelos receptores que fazem uma leitura simultânea transformando as vibrações do solo (sinal sísmico) em sinais elétricos. Em intervalos de tempo dt , os receptores verificam se existem ou não sinais retornando à superfície. Cada receptor registra um traço sísmico, que é uma função do tempo de ida e volta que a onda percorreu desde a fonte até o receptor. Ao conjunto de traços registrados (pelo arranjo de geofones) dá-se o nome de sismograma.

Os dados de um sismograma podem ser representados por uma matriz em que cada coluna e cada linha correspondem, respectivamente, aos índices dos receptores e aos índices dos instantes de tempo. Em seu conteúdo ficam armazenadas as amplitudes das ondas captadas.

A partir de um sismograma, é possível gerar uma representação gráfica dele. A Figura

4(a) apresenta um exemplo de sismograma com quatro receptores (rec_1 , rec_2 , rec_3 e rec_4) e seis instantes de tempo (tt_1 , tt_2 , tt_3 , tt_4 , tt_5 e tt_6). O conteúdo do sismograma é dado pelas amplitudes de ondas capturadas por cada receptor rec_i no instante de tempo tt_i . A Figura 4(b) contém uma representação gráfica do sismograma da Figura 4(a) onde o eixo das abscissas corresponde ao afastamento (distância entre a fonte e cada receptor) e o eixo das ordenadas, ao tempo das leituras dos receptores. Cada coluna da matriz é agora representada por uma linha vertical igual a coluna da matriz, a qual denominamos traço sísmico. A curva presente nessa imagem, que deve ser aproximada a uma semi-hipérbole, é definida pelas amplitudes máximas existentes na imagem e representa um evento sísmico.

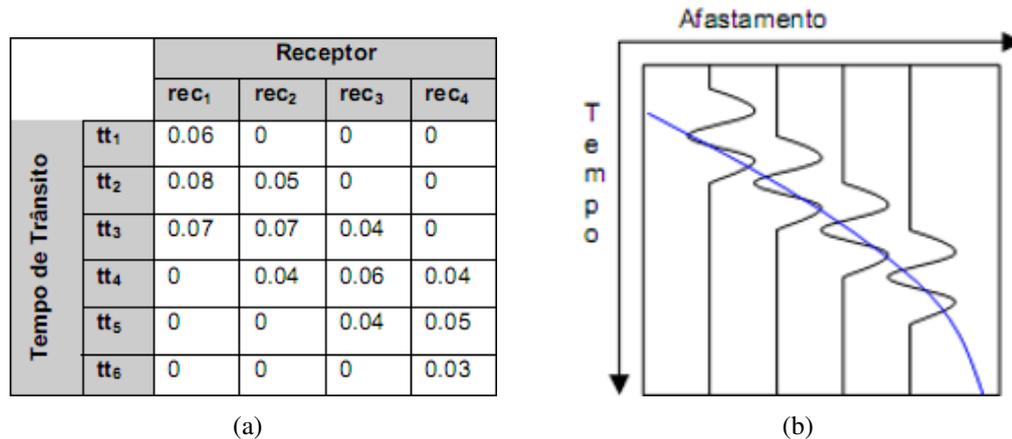


Figura 4: (a) Sismograma contendo as amplitudes lidas nos receptores em cada instante de tempo. (b) Representação gráfica do sismograma da esquerda composto por quatro traços sísmicos (MEDEIROS, 2005).

Depois de gerada uma imagem gráfica do sismograma, podemos convertê-la para uma imagem em nível de cinza, onde o valor de cada pixel é definido pela correspondente amplitude. Essa nova imagem é o principal objeto deste trabalho, onde as análises serão feitas, permitindo que se tenha uma visualização mais rápida dos dados e, assim, possibilitando uma interpretação dos mesmos com mais rapidez e eficiência.

A Figura 5 exemplifica esse processo de conversão. Na Figura 5(a) tem-se um exemplo de um sismograma em sua representação gráfica, como mostrado na Figura 4(b), e na Figura 5(b) tem-se a representação em tons de cinza desta figura.

2.3 Ponto Médio Comum

O principal arranjo de campo empregado na sismica é o chamado ponto médio comum (CMP, do inglês *common-midpoint*) (THOMAS, 2001), também conhecido por ponto comum em profundidade (CDP, do inglês, *common depth point*). O CMP indica que um conjunto de dados sísmicos possui um ponto médio comum entre fonte e receptor. Porém, uma distinção deve ser feita: o termo original CDP é válido somente em casos de refletos planos que não apresentem nenhuma variação lateral de velocidade.

Nos primeiros levantamentos de dados sísmicos, era realizada apenas uma única leitura para cada ponto de reflexão da onda na interface (SHERIFF; GELDART, 1995). Um único disparo era realizado e os receptores armazenavam as amplitudes das ondas refletidas. Com esses

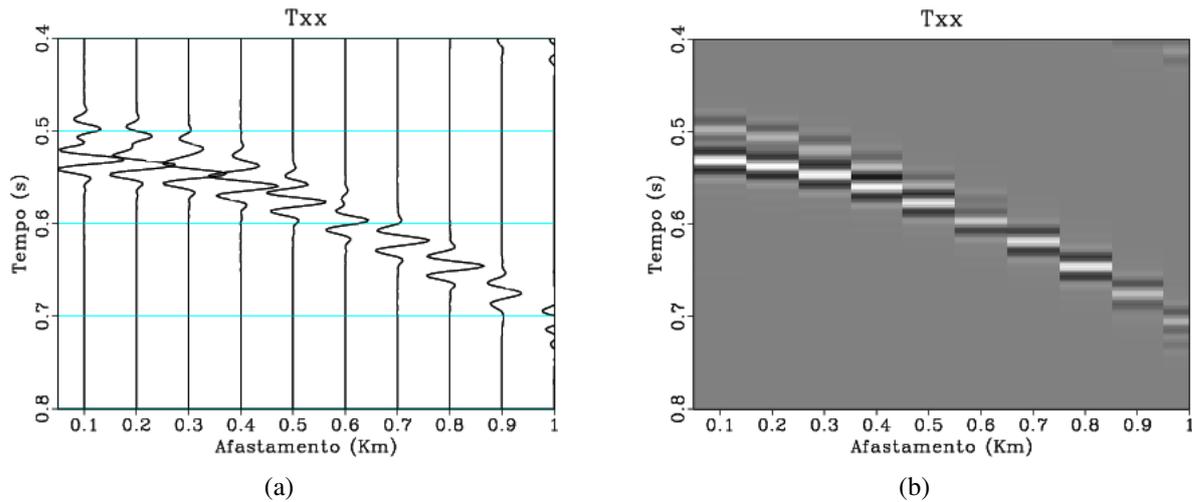


Figura 5: (a) representação gráfica de um sismograma e (b) representação gráfica do sismograma (a) em tons de cinza. Adaptado de (FOMEL, 2009).

dados, realizava-se a interpretação. Porém, por volta de 1960, esse quadro mudou passando-se a realizar várias leituras de um mesmo ponto na interface e utilizando-se vários disparos.

Para se formar um CMP, altera-se a posição da fonte a cada disparo, a fim de ler o mesmo ponto várias vezes e adquirir dados com redução do ruído e, assim, melhorar a relação sinal-ruído.

Com a utilização de vários receptores para cada disparo e de uma grande área a ser mapeada, gera-se uma enorme quantidade de dados. Utilizando-se a soma de amplitudes dos traços, também chamada de empilhamento, reduz-se bastante a redundância de informações presentes nos dados, permitindo a diminuição de uma série de ruídos e o enriquecimento do sinal sísmico, ou seja, o aumento dos sinais de interesse obtidos pelas ondas sísmicas que foram refletidas pelas camadas no subsolo e captadas pelo sensores na superfície.

A distância entre fonte e receptor é denominada *offset*, ou afastamento. Quanto maior o afastamento, maior será o tempo de chegada da onda na superfície, refletida por uma mesma interface, caracterizando um atraso na captação do sinal pelos receptores posicionados cada vez mais distantes da fonte. Para corrigir esse efeito, utilizamos a correção sobretempo normal (NMO, do inglês *normal moveout*).

Seja $t(x)$ o tempo estimado do evento sísmico, t_0 o tempo de trânsito (tempo que a onda leva para se propagar da fonte até o receptor) com deslocamento zero (*zero-offset*), ou seja, fonte e receptor estão no mesmo ponto, x a posição do receptor e V_1 é a velocidade do meio, a NMO é dado pela seguinte igualdade

$$t(x) = \sqrt{t_0^2 + \left(\frac{x}{V_1}\right)^2}.$$

2.4 Velocidades Sísmicas

Na aplicação dos métodos sísmicos, a velocidade de propagação das rochas constitui a característica mais importante delas, principalmente no que diz respeito às ondas longitudinais (ondas-P), que como foi dito anteriormente, são as mais rápidas e, portanto, as primeiras a chegarem e a serem registradas. A reflexão, bem como a refração, da onda sísmica está relacionada às diferenças de velocidades das ondas durante uma passagem de um meio para outro, sendo indispensável o conhecimento das velocidades das ondas nas rochas para o processo de interpretação sísmica.

As velocidades sísmicas mais conhecidas são (SHERIFF; GELDART, 1995):

1. Velocidade Intervalar - V_i : é a média de um intervalo de uma determinada profundidade para uma camada. É a velocidade que mais se aproxima do conceito de velocidade característica de um meio físico-geológico. Em uma camada isotrópica e homogênea, a velocidade intervalar corresponde à velocidade da camada.
2. Velocidade Média - V_{bar} : resulta do cálculo da média de todas as velocidades das camadas, desde a superfície até uma determinada profundidade.
3. Velocidade *Normal Moveout* - V_{nmo} : é o parâmetro que descreve a linha da reflexão sísmica armazenada para diferentes distâncias de *offset* (distância entre a fonte e o receptor). Determinar essas velocidades é o passo inicial para uma estimativa das velocidades das camadas. Para isto, é necessário o conceito da velocidade RMS (*root-mean-square*) e o algoritmo que se desdobra a partir dela, chamado algoritmo de DIX (DIX, 1955) que é usado para encontrar as velocidades intervalares.
4. Velocidade *Root-Mean-Square* - V_{rms} : essa velocidade diz respeito a uma trajetória que a onda percorreu e representa a velocidade unificada de todas as camadas por onde essa onda passou. Se V_{rms} é a média quadrática das velocidades intervalares das camadas por onde a onda passou, então podemos obter a velocidade intervalar de cada uma dessas camadas.

2.5 Inversão Sísmica

O problema inverso consiste em determinar estruturas dos dados de subsolo a partir da prospecção geológica, tendo como principal objetivo conseguir uma seção geológica ou um modelo 3D (LINDEN, 2006).

Pode-se descrever o problema inverso como a teoria matemática que mostra como se pode conseguir informações de parâmetros não observáveis do sistema a partir de observações feitas sobre esse próprio sistema, de relações teóricas entre os parâmetros não observáveis do sistema e das observações, e das informações que se tem *a priori* sobre os parâmetros não observáveis do sistema.

A partir de propriedades sísmicas, pode-se quantificar atributos físicos das rochas, tais como a velocidade, a densidade, dentre outros. A partir de um sismograma e utilizando-se um método de inversão, obtêm-se esses e outros atributos.

2.6 Sinal AVO

Os dados utilizados neste trabalho, correspondem ao sinal AVO. À medida que o afastamento aumenta, a amplitude registrada geralmente tende a decrescer (*falls*), conforme pode ser observado na Figura 6. Nessa figura, elevadas amplitudes são registradas nos pontos próximos à origem, isto é, mais próximos da fonte. Estas amplitudes são representadas pela cor branca no sismograma. À medida que o afastamento aumenta, são registradas amplitudes mais baixas, uma vez que as ondas acústicas têm que percorrer um percurso maior entre a fonte e os sensores mais distantes. Durante esse percurso, elas tendem a perder mais energia para as camadas da subsuperfície. Estas amplitudes são representadas pela cor preta. Assim, temos um gradiente sobre a traça sísmica variando entre o branco, nos pontos mais próximos à origem, e o preto, nos mais distantes.

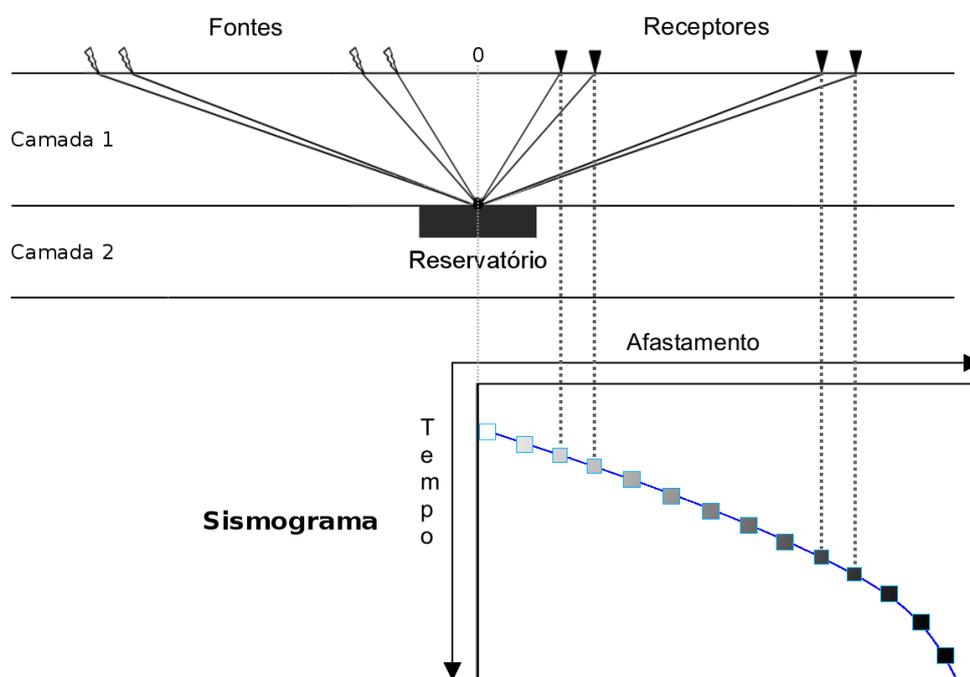


Figura 6: Sinal AVO.

Diferentemente dos dados invertidos, este tipo de sinal não mostra as reais posições dos objetos presentes no interior dos subsolo. Por exemplo, na Figura 7, os sensores mais distantes (os dois mais à direita) estão captando informações sobre o reservatório que está na origem e registrando os pixels em um ponto diferente dela no sismograma. Porém, o correto seria que estes pixels estivessem também próximos à origem, como no caso real.

Segundo Kisuka e Andrade (2006), na década de 60 descobriu-se que a presença de gás estava geralmente associada às regiões de elevadas amplitudes na seção sísmica, conhecidas como *bright spots* (aumento da amplitude), uma vez que normalmente a amplitude registrada decresce com o afastamento.

Todavia, na presença de gás, ou vapor, há um crescimento irregular (*rising*) conhecido como anomalia. A Figura 7 ilustra esse processo. Mediante este comportamento, dá-se a importância da análise de AVO como instrumento para detecção direta de hidrocarbonetos (OSTRANDER, 1984).

Uma vez que o vapor é injetado no reservatório, como nos dados reais utilizados neste tra-

balho, entra em cena uma outra propriedade: a impedância acústica. Impedância acústica é definida como o produto entre a velocidade compressional e a densidade da rocha (SANCEVERO *et al.*, 2006) e está relacionada com a resistência ou dificuldade do meio à passagem das ondas acústicas. Assim, nos dados adquiridos após a injeção de vapor, ele pode ser facilmente identificado nas sequências sísmicas, pois apresenta-se com elevados valores de amplitudes, estas por sua vez, relacionadas ao contraste de impedância acústica entre os meios. Estes valores podem ser observados na Figura 7 por pixels brancos sobre a traça sísmica, principalmente nos dois pontos mais distantes da origem. Além disso, por possuir uma densidade menor, ele tende a permanecer no topo do reservatório agregado ao óleo/gás presente nele. Assim, monitorando-se o vapor, também se está monitorando os fluidos.

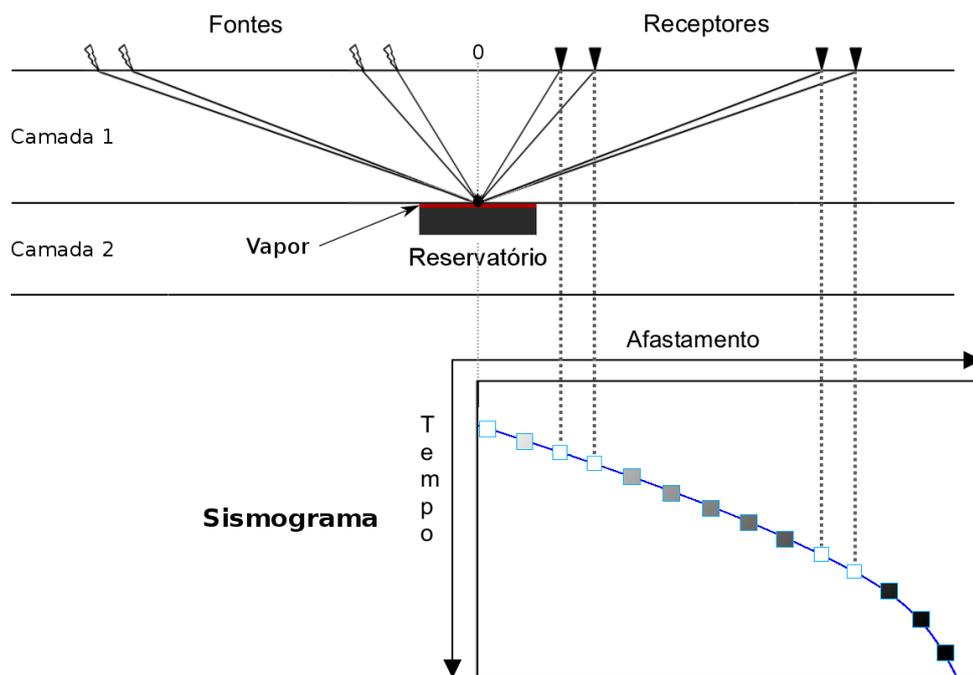


Figura 7: Anomalia AVO.

3 Predição de Fluidos

Os reservatórios em regime de produção sofrem constante mudanças, sejam elas no formato, no volume ou em seu fluxo no interior destes. Alguns fatores contribuem ainda mais para essa situação, tais como os descritos a seguir.

Da quantidade de petróleo existente nos reservatórios, na prática, apenas uma pequena fração consegue ser retirada (CURBELO, 2006). Isso ocorre principalmente devido à redução da pressão interna. Além disso, um outro fator contribui para isso: a permeabilidade do reservatório.

Com o objetivo de aumentar esta quantidade, nos poços em regime de produção, uma das técnicas mais utilizadas são os métodos de recuperação. Estes, por sua vez, têm a finalidade de tentar manter a pressão do reservatório e deslocar o óleo para o poço produtor. Para tanto, dentre as técnicas empregadas para este fim, estão a injeção de água, gás, vapor d'água, dentre outros. A Figura 8 a seguir, ilustra este processo.

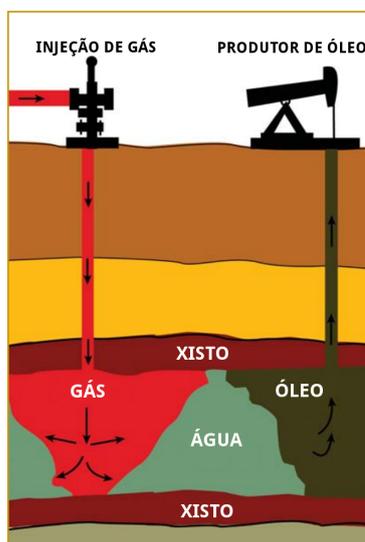


Figura 8: Injeção de gás em um poço. Adaptado de (LYONS, 2010).

Dependendo das características pertencentes às camadas vizinhas ao reservatório, como a porosidade, por exemplo, grande parte do óleo tende a deslocar-se para outras regiões e não só para os poços.

Também, durante a produção, algumas outras alterações nos reservatórios podem ser observadas, tais como a densidade e a compressibilidade do reservatório. Estas alterações são acarretadas por uma série de mudanças, principalmente a pressão de poro e a temperatura (COSTA, 2005), e podem ser detectadas mediante a análise de repetidos levantamentos de dados sísmicos.

Desta forma, é necessário o uso de técnicas que permitam monitorar o deslocamento desses fluidos no interior destas camadas. Para abordar este problema, tem-se a sísmica 4D que baseia-se na aquisição de dados em intervalos de tempo, necessitando de elevados investimentos que são essenciais para o processo de aquisição de sucessivas imagens 3D sobre uma mesma área – o que a torna bastante dispendiosa. Todavia, esse processo é fundamental para se realizar o monitoramento do movimento dos fluidos no reservatório e até mesmo possibilitar a identificação de áreas com óleo ainda não drenado.

Neste contexto, injeta-se vapor d'água nos poços e monitora-se seu deslocamento. Com isso, tem-se basicamente que, devido às suas propriedades químicas, ele tende a deslocar parte do fluido no reservatório. Além disso, ele tende a permanecer no topo do reservatório agregado ao óleo/gás presente nele, já que possui uma densidade menor. Desta forma, rastreando-se o movimento do vapor d'água, também se rastreia-se o dos fluidos.

Em seguida, são realizados levantamentos sísmicos em períodos de tempo diferentes. Neste ponto entra uma outra característica do vapor, a alta impedância acústica, que conseqüentemente está diretamente relacionada às elevadas amplitudes nas traças.

Uma vez coletados estes dados podem ser processados, por algoritmos específicos, de forma a possibilitar o mapeamento do deslocamento de fluidos no reservatório e a partir deste ponto, até mesmo predizer sua posição futura. Com os resultados deste processo, é possível guiar inteligentemente o uso dos dispendiosos processos de prospecção. A modelagem de fluxo de fluido e sua previsão no tempo pode ser aproximada através de técnicas de extrapolação, empregando métodos de previsão de séries temporais, descritos a seguir.

3.1 Métodos de Previsão de Séries Temporais

Uma série temporal é um conjunto de observações geradas sequencialmente ao longo do tempo e que apresentam uma dependência entre elas.

Alguns exemplos de séries temporais são:

1. Quantidade anual de chuva;
2. Valores diários da temperatura de uma determinada cidade;
3. Índices diários de bolsa de valores;

No entanto, todos estes exemplos, bem como a maioria das aplicações que envolvem o uso de métodos de previsão, trabalham no espaço 1D. Porém, nossas aplicações, por se tratarem de imagens, estão no espaço 2D. Todavia, elas podem ser convertidas para um vetor, convertendo-se assim, para o espaço unidimensional.

Dentre os objetivos da análise de séries temporais, está a realização de previsões dos valores futuros da série, que é o objetivo deste trabalho, podendo ser a curto ou longo prazo. Segundo Moretin e Toloí (2004) a maioria dos métodos de previsão de séries temporais baseia-se na idéia que o comportamento futuro depende de seu comportamento passado. Os métodos de previsão utilizados na prática podem variar de modelos simples e intuitivos até mais os complexos.

O métodos de previsão baseados em regressão consideram o comportamento passado da série, ou seja, um histórico. Para este trabalho, foram desenvolvidos dois métodos de previsão de séries temporais baseados em regressão linear e polinomial, e utilizados outros dois disponíveis no Weka. Estes, por sua vez, são o processo linear e o gaussiano. Todos eles são descritos a seguir.

3.1.1 Regressão Linear

A regressão linear é um modelo que expõe uma relação entre uma variável aleatória dependente y e uma variável independente x . Neste modelo, o objetivo é encontrar uma reta que

se ajuste ao máximos nas observações. A Figura 9 exibe esse processo. Neste caso, os exemplos de treinamento são pares ordenados na forma (x, y) , mostrados como pontos na figura. A equação de regressão linear tem a forma $y = ax + b$, mostrada como uma linha na Figura 9. O coeficiente a revela a inclinação da reta, isto é, o acréscimo ou decréscimo do valor de y em relação a x ; e b , ou intercepto, corresponde ao ponto de interseção da reta em relação a y , isto é, o valor de y quando $x = 0$.

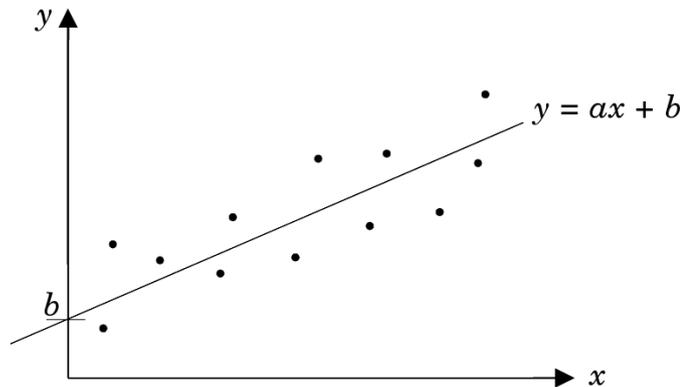


Figura 9: Modelo de regressão linear.

Segundo BUSTOS *et al.* (2011a) e BUSTOS *et al.* (2011b), a equação ajustada pelo modelo de regressão linear simples é dada por

$$I_1(x, y, t) = a_1(x, y)t + b_1(x, y), \quad (1)$$

em que $I_1(x, y, t)$ representa o tom de cinza do pixel associado à amplitude acústica de cada sequência na t -ésima tentativa, ou observação; (x, y) correspondem às coordenadas do pixel de interesse para cada sequência sísmica; $t = 1, 2, \dots, n$ corresponde à variável independente do tempo; o sub-índice 1 indica que um ajuste linear entre duas sequências sucessivas da sinal AVO foi usado para calcular a extrapolação deste sinal em um tempo futuro posterior à segunda sequência do sinal AVO; a_1 é o coeficiente de regressão (coeficiente angular); e, b_1 representa o coeficiente linear. Assim, a_1 e b_1 representam os parâmetros do ajuste da regressão linear no tempo.

Sendo este um modelo de regressão baseado em um ajuste linear da série temporal associada ao sinal AVO, então este se encaixa dentro do modelo Auto-Regressivo (AR) generalizado (WIENER, 1964), uma vez que o valor futuro da série pode ser obtido a partir de uma combinação linear dos valores passados da mesma.

Para encontrar uma imagem do sinal AVO extrapolada no tempo é necessário ter uma sequência em um determinado período de tempo, a qual corresponde ao instante de tempo inicial, ou seja, $t = 0$, e outra sequência em um instante de tempo posterior. Também, é necessário determinar os valores de a_1 e b_1 . Para tanto, b_1 pode ser definido assumindo que t é igual a 0 sobre a Equação 1. Da mesma forma, a_1 pode ser estimado através da Equação 1 assumindo que t corresponde ao período de tempo posterior em que é amostrada a segunda sequência do sinal AVO e substituindo o valor de b_1 .

Regressão Linear no software WEKA

O processo linear no WEKA usa a guia *Classifier* no *Weka Explorer*, especificamente na pasta rotulada *functions* que produz um modelo de uma função linear (isto é, uma soma ponderada) dos atributos de entrada.

No pacote WEKA, o formato geral da equação de regressão linear é dada por:

$$y = w_1x_1 + w_2x_2 + \dots + w_nx_n + c$$

onde y é o atributo de saída; x_1, \dots, x_n ; w_1, \dots, w_n são pesos numéricos; e c é uma constante adicional numérica.

A regressão linear encontra uma reta que melhor se ajusta aos dados calculados, isto é, a reta que minimiza a distância média de todos os pontos a ela. Seguindo a linha à frente no tempo, pode-se estimar pixels futuros, uma vez que se possa seguramente assumir que o crescimento continuará a ser linear.

3.1.2 Processo Gaussiano

Processo gaussiano (SEEGER, 2004) no WEKA implementa este método já bem conhecido para regressão e representa um conjunto de variáveis aleatórias, na qual qualquer subconjunto finito tem uma distribuição gaussiana conjunta (WILLIAMS; RASMUSSEN, 2006).

De acordo com MacKay (2009), o processo gaussiano pode ser definido como uma distribuição de probabilidade sobre um espaço de funções $G(y)$. Seja o valor médio \bar{y} e σ o desvio padrão, uma variável aleatória y tem uma distribuição gaussiana se sua função de densidade de probabilidade tiver a forma

$$G(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\bar{y}}{\sigma}\right)^2}.$$

O processo gaussiano, da mesma forma que o processo linear, usa no WEKA a guia *Classifier* no *Weka Explorer*, especificamente na pasta rotulada *functions*.

3.1.3 Regressão Quadrática

A regressão quadrática, assim como a regressão linear, expõe um modelo de relação entre uma variável aleatória dependente y e uma variável independente x . Porém, o objetivo é encontrar uma parábola que se ajuste, tanto quanto possível, às observações. A Figura 10 exhibe esse processo.

A equação de regressão quadrática tem a forma $y = ax^2 + bx + c$, mostrada como uma parábola na Figura 10. Os coeficientes a , b e c são chamados, respectivamente, de coeficiente quadrático, coeficiente linear e coeficiente constante ou termo livre, e são constantes, onde $a \neq 0$, pois caso contrário, a equação torna-se linear.

De acordo com BUSTOS *et al.* (2011a) e BUSTOS *et al.* (2011b), a equação ajustada pelo modelo de regressão quadrática é definida por

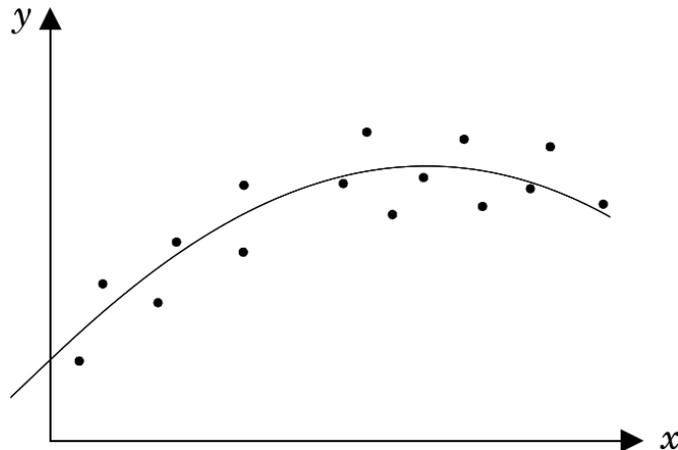


Figura 10: Modelo de regressão quadrática.

$$I_2(x, y, t) = a_2(x, y)t^2 + b_2(x, y)t + c_2(x, y), \quad (2)$$

em que $I_2(x, y, t)$, do mesmo modo que $I_1(x, y, t)$, representa o tom de cinza do pixel associado à amplitude acústica de cada sequência na t -ésima tentativa, ou observação; da mesma forma, $t = 1, 2, \dots, n$ corresponde à variável independente no tempo; porém, o sub-índice 2 indica que um ajuste polinomial quadrático entre três sequências sucessivas foi usado para calcular a extrapolação em um tempo futuro posterior à terceira sequência do sinal AVO; a_2 , b_2 e c_2 são parâmetros da equação quadrática e representam os parâmetros do ajuste da regressão quadrática.

Para encontrar uma imagem extrapolada do sinal AVO usando a regressão polinomial quadrática é necessário ter: uma sequência em um determinado período de tempo, a qual corresponde ao instante de tempo inicial, ou seja, $t = 0$; e outras duas sequências em instantes de tempos posteriores diferentes. Da mesma forma, é preciso encontrar os valores de a_2 , b_2 e c_2 . Para tanto, c_2 pode ser definido assumindo que t é igual a 0 sobre a Equação 2. Para definir a_2 e b_2 , é necessário resolver o sistema linear que se forma ao substituir os diferentes instantes de tempo posteriores na Equação 2 que estão associados às amostras das séries temporais conhecidas nestes períodos de tempo e substituir o valor de c_2 nelas.

Os levantamentos podem apresentar imagens sísmicas degradadas por ruídos sísmicos, o que dificulta o monitoramento sísmico do reservatório no processo de produção. Dente estes ruídos está um ruído coerente chamado ruído de rolamento superficial, ou *ground roll* comum em dados sísmicos obtidos em exploração sísmica feita em terra e que tem como características a presença de grandes amplitudes e baixas frequências (LEITE, 2007). No caso de imageamento em terra (*onshore*), as ondas sísmicas são produzidas pela reverberação entre a superfície e as camadas do subsolo próximo, gerando assim este tipo de ruído.

Dados com uma boa definição são necessários, pois podem prover informações relevantes dos parâmetros físicos do reservatório, tais como nível de saturação dos fluidos nas rochas-reservatórios, pressão dos poros nas rochas-reservatório, percentagem de porosidade nestas e a permeabilidade do reservatório. Esta, por sua vez, é necessária para predição dos volumes e caminhos da migração dos fluidos (CUNHA, 2002), tornando-se o principal parâmetro físico para o monitoramento de fluidos. Portanto, gerar imagens sísmicas de reservatórios com boa

qualidade e alta resolução espacial também é de grande importância estratégica para a indústria de gás e petróleo.

A sísmica 4D e mapas de diferenças podem identificar o fluxo dos fluidos no reservatório. Com isso, é possível prever o comportamento e as posições futuras destes fluidos por extrapolação. Neste ponto, o perfeito alinhamento destas sequências é extremamente necessário. Uma dessas técnicas usadas para isto, é o alinhamento geométrico, também chamado co-registro ou registro de imagens. Em sísmica, ele pode ser muito importante para rastrear o vapor d'água entre as sequências sísmicas 2D ou 3D, possibilitando o mapeamento do deslocamento de fluidos no reservatório através destas sequências (FOMEL; JIN, 2007) e atenuar parte do ruído presente nelas.

4 Co-Registro de Imagens

O co-registro de imagens (alinhamento geométrico) pode ser definido como a sobreposição de duas ou mais imagens – uma imagem base e uma imagem monitor. Isso pode ser feito mediante algumas operações, tais como rotação, translação e escala, que posicionam alinhadamente as imagens, gerando uma matriz dessa transformação (RODRIGUES, 2010). As imagens utilizadas na sobreposição geralmente possuem informações complementares e, portanto, podem ser uma importante ferramenta no monitoramento de fluidos.

Em aplicações de sísmica, o co-registro de imagens pode permitir o monitoramento dinâmico de fluidos em região de interesse de um reservatório em regime de produção. Para isso, esta técnica pode ser empregada para monitorar o vapor d'água presente no mesmo, pois o vapor tem a propriedade de agregar-se aos fluidos. Assim, rastreando-o, monitora-se também o deslocamento dos fluidos.

De acordo com Rodrigues (2010), o co-registro de imagens pode ser dividido em dois tipos: co-registro rígido e co-registro não rígido (ou elástico). O co-registro rígido é definido por quatro operações que podem ser aplicadas à imagem para o alinhamento: rotação, translação, escala e cisalhamento. Em alguns casos, quando, por exemplo, a imagem apresenta deformações elásticas, o co-registro rígido não é suficiente para o alinhamento correto. Assim, deve-se aplicar um co-registro não rígido, cuja transformação geométrica para o alinhamento é mais complexa, pois envolve correções elásticas.

A Figura 11 mostra um exemplo de co-registro rígido, onde apresenta duas imagens desalinhadas por translação e rotação e o resultado após a aplicação do co-registro de imagens. Na primeira linha, da esquerda para direita, tem-se: imagem modelo; imagem alvo; e imagem da diferença antes do co-registro de imagens. Já na segunda linha tem-se a imagem da diferença depois do co-registro de imagens. Pode-se observar que, quanto mais as regiões estiverem desalinhadas, mais pontos escuros aparecerão na imagem da diferença.

Essa é uma aplicação que se refere à área forense. Nela busca-se saber se a digital coletada (primeira linha, centro) corresponde à de uma pessoa cuja digital está armazenada, por exemplo, em um banco de dados (primeira linha, esquerda). Sem o co-registro de imagens, uma falsa conclusão é feita, pois as imagens estão desalinhadas. Isto é visivelmente percebido através da imagem da diferença realizada entre elas (primeira linha, direita). Porém, após a aplicação do co-registro de imagens, que as alinhou corretamente antes de criar a imagem da diferença, vemos que a digital coletada realmente pertence à da pessoa cuja digital está armazenada no banco de dados.

Todavia, vale salientar que, para esta aplicação, é recomendável a aplicação do co-registro rígido. O uso do co-registro não rígido que, mediante correções elásticas, tentaria alinhar as imagens, poderia “viciar” os resultados, causando falsas conclusões.

A técnica de co-registro de imagens é bastante utilizada com sucesso em outras áreas, como medicina (HOLDEN *et al.*, 2000), (TANG *et al.*, 2006). Além disso, outras áreas se beneficiam de suas características, onde sua aplicação pode ir desde a medição de deformação do solo a partir de imagens aéreas multitemporais (AYOUB *et al.*, 2009), até o reconhecimento de faces, como proposto por (QUEIROLO *et al.*, 2007). Desta forma, sua utilização é absolutamente dependente do contexto do problema a ser abordado (van den Elsen *et al.*, 1993).

Existem diversas maneiras para se efetuar o co-registro de imagens. Pode-se por exemplo,

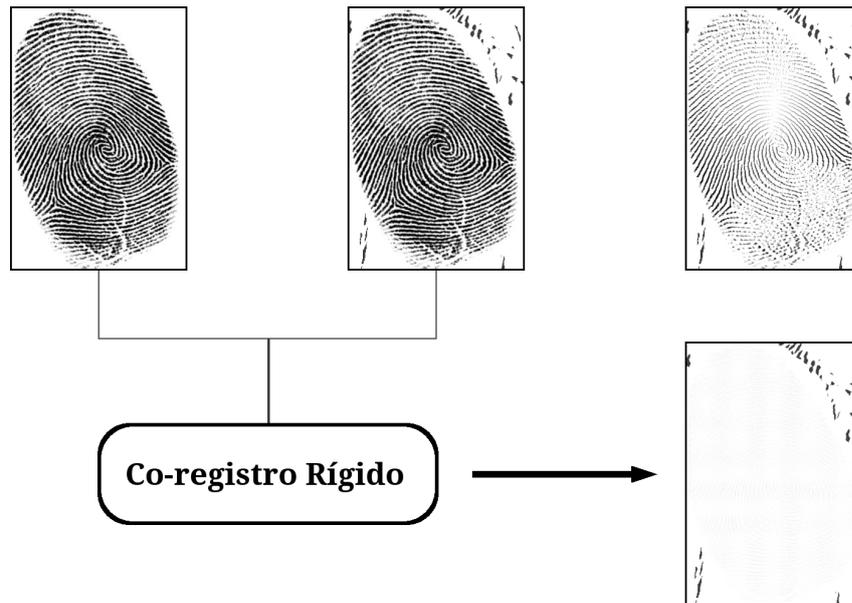


Figura 11: Co-registro rígido.

utilizar métodos baseados na seleção de pontos característicos, no valor de intensidade do pixel ou na segmentação de bordas. Os métodos baseados na intensidade do pixel, também chamados de *Block Matching*, comparam estas intensidades e, mediante operações, tentam estabelecer a correspondência entre pontos nas imagens. Esta correspondência pode ser encontrada utilizando simples comparação de regiões nas imagens, definindo-se um critério de comparação que envolve uma métrica de similaridade.

A escolha da métrica é importantíssima para o co-registro de imagens, pois ela refletirá diretamente na qualidade deste. Dentre elas, está o atributo de similaridade local (FOMEL, 2007a) que será bastante relevante para esta pesquisa.

4.1 Atributo de Similaridade Local

Fomel (FOMEL, 2007a) define atributos locais utilizando inversão regularizada e demonstra a sua utilidade para medir frequências locais de sinais sísmicos e similaridade local entre diferentes conjuntos de dados. Estes atributos são úteis para monitoramento sísmico, co-registro de dados multicomponentes e análise tanto de velocidades quanto de amplitudes. Além disso, o atributo de similaridade local é definido usando o método de regularização *shaping* (FOMEL, 2007b) que suaviza a imagem, porém preserva suas bordas. Ele é utilizado para controlar a localização e suavização dos atributos locais.

Segundo Fomel e Jin (2007), o coeficiente de correlação c entre duas sequências de dados a_t e b_t , em que t é uma função do tempo, é definido como

$$c = \frac{\sum_t a_t b_t}{\sqrt{\sum_t a_t^2 \sum_t b_t^2}}$$

e varia entre 1 (correlação perfeita) e -1 (correlação perfeita de sinais com polaridade diferente).

Para definir o atributo de similaridade local (FOMEL, 2007a) parte-se da observação de que o quadrado do coeficiente de correlação pode ser representado como o produto de duas quantidades, isto é, $c^2 = pq$, onde

$$p = \frac{\sum_t a_t b_t}{\sum_t b_t^2}$$

corresponde à solução do problema de minimização por mínimos quadrados,

$$\min_p \sum_t (a_t - p b_t)^2$$

e

$$q = \frac{\sum_t a_t b_t}{\sum_t a_t^2}$$

é a solução de minimização por mínimos quadrados

$$\min_q \sum_t (b_t - q a_t)^2 .$$

Sendo R um operador de regularização introduzido para reforçar um comportamento desejado, como a suavização, a similaridade local γ_t é definida como um sinal variável representada pelo produto de dois sinais variáveis p_t e q_t , ou seja, $\gamma_t = p_t q_t$. Estes sinais variáveis, por sua vez, correspondem as soluções dos problemas por mínimos quadrados regularizados representados pelas igualdades

$$\min_{p_t} \left(\sum_t (a_t - p_t b_t)^2 + R[p_t] \right)$$

e

$$\min_{q_t} \left(\sum_t (b_t - q_t a_t)^2 + R[q_t] \right) .$$

A aplicação da similaridade local para o problema de co-registro de imagens em lapso de tempo, consiste em “apertar” e “esticar” a imagem monitor em relação à imagem base, enquanto calcula-se o atributo de similaridade local. Em seguida, é selecionada a tendência mais forte de similaridade do painel de atributos e aplica-se o deslocamento correspondente à imagem (FOMEL; JIN, 2007).

5 Metodologia

5.1 Tecnologia de Software

Uma das principais ferramentas utilizadas pelas indústrias de exploração de hidrocarbonetos são os softwares de processamento sísmico. Estes por sua vez, surgem visando facilitar ainda mais o trabalho de tais empresas. Todavia, esses programas costumam ser, na sua grande maioria, proprietários e caros. Assim, somente empresas com um alto poder aquisitivo podem adquiri-los, enquanto que a maioria das instituições de ensino e pesquisa ficam fora desse seleto grupo.

Porém, visando contornar essa situação, softwares poderosos de código aberto são desenvolvidos e disponibilizados no mundo inteiro. Entre eles, está uma das mais importantes bibliotecas de processamento de sinais e imagens sísmicas: o Madagascar. Por ser de distribuição livre e código aberto, passou-se a utilizá-lo amplamente em todo o mundo. Em seu repositório estão presentes os dados sintéticos e reais que foram utilizados neste trabalho. Além disso, ainda são disponibilizados programas para co-registro de imagens, o que constitui um ponto relevante para sua utilização nesta pesquisa, pois estes poderão ser modificados para adequarem-se às nossas necessidades.

Além disso, como geralmente as soluções do campo de processamento de imagens requerem um extenso trabalho experimental envolvendo simulações e testes, para a implementação dos algoritmos para criação dos mapas de diferenças e o co-registro de imagens foi utilizado o software Madagascar.

5.1.1 Madagascar

O Madagascar é um pacote de programas em código aberto e distribuído pela licença padrão Licença Pública Geral (GPL, do inglês *General Public License*). Ele permite realizar processamento de imagens e dados sísmicos, possibilitando trabalhar com dados multidimensionais e experimentos computacionais reprodutíveis.

Basicamente, o Madagascar destina-se a pesquisadores que trabalham com imagens digitais, processamento de dados geofísicos e áreas afins. Embora iniciado em 2003 pelo Dr. Sergey Fomel da *Bureau of Economic Geology* (BEG) da University of Texas at Austin, somente em 2006 ele foi liberado, contando hoje com mais de 300 programas principais, 3000 testes, documentos reprodutíveis e *papers*. Também é gerenciado por diferentes grupos de pesquisas incluindo o *Seismic Laboratory for Imaging and Modeling* (SLIM) e *Uniform Building Code* (UBC).

De acordo com Fomel e Hennenfent (2007), o Madagascar é um pacote de software multicamadas, conforme pode ser observado na Figura 12. Desta forma, os usuários podem utilizá-lo de diferentes maneiras: mediante linha de comando, haja vista que o Madagascar é antes de tudo uma coleção de programas de linha de comando; com fluxos de processamento utilizando *Software Construction* (SCons) para realizar experimentos numéricos reprodutíveis que podem ser feitos não apenas usando linha de comando, mas também em Matlab, Mathematica, Python, ou outros pacotes sísmicos (por exemplo, Seismic Unix); e por fim, a documentação, que corresponde à camada mais superior e que faz uso do Latex para gerar arquivos em Formato de Documento Portátil (PDF – do inglês *Portable Document Format*), Linguagem de Marcação

de Hipertexto (HTML – do inglês *HyperText Markup Language*) e Media Wiki (MEDIAWIKI, 2010), um software Wiki escrito em PHP e desenvolvido originalmente para o uso no Wikipédia.

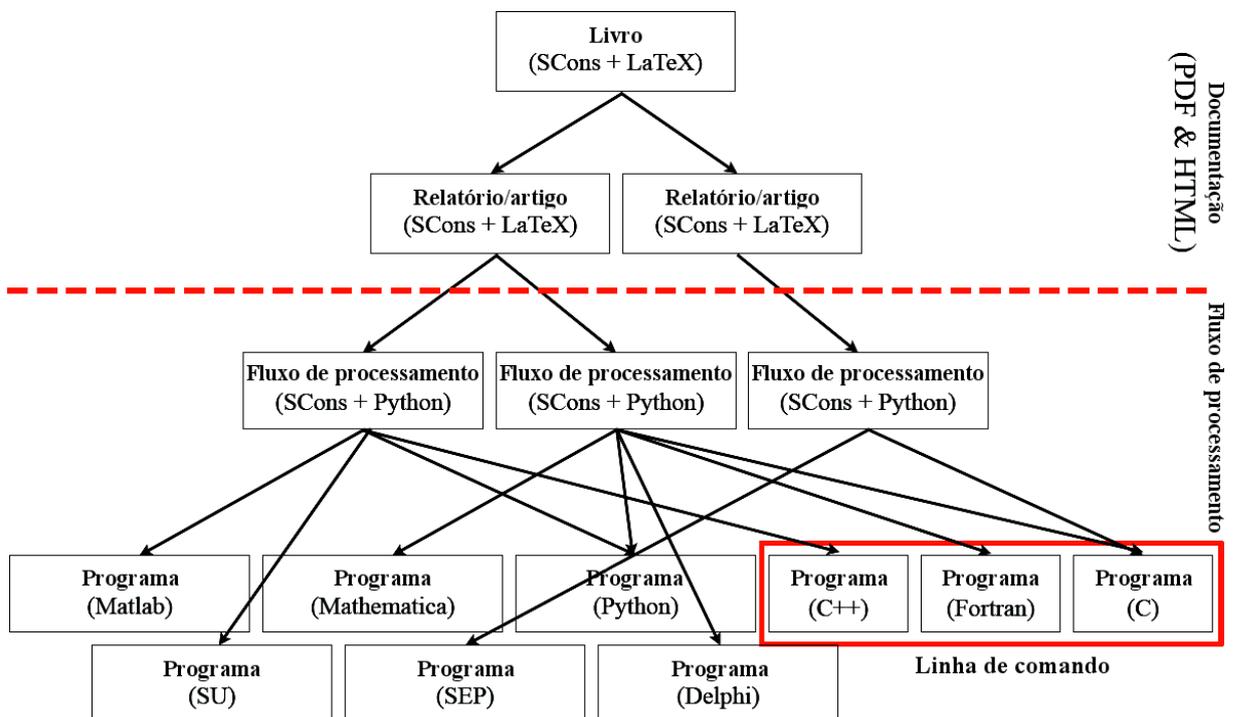


Figura 12: Visão geral do Madagascar. Adaptado de (FOMEL; HENNENFENT, 2007)

O software Madagascar usa o *Regularly Sampled Format* (RSF) para arquivos de dados, semelhante ao utilizado na biblioteca *Stanford Exploration Project Library* (SEPlib) desenvolvida pelo Projeto de Exploração Stanford (PES). Este formato de arquivo descreve hipercubos regularmente amostrados em até nove dimensões e se baseia na filosofia do Unix, cada arquivo RSF corresponde a um texto legível e simples.

O Madagascar contém programas para conversão entre diferentes formatos de dados como SEG-Y e os utilizados pelo *Seismic Unix* (SU). Atualmente, ele adota o formato de arquivo *Vplot* para arquivos gráficos, que também foi desenvolvido no SEP.

Tal como o SU e o SEPlib, os programas do Madagascar podem ser executados mediante o uso de linhas de comandos. No entanto, linhas de comando tornam o trabalho rotineiro e cansativo, tornando-se inconveniente para algumas pessoas. Desta forma, o Madagascar, visando manipular esses arquivos mais facilmente, provê uma alternativa fazendo uso de um construtor chamado SCons. Este construtor é um pacote com código-fonte aberto para construção de software, sendo uma alternativa para outros construtores como o Make. No SCons, seus arquivos estão dentro de *scripts* chamados *SConstructs* que são escritos em Python, uma linguagem de programação dinâmica orientada a objeto.

5.1.2 WEKA

O software WEKA é um programa desenvolvido na Universidade de Waikato na Nova Zelândia, implementado em Java e de distribuição livre e código aberto. Este programa possui várias técnicas de aprendizagem de máquina que fornecem relatórios com informações analíticas e estatísticas dos dados utilizados.

O Weka apresenta basicamente quatro opções: um modo simples de linha de comando; um modo *Explorer*, onde é possível explorar os dados e os algoritmos disponíveis; um modo experimental; e um outro modo que permite uma definição de fluxo de operações.

O modo *Explorer* do WEKA apresenta várias seções com diferentes algoritmos para manipular os dados, cada um com seu objetivo. Dentre estas seções, está a chamada *Forecasting* que permite a utilização do pacote *Time Series Forecasting* (versão 1.0.6). Este pacote é responsável pela realização dos experimentos utilizando regressão linear e processo gaussiano pelo WEKA e pode ser instalado diretamente por ele.

5.2 Dados Utilizados

Fomel e Long (FOMEL; JIN, 2007) ilustram um método que utiliza o atributo de similaridade local no co-registro de imagens e demonstram a sua eficácia, utilizando experimentos sintéticos e dados reais. Estes, por sua vez, estão disponíveis no repositório de dados do Madagascar e no caso deste último, foram utilizados neste trabalho.

A Figura 13 mostra dados de campo reais para monitoramento de fluidos em Duri, Indonésia reproduzido por (FOMEL; JIN, 2007) de (LUMLEY, 1995a, 1995b). Neste campo, um registro base foi adquirido antes da injeção de vapor para demonstrar a repetibilidade dos dados. Depois, outros cinco registros foram feitos em uma média de intervalos de 4 meses, que varia de 2 a 19 meses após o início da injeção do vapor. Observando-se o eixo do tempo, a parte superior do reservatório principal está em cerca de 1,6 s e a base está um pouco abaixo dos 2 s (LUMLEY, 1995b).

Os dados de campo reais (Figura 13) mostram as seções 3D da linha sísmica empilhada (LUMLEY, 1995b) das traças brutas, representando o sinal AVO que corresponde a variação da amplitude da onda sísmica refletida com o afastamento (distância entre a fonte e o receptor). Além disso, estas traças devem apresentar um comportamento hiperbólico, que não pode ser visualizado nas imagens devido ao afastamento ser muito curto (pouco mais de 600 m).

As anomalias no sinal AVO podem ser melhor identificadas observando-se uma representação colorida, como a realizada com a Figura 13 e mostrado na Figura 14. Neste processo de conversão, que foi realizado no ambiente do Madagascar, quanto mais próximo do branco for o valor do pixel, mais vermelho será sua representação e quanto mais próximo do preto, mais azul ele será.

Desta forma, a cor azul nas imagens representa a amplitude registrada que tende a decrescer conforme aumenta-se o afastamento, tendência conhecida como *falls*. Também, esta cor corresponde às regiões com baixas amplitudes e impedância acústica, representando a interface entre a base do reservatório e as outras camadas abaixo dele. Entretanto, o crescimento irregular (*rising*), ou anomalia, é representado pela cor vermelha e corresponde às regiões com grandes amplitudes e impedância acústica. Esta cor está associada a interface entre o topo do reservatório (e conseqüentemente ao vapor injetado nele) e as outras camadas acima dele. Assim, tem-se que quanto mais próximo da cor vermelha for um pixel, maior é a sua amplitude e quanto mais próximo do azul, menor é a amplitude naquele ponto.

Embora o reservatório esteja entre cerca de 1,6 s e 2 s, várias outras interfaces coloridas com vermelho e azul podem ser identificadas por toda a imagem, principalmente no topo (entre 0,4 s e 1 s) e na parte inferior (um pouco abaixo de 2 s). Isso constitui uma espécie de ruído,

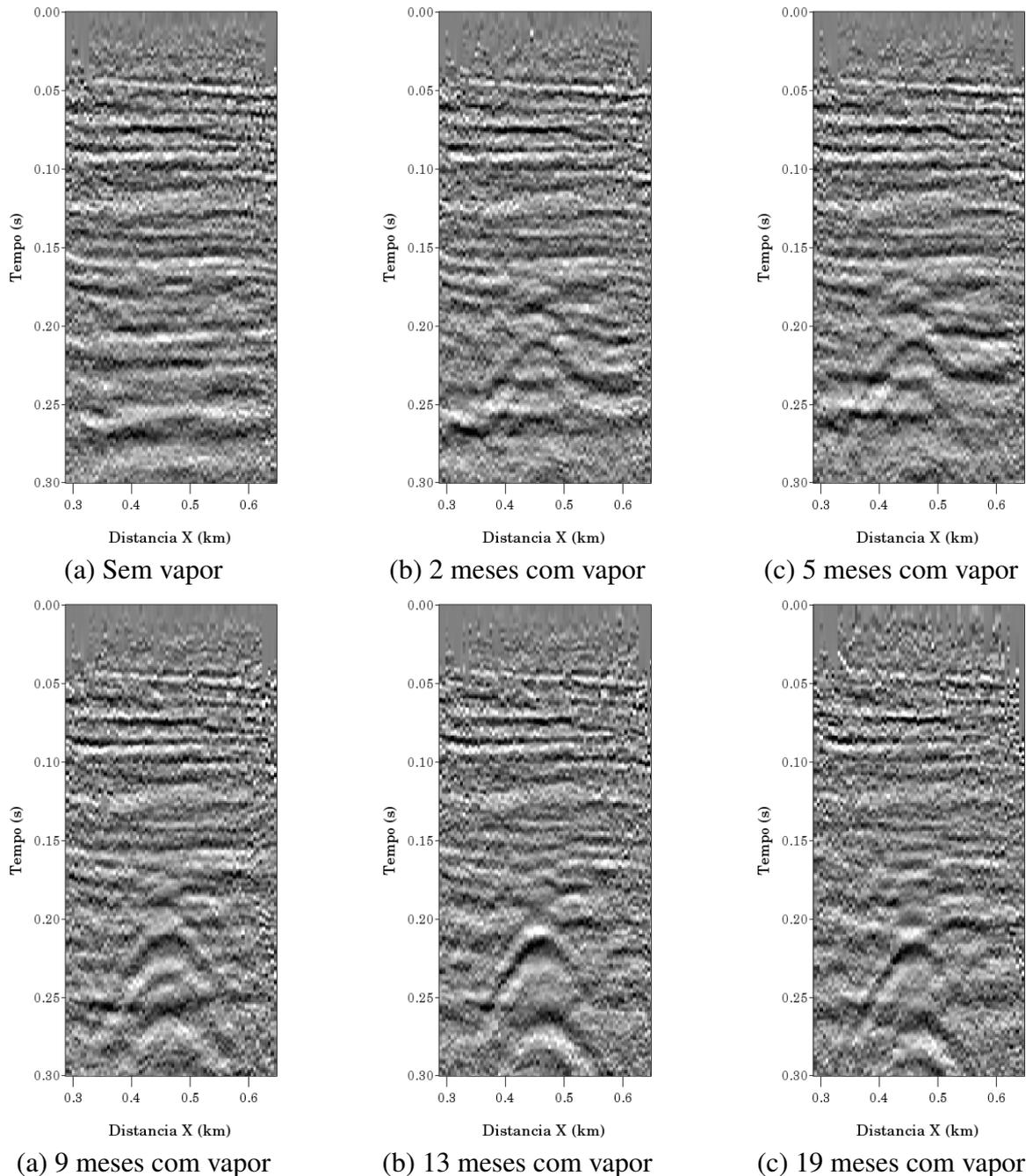


Figura 13: Campo de Duri, Indonésia. Adaptado de (FOMEL; JIN, 2007).

pois comprometem o resultado final, haja vista que um especialista poderia interpretá-las como possíveis reservatórios e não são.

5.2.1 Dados Utilizados no WEKA

Como dito anteriormente, o topo do reservatório principal está em cerca de 1,6 s e a base um pouco abaixo de 2 s. Esta área corresponde à região de interesse a ser monitorada. Desta forma, as outras regiões da imagem podem ser desconsideradas, uma vez que o principal objeto de interesse que se quer monitorar é o reservatório.

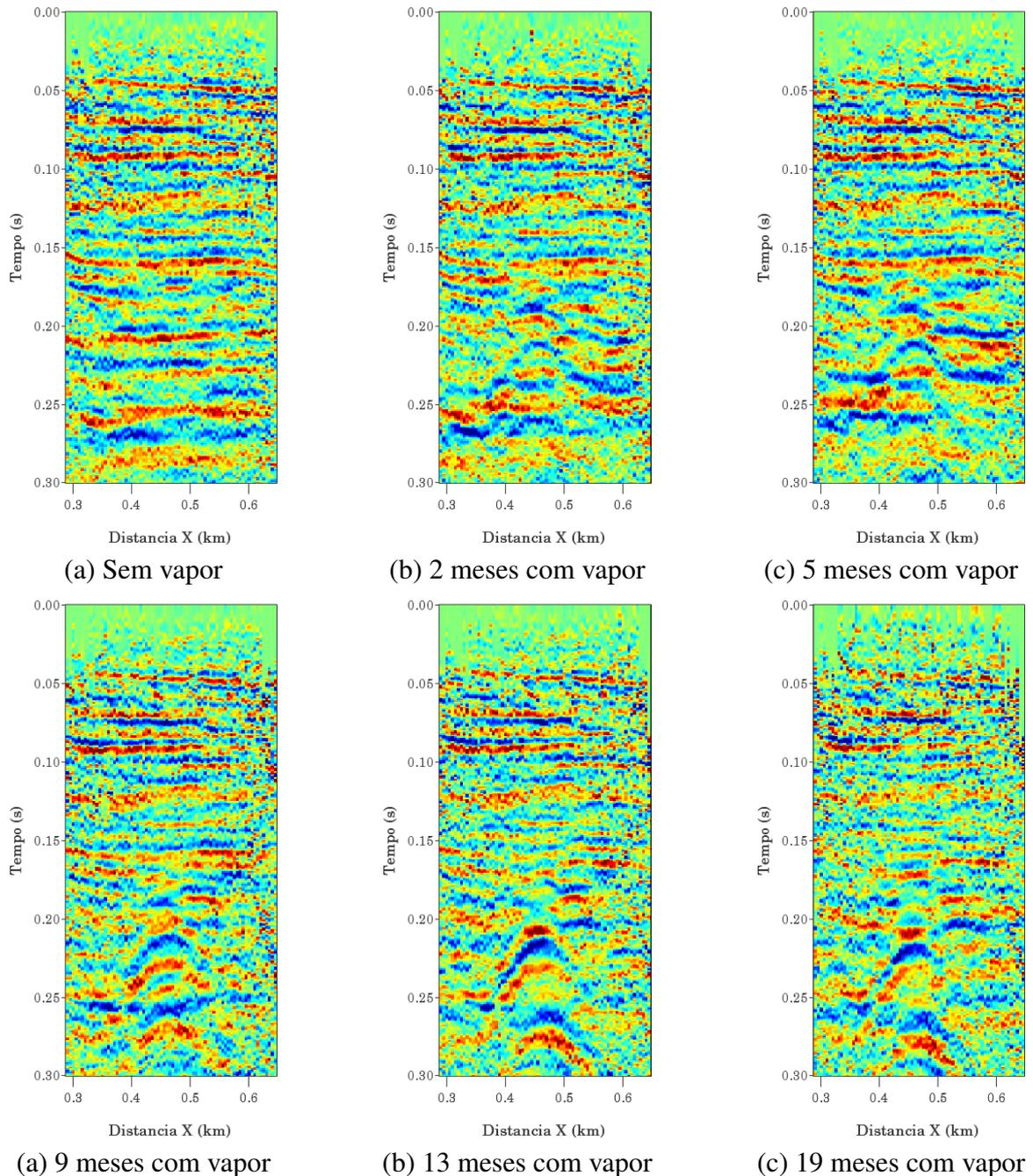


Figura 14: Modelo real do campo de Duri, Indonésia, colorido. Adaptado de (FOMEL; JIN, 2007).

Este procedimento foi realizado visando reduzir as limitações computacionais do processamento no WEKA. Desta forma, realizou-se um corte sobre as imagens utilizadas nesta região, isto é, entre 1,5 s e 2,1 s, onde as técnicas de regressão do WEKA foram aplicadas. As sequências utilizadas foram mapas de diferenças obtidos após a aplicação do co-registro de imagens, ambos descritos a seguir, uma vez que os resultados obtidos a partir deles mostraram-se superiores, como apresenta a Seção 6. A Figura 15 exhibe estas sequências, enquanto a Figura 16 mostra uma versão colorida destas imagens.

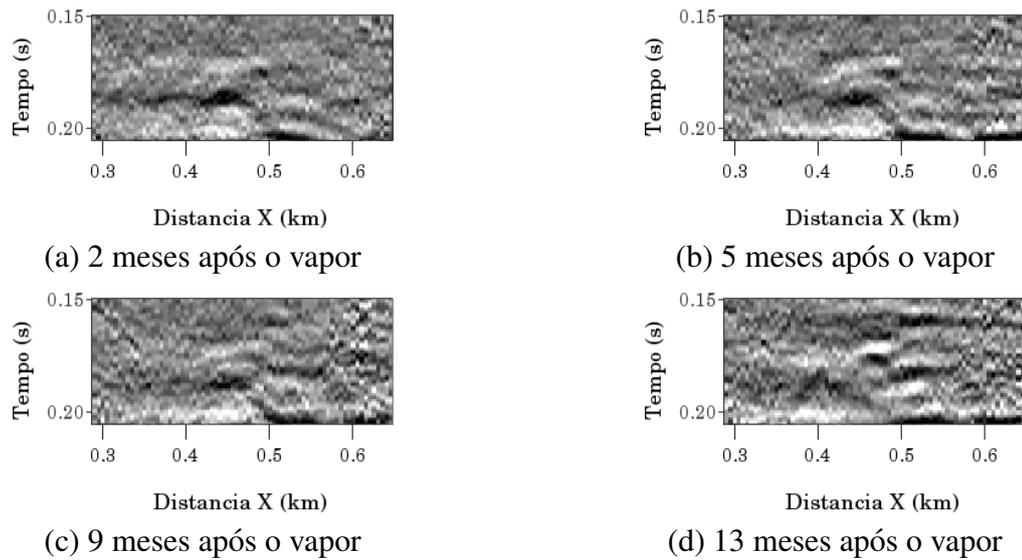


Figura 15: Corte após o co-registro de imagens. Adaptado de (FOMEL; JIN, 2007).

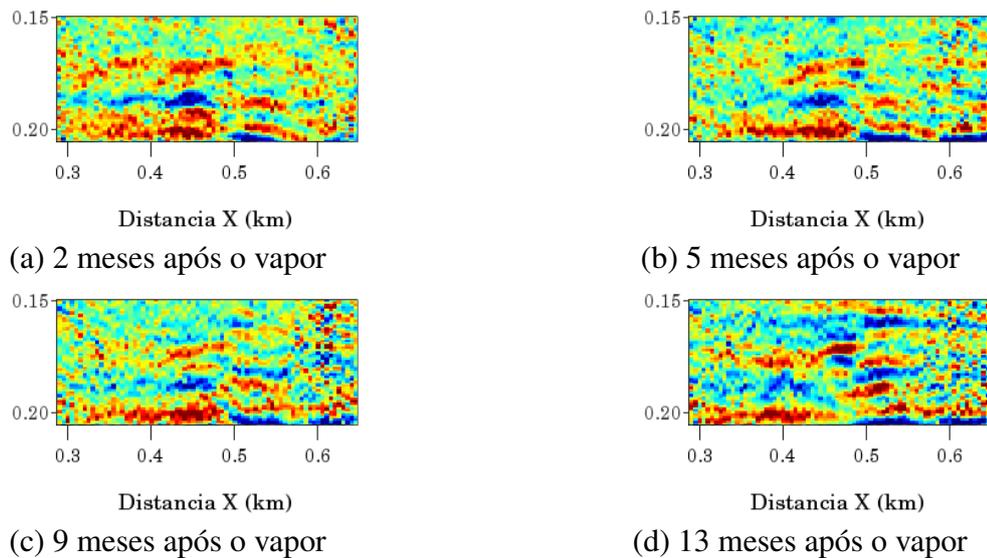


Figura 16: Representação colorida do corte após o co-registro de imagens. Adaptado de (FOMEL; JIN, 2007).

5.3 Metodologia Proposta

5.3.1 Procedimento de Mapas de Diferenças

Os mapas de diferenças correspondem ao principal objeto deste trabalho, pois é sobre eles que são aplicadas as técnicas de extrapolação. A Figura 17 mostra dados sintéticos utilizados por (FOMEL; JIN, 2007) para demonstrar a eficiência do seu algoritmo de co-registro de imagens. Estas imagens representam o interior da terra, onde dois levantamentos sísmicos teriam ocorrido em dois instantes de tempo distintos. Na Figura 17(a) tem-se um modelo adquirido em um determinado momento. A Figura 17(b) representa a aquisição deste mesmo modelo em um instante de tempo posterior. Nela, assume-se que as mudanças ocorrem tanto no reservatório (objeto no centro das imagens) como na subsuperfície rasa (topo das imagens) (FOMEL; JIN,

2007).

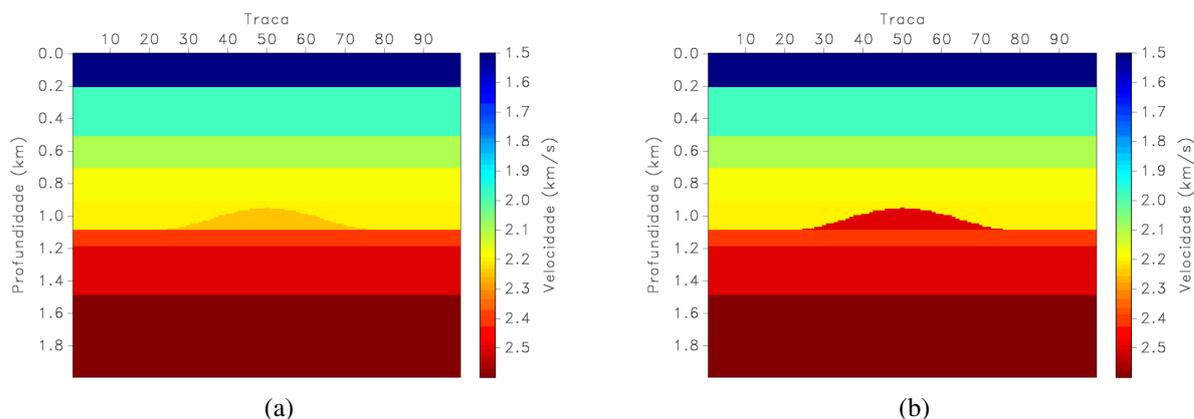


Figura 17: Modelo sintético em (a) um instante de tempo e (b) em um instante posterior. Adaptado de (FOMEL; JIN, 2007).

Para a geração dos mapas de diferenças, é realizada a subtração entre as imagens em lapso de tempo posteriores e a imagem base, ou seja, em um instante de tempo inicial. No modelo sintético, isto é feito fazendo-se a subtração entre a Figura 17(b) e a Figura 17(a). Neste caso, este processo foi realizado antes do co-registro de imagens. O resultado pode ser visto na Figura 18 e é possível observar diversas linhas horizontais obtidas a partir do não alinhamento das camadas entre estas imagens. Estas linhas são desprezíveis, uma vez que deseja-se monitorar apenas a região do reservatório (centro da imagem).

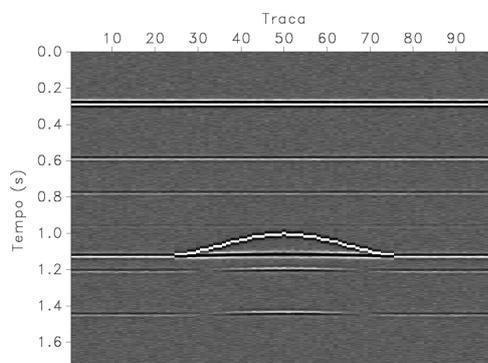


Figura 18: Mapa de diferença para os dados sintéticos. Adaptado de (FOMEL; JIN, 2007).

A Figura 19 mostra os mapas de diferenças obtidos a partir de dados de campos reais. A partir destas imagens monitoram-se as mudanças que ocorreram no reservatório no decorrer do tempo. Antes de aplicar o co-registro de imagens, as diferenças reais nas imagens entre 2 e 19 meses são obscurecidas por artefatos coerentes (*ground roll*), que são causados por mudanças na velocidade, tanto na superfície rasa quanto na região do reservatório (FOMEL; JIN, 2007). Estes artefatos também são visivelmente percebidos na área abaixo do reservatório e no topo das imagens, principalmente nos registros de 13 e 19 meses após a injeção do vapor d'água. A utilização do co-registro de imagens também pode diminuir estes artefatos indesejáveis, ressaltando ainda mais sua importância.

Além disso, novos mapas de diferenças foram gerados, estes por sua vez, após a aplicação do co-registro de imagens.

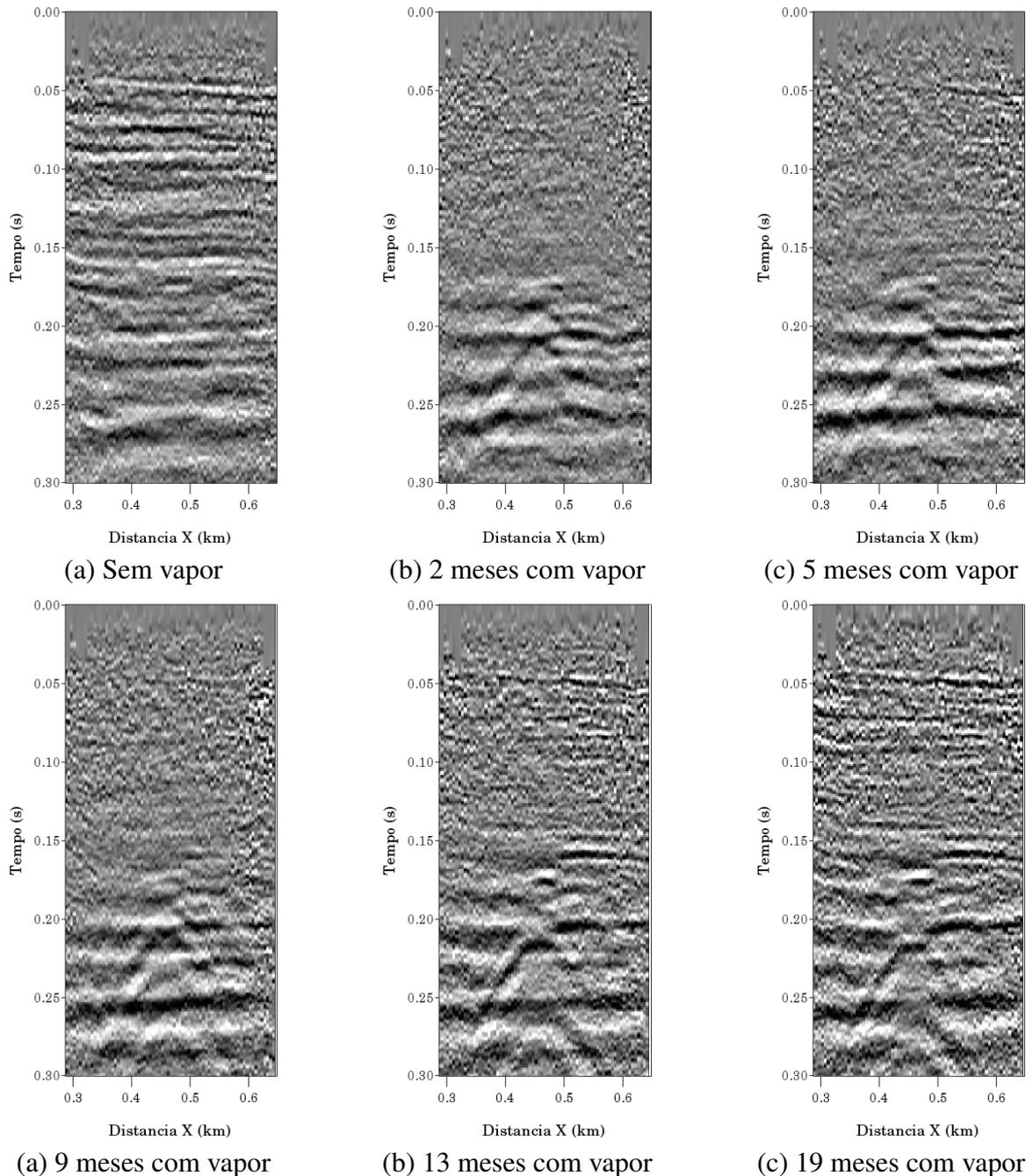


Figura 19: Diferença antes do co-registro de imagens. Adaptado de (FOMEL; JIN, 2007)

5.3.2 Procedimento de Co-registro

O co-registro de imagens representa a sobreposição de duas ou mais imagens de forma que haja correspondência entre seus pontos. Ele pode permitir o monitoramento de fluidos em um reservatório em regime de produção. Neste trabalho, o co-registro foi realizado com o algoritmo descrito na Subseção 4.1 e aplicado sobre as imagens mostradas na Figura 13. A Figura 20 exibe o resultado da criação do mapa de diferença utilizando os dados sintéticos e após a aplicação do co-registro de imagens. Comparando-se a Figura 18 com a Figura 20, que representam os mapas de diferenças obtidos antes e depois do co-registro de imagens, respectivamente, é possível observar que sua aplicação permitiu realçar a região do reservatório e diminuir os

efeitos causados pelo não alinhamento entre as imagens. No entanto, ainda persiste uma linha horizontal no topo da imagem.

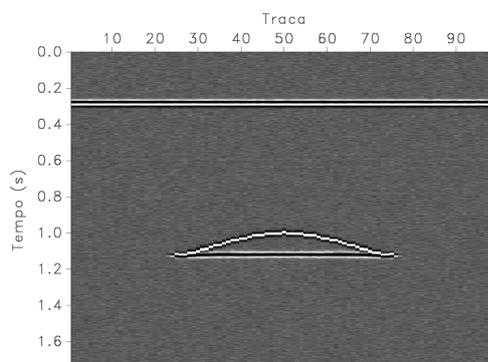


Figura 20: Mapa de diferença para os dados sintéticos após o co-registro de imagens. Adaptado de (FOMEL; JIN, 2007).

A Figura 21 mostra os resultados de campos reais depois da aplicação do método proposto por (FOMEL; JIN, 2007) sobre os registros de 2 a 19 meses. Pode-se perceber que parte dos artefatos coerentes foram eliminados, comprovando assim a eficiência do método quando comparado à imagem sem co-registro de imagens (Figura 19). O algoritmo é utilizado para aumentar a precisão da interpretação que o especialista possa fazer das feições do reservatório. Embora melhor, este ainda não é o resultado ideal, pois ainda persistem artefatos de imagens produzidos pelo ruído coerente *ground roll* nas sequências das seções sísmicas. Isso pode ser constatado observando que, mesmo com a aplicação do método, os efeitos indesejáveis situados abaixo do reservatório ainda persistem.

5.3.3 Procedimento para Extrapolação

O processo de extrapolação, no contexto deste trabalho, busca estimar o comportamento futuro dos fluidos em reservatórios petrolíferos. Para realizar a extrapolação, optou-se por implementar com uma interface gráfica, o algoritmo que obteve melhor resultado, este mostrado na Seção 6. Assim, usando a linguagem de programação Java, implementou-se o algoritmo regressão linear definido pela Equação 1. Esta linguagem foi escolhida devido à familiaridade com a sua sintaxe e sua manipulação com objetos gráficos, além de ter vários métodos já implementados para trabalhar com processamento de imagens.

A Figura 22 mostra o programa criado para realizar a extrapolação linear no tempo. Para utilizar tal aplicativo, o usuário deve clicar no botão “Imagem Base”, onde deverá selecionar uma imagem em um tempo inicial. Em seguida, deverá selecionar uma sequência em um tempo posterior à primeira imagem selecionada, isto realizado através do botão “Imagem Posterior”. Em seguida, deverá preencher os campos “Variação de Tempo” que corresponde à variação de tempo entre as duas sequências carregadas pelos botões anteriores e “Extrapolar Para”, que solicita para que instante os dados devem ser extrapolados a partir da imagem base.

Ao acionar o botão “OK”, o usuário será levado para uma outra janela, esta mostrada na Figura 23. Nela, o usuário poderá visualizar a imagem base (à esquerda), a imagem posterior à imagem base (no centro) e a sequência obtida pela extrapolação linear. Também, é possível voltar para a janela anterior, sair do aplicativo e principalmente, salvar os dados nos formatos de imagens mais conhecidos.

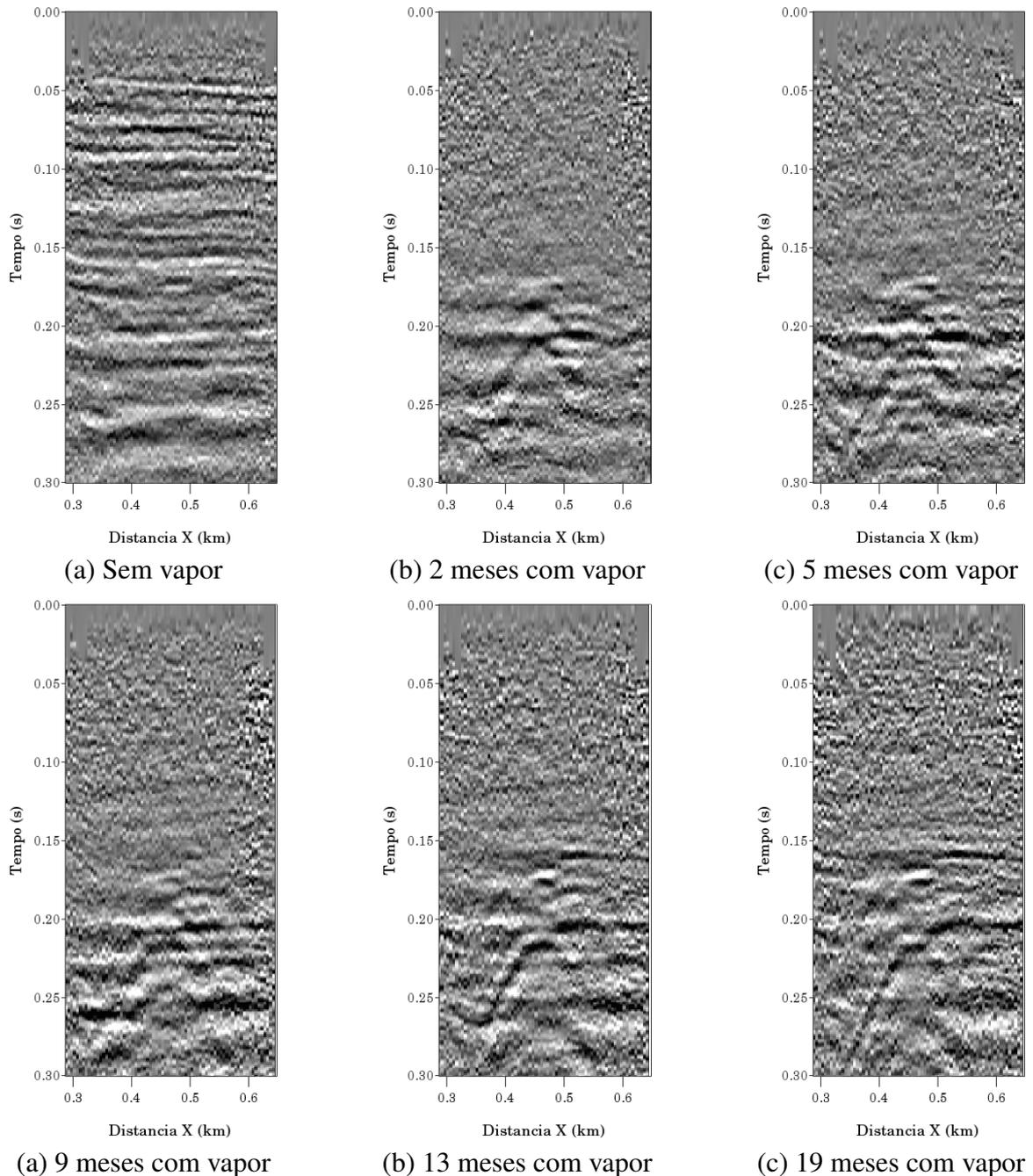


Figura 21: Resultado após o co-registro de imagens. Adaptado de (FOMEL; JIN, 2007).

Extrapolação usando o WEKA

Utilizando os métodos de regressão linear e o processo gaussiano implementados no WEKA, foi gerada uma sequência conhecida para fins de validação dos algoritmos. Desta forma, buscou-se gerar a sequência que corresponde a 13 meses após a injeção do vapor e depois da aplicação do co-registro. Em ambos os métodos, foram usadas duas imagens anteriores a esta, isto é, as sequências de 5 e 9 meses.

As extrapolações por meio de regressão linear e processo gaussiano foram executadas a partir do pacote *Time Series Forecasting* (versão 1.0.6), que fornece um ambiente para previsão

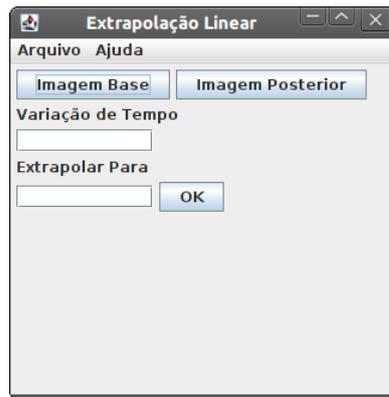


Figura 22: Extrapolador Linear - Parte 1.

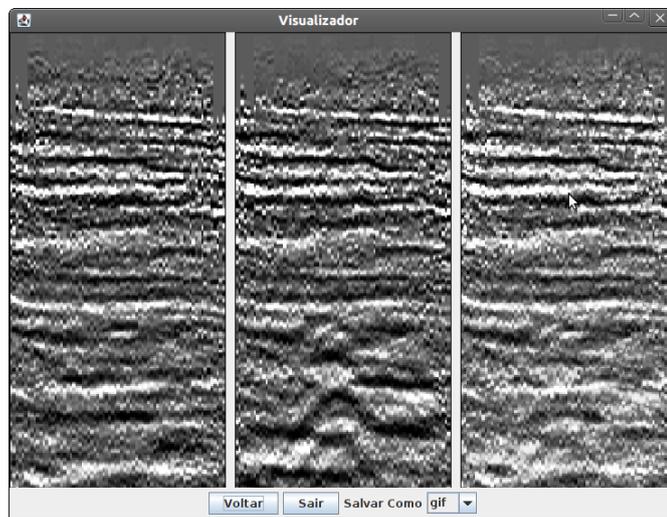


Figura 23: Visualizador.

de séries temporais para o WEKA (versão 3-7-5).

Antes de serem processadas no WEKA, inicialmente essas imagens foram convertidas para o formato de texto, especificamente para Portable Gray Map (PGM). Neste tipo de arquivo, cada pixel pode ser representado como um número inteiro positivo em Código Padrão Americano para o Intercâmbio de Informação (ASCII, do inglês American Standard Code for Information Interchange) que varia entre 0 (preto) a 255 (branco), sendo o intervalo entre estes valores a variação de níveis de cinza. Em seguida, devido à limitação do computador, estes pixels foram divididos em arquivo Comma-Separated Values (CSV) e processados pelo WEKA em grupos de 500 pixels, onde os algoritmos foram executados.

Depois de todo o processamento no WEKA, cada saída gerada, ou seja, os valores futuros estimados dos novos pixels, foram todos agrupados em uma única tabela, onde aplicou-se técnicas de arredondamento para os pixels com valores decimais. Além disso, como o valor dos pixels no formato PGM varia entre 0 e 255, valores abaixo de 0 ou acima de 255 devem ser tratados para que estes não fiquem acima ou abaixo do fluxo de pixel. Assim, também foram aplicadas técnicas de *overflowing* para tratar essas exceções.

Por fim, após as imagens serem reconstruídas, estas foram convertidas para o formato Portable Network Graphics (PNG), onde determinou-se o erro para saber o nível de similaridade entre a imagem real e suas extrapoladas. Este erro é discutido a seguir.

5.3.4 Determinação do erro

Para verificar a eficiência e o desempenho dos algoritmos propostos, buscou-se extrapolar sequências sísmicas conhecidas. Desta forma, procurou-se obter extrapolações para as sequências de 13 e 19 meses utilizando os métodos de regressão linear e quadrática, definidos pelas equações 1 e 2, respectivamente. Desta forma, é possível quantificar numericamente o quão exato é o processo adotado, pois permite comparar as imagens extrapoladas com as próprias.

Assim, para cacular o nível de semelhanças entre as imagens extrapoladas e a sequência sísmica real, foi utilizado a biblioteca ProEikon (KIM, 2010). O ProEikon é uma biblioteca com algoritmos de Processamento e Análise de Imagens e Vídeos (PAIV). Esta biblioteca possui um programa chamado DistG que permite calcular a diferença entre duas imagens em níveis de cinza, possibilitando assim, quantificar numericamente o quão exatas são as imagens geradas.

O DistG possui, dentre outros estimadores, o Erro Médio Absoluto (MAE – do inglês *Mean Absolute Error*), que é uma grandeza utilizada para medir o quão próximas as previsões e as predições estão dos resultados finais, além de ser uma medida comum de erro de previsão em análise de séries temporais. Desta forma, este estimador permite avaliar, em porcentagem, o erro entre os dados previstos pela extrapolação numérica e os dados reais utilizados para gerar o sinal AVO da sequência sísmica amostrada.

Seja x_i o valor real obtido, y_i os dados previstos e n o número de avaliações consideradas, o MAE é representado pela seguinte equação

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|.$$

Um outro estimador disponível nesse programa é o Erro Médio Quadrático (RMSE – do inglês *Root Mean Squared Error*). O RMSE mede a variação dos valores estimados ao redor dos valores medidos, enquanto que o MAE indica o afastamento (desvio) médio absoluto dos valores previstos em relação aos valores observados (ZACHARIAS *et al.*, 1996). O ideal é que os valores do MAE e do RMSE sejam o mais próximo de zero possível. O RMSE é dado pela seguinte igualdade

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}.$$

Os resultados obtidos por esses estimadores são apresentados no próximo capítulo.

5.3.5 Verificando a Estacionariedade

Quando se busca estimar uma série temporal, a principal idéia que se deve ter é o conceito de estacionariedade (BUENO, 2008). Uma série é dita estacionária se sua média e sua variância são constantes ao longo do tempo e se o valor da covariância entre dois períodos de tempo depende apenas da distância entre esse dois períodos (ENDERS, 2004). Isso significa que o comportamento da série não se altera com o passar do tempo.

Assim, para verificar a estacionariedade da série temporal, foram calculadas as médias, as variâncias e as covariâncias sobre cada sequência sísmica utilizada para gerar as estimativa futuras, estas por sua vez, antes e depois da aplicação do co-registro de imagens.

Seja x_i o valor do pixel e n o número de pixels da imagem, a equação ajustada para o cálculo da média, \bar{x} , sobre cada imagem é definida pela equação

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

A variância σ^2 da variável aleatória x é dada pela igualdade

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

e geralmente é designada por $\text{var}(x)$, σ_x^2 , ou simplesmente σ^2 .

A variância e a covariância são na teoria conceitos muito próximos. A diferença é que a variância depende de uma única variável enquanto que a covariância mede a relação entre duas variáveis distintas. Dadas duas sequências em instantes de tempo diferentes (x_i e y_i), a covariância pode ser determinada pela equação

$$\sigma_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_t - \bar{x}_t).$$

As Tabelas 1, 2 e 3 exibem os cálculos da média, da variância e da covariância sobre a série temporal utilizada para gerar as estimativas, isto é, as previsões, antes e depois do co-registro. Estes algoritmos foram desenvolvidos na linguagem de programação Python. A partir de seus resultados, pode-se identificar se esta série é ou não estacionária.

Analisando a Tabela 1, pode-se observar que suas médias flutuam em torno do valor 128. Assim, considera-se que, no nível da média, a série é estacionária, uma vez que seus valores não variam ao longo da série.

A Tabela 2 mostra as variâncias calculadas, onde pode-se perceber que elas não são constantes, pois seus valores mudam ao longo da série. Com isso, já pode-se descartar a hipótese da série ser estacionária.

Tabela 1: Média.

	2M	5M	9M	13M	19M
Antes do Co-Registro	128,57	128,28	128,23	128,22	128,23
Depois do Co-Registro	128,63	128,35	128,38	128,55	128,30

Tabela 2: Variância.

	2M	5M	9M	13M	19M
Antes do Co-Registro	1759,34	2322,44	2366,81	2867,45	3461,01
Depois do Co-Registro	1750,61	2267,46	2572,37	2454,83	3016,35

Tabela 3: Covariância.

	2M e 5M	5M e 9M	9M e 13M	13M e 19M
Antes do Co-Registro	1189,82	1528,41	1615,76	2232,85
Depois do Co-Registro	1278,38	1449,49	1314,92	1759,99

A Tabela 3 mostra as covariâncias calculadas para a série. Mediante sua análise, pode-se perceber que elas apresentam valores bem distintos. Desta forma, pode-se concluir que esta é uma série não-estacionária.

Assumimos que esta não-estacionariedade da série está atribuída principalmente ao ruído *ground roll* presente nos dados. Desta forma, sua remoção, por exemplo, utilizando técnicas de filtragem, provavelmente implicaria na estacionariedade da série.

6 Resultados

A utilização de tecnologias de lapso de tempo (4D) para realizar a análise e o monitoramento de fluidos e pressão em áreas de produção de óleo e gás tem crescido consideravelmente, principalmente nas últimas duas décadas (LUMLEY, 2004). Este problema tem sido alvo de várias pesquisas e estudos de caso na literatura (van Gestel *et al.*, 2008), (LAFET *et al.*, 2009). Técnicas advindas de outras áreas podem ser úteis na solução deste problema, tais como o co-registro de imagens e a predição de séries temporais.

A partir dos dados reais, tem-se que a última sequência sísmica conhecida é a de 19 meses após a injeção do vapor. Dada a importância de se estimar o comportamento futuro do reservatório, torna-se necessário estimar uma sequência que ainda não se conhece, por exemplo, 23 meses após a injeção de vapor d'água. Vale ressaltar que para validar os algoritmos desenvolvidos, buscou-se extrapolar sequências sísmicas conhecidas. Portanto, procurou-se obter extrapolações para as sequências de 13 e 19 meses utilizando os métodos de regressão linear e quadrática, antes e depois do co-registro de imagens.

Assim, a Figura 24 corresponde aos processos realizados sobre a sequência sísmica de 13 meses antes do co-registro de imagens, na qual a Figura 24(a) exibe a imagem de 13 meses original, enquanto que na Figura 24(b) tem-se a sequência sísmica obtida pelo processo de extrapolação usando regressão linear e a Figura 24(c) representa a sequência sísmica obtida pelo processo de extrapolação utilizando regressão quadrática. Para melhor visualizar os atributos sísmicos, tem-se uma versão colorida destas imagens, exibida na Figura 25.

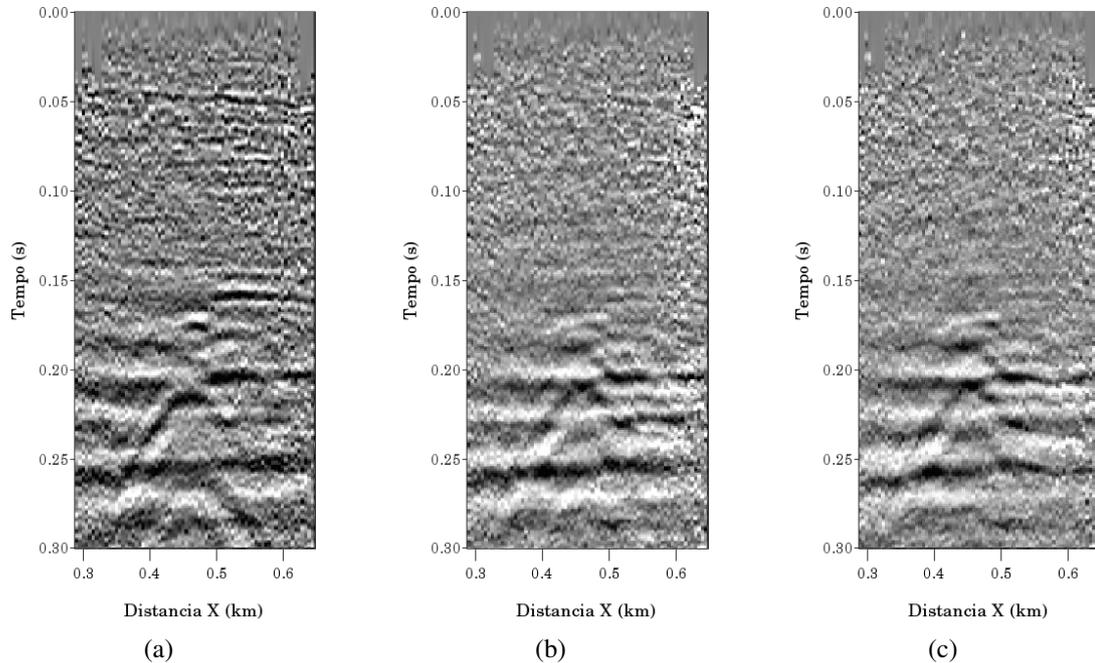


Figura 24: Sequências sísmicas de 13 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.

A Figura 26 corresponde aos processos efetuados sobre a sequência sísmica de 13 meses depois do co-registro de imagens, onde na Figura 26(a) tem-se a imagem de 13 meses original, na Figura 26(b) tem-se a sequência sísmica obtida por regressão linear e na Figura 26(c) a

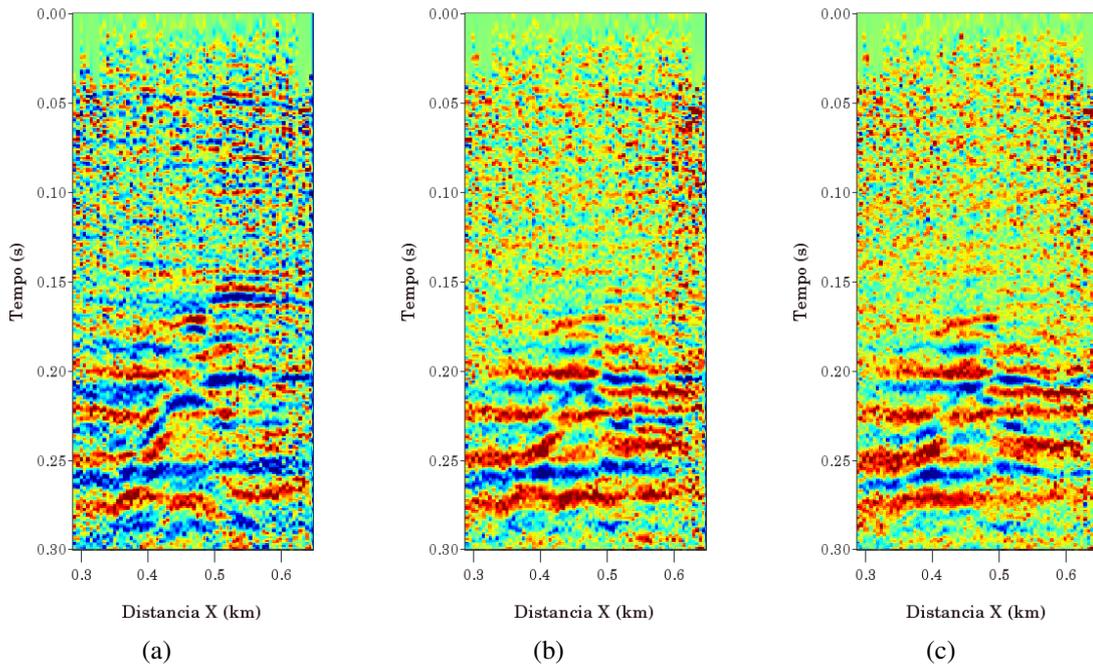


Figura 25: Sequências sísmicas coloridas correspondendo a 13 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.

sequência sísmica obtida por regressão quadrática. Já a Figura 27 mostra uma versão colorida destas imagens.

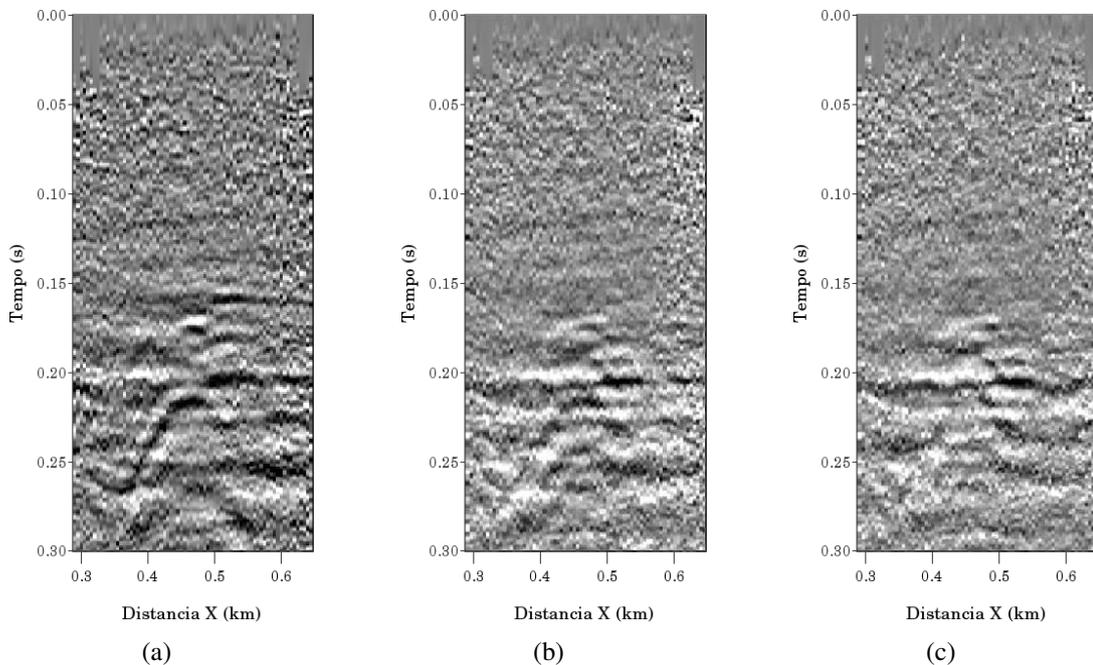


Figura 26: Sequências sísmicas de 13 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.

Na Figura 28 tem-se os processos realizados sobre a sequência sísmica de 19 meses antes do co-registro de imagens, na qual a Figura 28(a) exibe a imagem de 19 meses original, enquanto

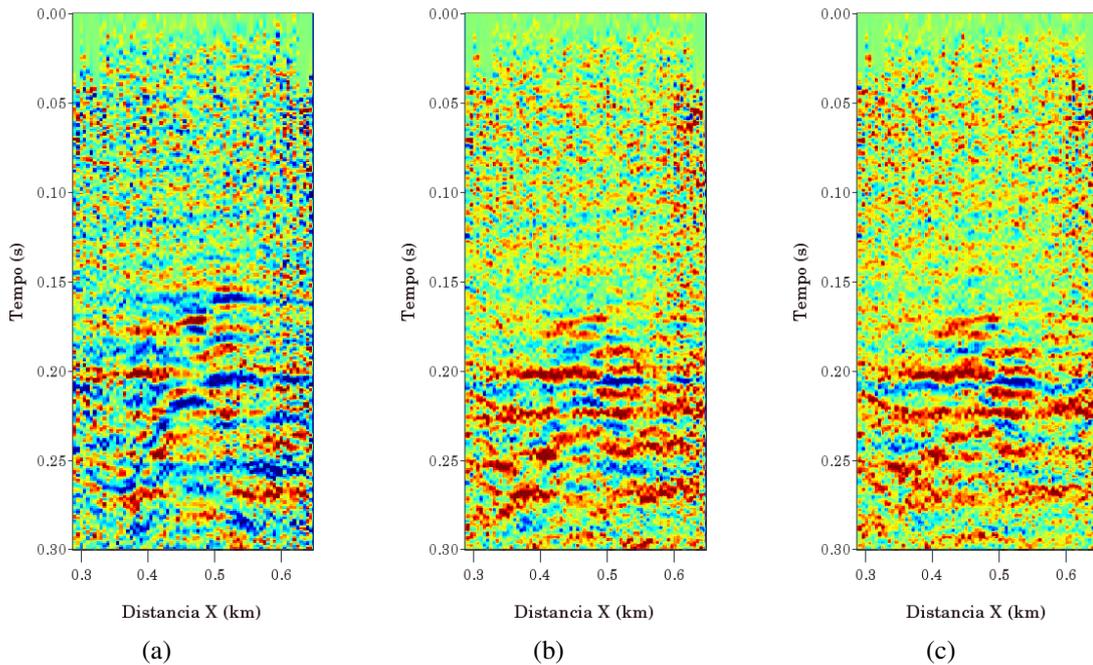


Figura 27: Sequências sísmicas coloridas correspondendo a 13 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.

que na Figura 28(b) tem-se a sequência sísmica obtida pelo processo de extrapolação regressão linear e a Figura 28(c) representa a sequência sísmica obtida pelo processo de extrapolação regressão quadrática. Da mesma forma, a Figura 29 mostra uma versão colorida destas imagens.

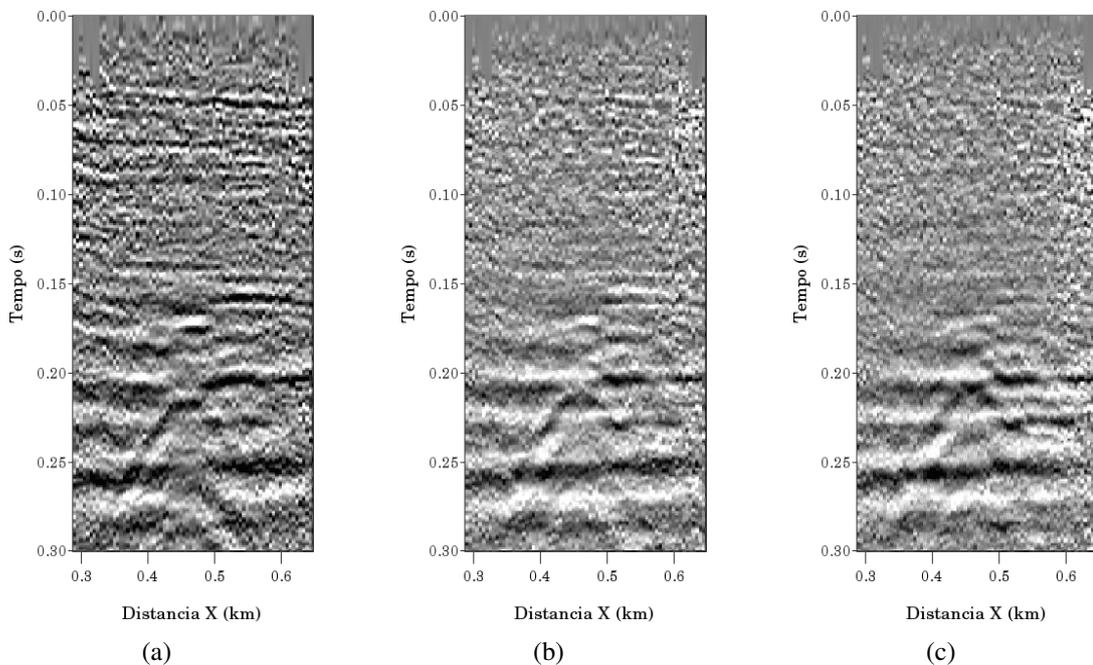


Figura 28: Sequências sísmicas de 19 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.

Já a Figura 30 corresponde aos processos efetuados sobre a sequência sísmica de 19 meses

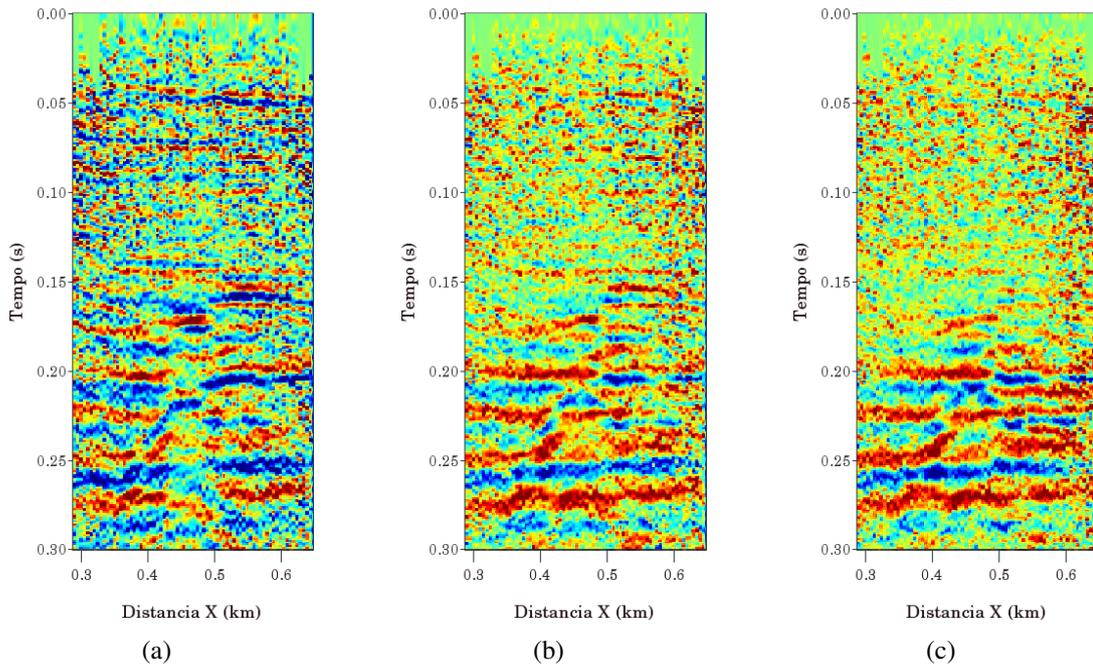


Figura 29: Sequências sísmicas coloridas correspondendo a 19 meses antes do co-registro de imagens (a) original, extrapolada por (b) regressão linear e por (c) regressão quadrática.

depois do co-registro de imagens, onde na Figura 30(a) tem-se a imagem de 19 meses original, na Figura 30(b) tem-se a sequência sísmica obtida pelo processo de extrapolação regressão linear e na Figura 30(c) a sequência sísmica obtida pelo processo de extrapolação regressão quadrática. Do mesmo modo, a Figura 31 mostra uma versão colorida destas imagens.

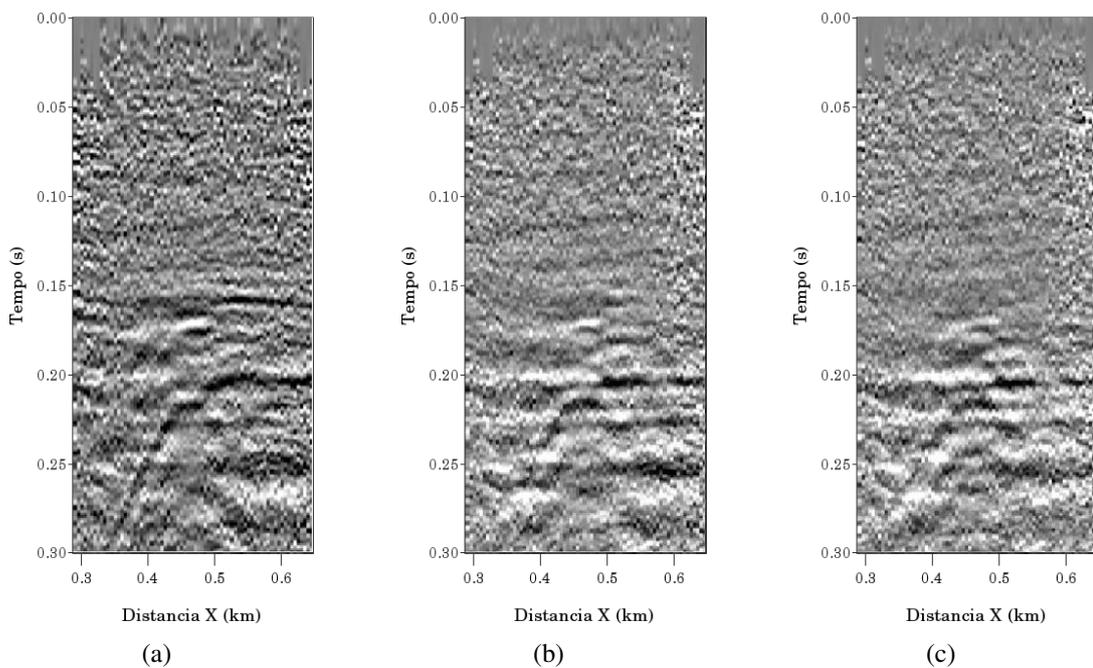


Figura 30: Sequências sísmicas de 19 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.

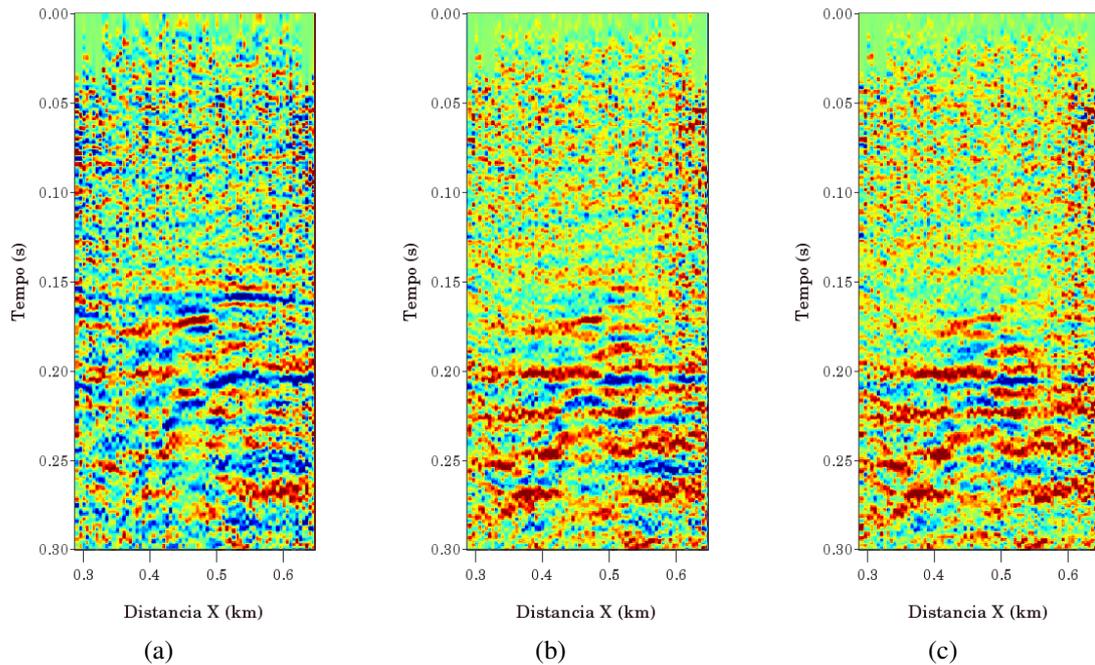


Figura 31: Sequências sísmicas coloridas correspondendo a 19 meses depois do co-registro de imagens (a) original, (b) extrapolada por regressão linear e (c) por regressão quadrática.

Uma vez definido um modelo numérico para as sequências conhecidas, pode-se cacular o nível de semelhanças entre elas e as sequências sísmicas reais. Para tando, foi utilizado os métodos descritos na Seção 5.3.4.

As Tabelas 4 e 6 exibem as saídas do algoritmo MAE após a execução do DistG, onde o índice da coluna representa o método de regressão utilizado, o índice da coluna representa as imagens da diferença em lapso de tempo usadas (antes e depois do co-registro de imagens) e o conteúdo da célula mostra o MAE obtido (máximo 100%).

Da mesma forma, as Tabelas 5 e 7 exibem as saídas do algoritmo RMSE após a execução do DistG, onde o índice da coluna representa o método de regressão utilizado, o índice da coluna representa as imagens da diferença em lapso de tempo usadas (antes e depois do co-registro de imagens) e o conteúdo da célula mostra o RMSE obtido (máximo 100%).

A Tabela 4 mostra o resultado do MAE analisando a imagem de 13 meses após a injeção do vapor d'água e suas versões extrapoladas obtidas por regressão linear e quadrática, antes e depois do co-registro de imagens.

Tabela 4: MAE (%) calculado a partir do registro de 13 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.

	Regressão Linear	Regressão Quadrática
Antes do Co-Registro	13,7	15,31
Depois do Co-Registro	13,07	14,72

Na Tabela 5 tem-se o resultado da análise usando o RSME realizado sobre a imagem de 13 meses após a injeção do vapor d'água e suas versões extrapoladas obtidas por regressão linear

e quadrática, antes e depois do co-registro de imagens.

Tabela 5: RMSE (%) calculado a partir do registro de 13 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.

	Regressão Linear	Regressão Quadrática
Antes do Co-Registro	18,28	20,36
Depois do Co-Registro	17,58	19,57

Já a Tabela 6 tem-se o resultado usando o MAE sobre a imagem de 19 meses após a injeção do vapor d'água e suas versões extrapoladas obtidas por regressão linear e quadrática, antes e depois do co-registro de imagens.

Tabela 6: MAE (%) calculado a partir do registro de 19 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.

	Regressão Linear	Regressão Quadrática
Antes do Co-Registro	15,13	17,69
Depois do Co-Registro	14,19	15,64

Por fim, a Tabela 7 exibe o resultado usando o RMSE sobre a imagem de 19 meses após a injeção do vapor d'água e suas versões extrapoladas obtidas por regressão linear e quadrática, também antes e depois do co-registro de imagens.

Tabela 7: RMSE (%) calculado a partir do registro de 19 meses, antes e depois do co-registro de imagens, e suas versões extrapoladas.

	Regressão Linear	Regressão Quadrática
Antes do Co-Registro	19,84	23,29
Depois do Co-Registro	18,71	20,43

Através dessas tabelas, é possível observar que o método de regressão quadrática obteve um resultado um tanto inferior ao método de regressão linear. Assim, fica evidente que a série apresenta uma tendência linear. Todavia, intuitivamente, é de conhecimento que o processo de ajuste polinomial da série poderia permitir obter resultados mais exatos na extrapolação do sinal AVO, em relação aos gerados pelo ajuste linear, pois utiliza três sequências de dados de entrada, quando comparados com apenas duas sequências de entrada do ajuste linear. Porém, os erros calculados para ambas as aproximações dos ajustes são quase similares.

Consideramos que tal fato pode ser explicado devido às sequências do sinal AVO terem sido registradas em intervalos de tempo relativamente próximos uns dos outros. Assim, tais sequências apresentam tendências no tempo relativamente idênticas uma das outras. Nesta hipótese, pode ser explicado então, por que um ajuste baseado em uma simples regressão linear adquire resultados um pouco melhores do que os obtidos por um ajuste polinomial (não-linear), na tentativa de prever a tendência no tempo do reservatório. Além disso, o fato da série ter média constante, também pode ter favorecido o modelo linear.

Mediante os resultados de MAE e RMSE, é possível observar que os resultados obtidos a partir dos dados co-registrados, mostraram-se superiores aos obtidos antes da aplicação co-registro de imagens. Isto pode ser explicado devido ao perfeito alinhamento entre cada pixel das sequências sísmicas em lapso de tempo, facilitando sua previsão no tempo.

Uma vez testados os métodos para extrapolar uma sequência conhecida e visto que os dados processados após o co-registro de imagens apresentaram resultados melhores, foram utilizados tais métodos para estimar o comportamento do reservatório 23 meses após a injeção do vapor, isto é, uma sequência futura e que ainda não se conhece. Também, como os erros gerados pela regressão linear e pela regressão quadrática foram aproximadamente similares, optou-se por utilizar estes dois métodos. Assim, a Figura 32(a) exhibe a extrapolação por regressão linear e a Figura 32(b) exhibe o resultado da extrapolação por regressão quadrática desta sequência. Já a Figura 33 mostra uma versão colorida destas imagens.

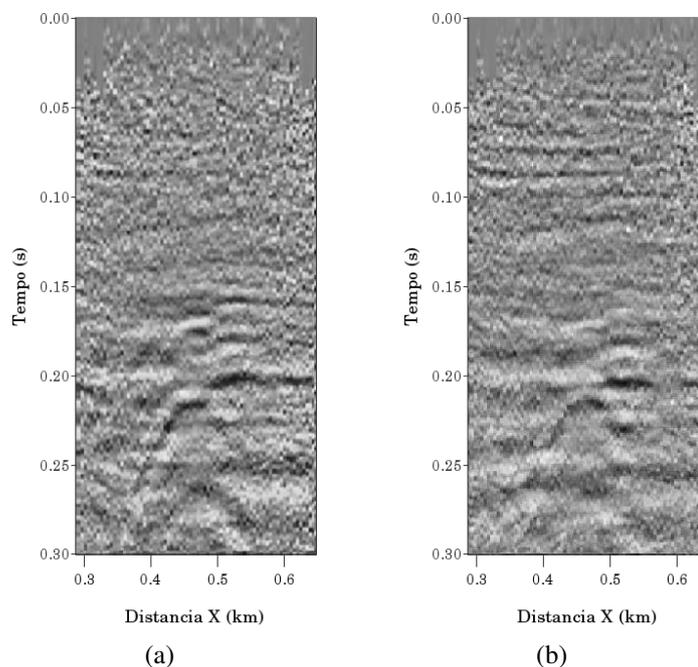


Figura 32: Sequências sísmicas correspondendo a 23 meses após a injeção do vapor extrapoladas por (a) regressão linear e (b) quadrática.

6.1 Resultados Com o WEKA

A Figura 34 corresponde ao processo visando estimar a sequência de 13 meses após a aplicação do co-registro de imagens, uma vez que os resultados após sua utilização mostraram-se melhores. Na Figura 34(a) tem-se a sequência de 13 meses original. Ela corresponde à imagem a qual deseja-se estimar. A Figura 34(b) representa a sua extrapolada pelo método de regressão linear desenvolvido e a Figura 34(c) pelo processo linear do WEKA. Já a Figura 34(d) tem-se a sequência obtida pelo processo gaussiano do WEKA e a Figura 34(e) por regressão quadrática.

As Tabelas 8 e 9 mostram o resultado após a execução do algoritmo DistG, ou seja, a análise entre a imagem de 13 meses após a injeção do vapor d'água e da aplicação do co-registro de

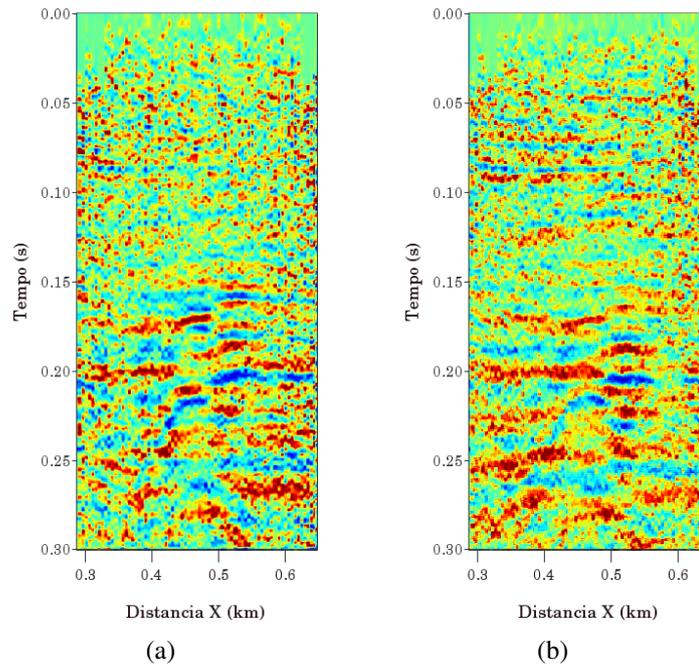


Figura 33: Sequências sísmicas coloridas correspondendo a 23 meses após a injeção do vapor extrapoladas por (a) regressão linear e (b) quadrática.

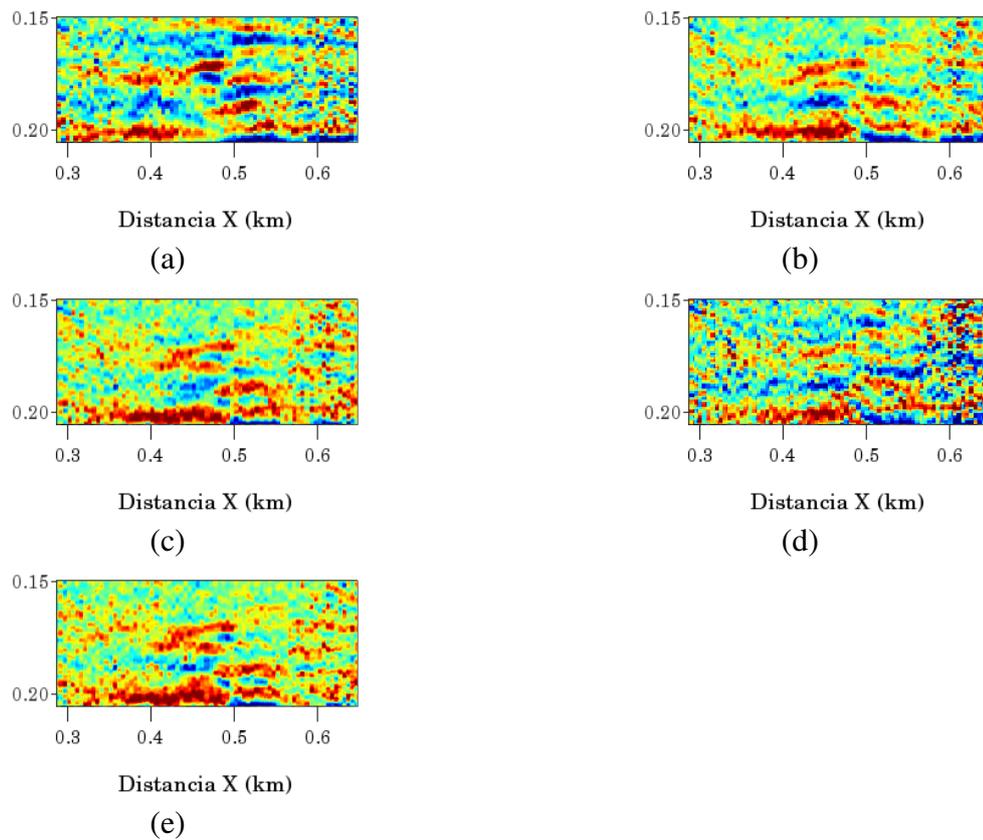


Figura 34: Sequências sísmicas de 13 meses coloridas (a) Original, extrapolada por (b) regressão linear (c) processo linear do WEKA, (d) processo gaussiano e (e) regressão quadrática.

imagens e suas extrapoladas. Na Tabela 8, tem-se os valores do MAE. Já o resultado do RMSE obtido após a aplicação do algoritmo DistG pode ser a Tabela 9.

Tabela 8: MAE (%) calculado a partir do registro de 13 meses depois do co-registro de imagens, e suas extrapoladas.

Regressão Linear	Processo Linear (WEKA)	Processo Gaussiano	Regressão Quadrática
14,45	13,16	16,45	16,36

Tabela 9: RMSE (%) calculado a partir do registro de 13 meses depois do co-registro de imagens, e suas extrapoladas.

Regressão Linear	Processo Linear (WEKA)	Processo Gaussiano	Regressão Quadrática
18,24	16,67	21,03	20,63

Embora os erros apresentados nas Tabelas 8 e 9 sejam similares, é possível perceber que os resultados obtidos a partir do processo linear implementado no WEKA obteve os melhores resultados. Porém, embora um pouco inferior, o algoritmo de regressão linear implementado obteve resultados próximos a este. Também, é possível observar que o algoritmo de regressão quadrática desenvolvido provou ter um desempenho similar ao processo gaussiano do software WEKA.

7 Conclusões

No que diz respeito à predição do comportamento futuro de um reservatório, esta é considerada uma questão em aberto e também uma área extremamente importante para a indústria petrolífera, pois ela almeja o aumento na recuperação final de hidrocarbonetos.

Neste trabalho foram descritos o processo de desenvolvimento e os resultados alcançados pelo uso de simples métodos de previsões de séries temporais, especificamente regressão linear e regressão quadrática, aplicados sobre mapas de diferenças obtidos a partir das sequências em lapso de tempo. Também, foram utilizados o processo linear e o gaussiano, ambos algoritmos de aprendizagem de máquina do Weka. Numa primeira etapa estes métodos foram aplicados sobre os mapas de diferenças antes do co-registro de imagens. Posteriormente, tais métodos também foram empregados sobre os dados resultantes desse processo de alinhamento.

Os resultados alcançados apresentados ao longo do texto demonstraram um desempenho considerado satisfatório para ambos os métodos, linear e quadrático. Isso pode ser constatado observando-se os erros exibidos nas Tabelas 4, 5, 6 e 7. A partir delas, pode-se notar que o método de ajuste linear descreve melhor o modelo de previsão feito sobre a série espaço-temporal do sinal AVO, uma vez que os erros do ajuste linear são menores para ambos os casos se comparado aos erros do ajuste polinomial.

Consideramos que tal fato pode ser explicado devido às sequências do sinal AVO terem sido registradas em intervalos, ou lapsos, de tempo relativamente próximos uns dos outros. Nesta hipótese, pode-se explicar então porque um ajuste baseado em uma simples regressão linear obteve melhores resultados do que os obtidos por um ajuste não-linear (polinomial de ordem 2), na tentativa de prever o comportamento no tempo do reservatório.

Também, observando-se as Tabelas 4, 5, 6 e 7, é notável que o co-registro de imagens provou ser um processo importante para as extrapolações em lapso de tempo, uma vez que os dados obtidos a partir das imagens das diferenças após o co-registro de imagens provaram ser melhores do que os obtidos antes do co-registro de imagens. Vale salientar também que os métodos de previsão usados neste trabalho podem ser aplicados a outros tipo de dados, desde que estejam em lapso de tempo, como por exemplo, imagens de satélite.

O uso do software WEKA, que utiliza algoritmos de aprendizagem de máquina, mostrou-se ser promissor para gerar modelos de predição inteligente. Comparando-se os algoritmos desenvolvidos com os implementados no WEKA, viu-se que eles apresentam resultados similares, conforme mostrado nas Tabelas 8 e 9. No entanto, os resultados obtidos a partir do processo linear implementado no WEKA, obteve os melhores resultados.

Comparando-se este algoritmo com o de regressão linear que foi implementado neste trabalho, pode-se perceber que os resultados são próximos. Todavia, o processo envolvendo os algoritmo do WEKA é mais dispendioso, pois é demorado, rotineiro e cansativo. Também, é possível observar que o algoritmo de predição usando regressão polinomial provou ter um desempenho similar ao processo gaussiano do WEKA.

A metodologia empregada para este trabalho apresenta algumas limitações quanto ao tratamento dos dados, mas que não invalidam os meios para alcançar seus objetivos. Isto refere-se à presença de ruídos nos dados que, provavelmente, comprometeu os resultados finais, aumentando assim, os erros gerados entre as sequência reais e as estimativas.

Como possíveis trabalhos futuros, pode-se apontar a implementação de outras funções de

extrapolações, tais como cúbica e hiperbólica. Além disso, métodos inteligentes podem ser desenvolvidos, estes, por sua vez, baseados em RNA. Um outro trabalho em andamento visa incorporar técnicas de reconstrução de imagem Bayesiana para previsão inteligente das sequências sísmicas (sinais AVO), obtidas a partir de reservatórios de hidrocarbonetos.

Referências

- ANDRADE, M. C. de; NETO, A. R.; ROSA, R. R.; SAWANT, H. S.; FERNANDES, F. C. R. Spatiotemporal patterns estimation using a multilayer perceptron neural network in a solar physics application. *Brazilian Society on Neural Networks*, Brazilian Society on Neural Networks, v. 2, n. 1, p. 14–21, 2004.
- AYOUB, F.; LEPRINCE, S.; AVOUAC, J. P. Co-registration and correlation of aerial photographs for ground deformation measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 64, p. 551–560, Novembro, 2009.
- BUENO, R. de Losso da S. *Econometria de Séries Temporais*. 1. ed. São Paulo: Cengage Learning, 2008.
- BUSTOS, H. I. A.; BANDEIRA, C. L. L.; SILVA, M. P. S.; RIBEIRO, E. C. Prediction of fluid in a petroleum reservoir using time series forecasting methods based on image co-registration. *International Conference Applied Computing*, Rio de Janeiro, Brasil, 2011.
- BUSTOS, H. I. A.; SILVA, M. P. S.; BANDEIRA, C. L. L.; RIBEIRO, E. C. Predição de fluidos em um reservatório petrolífero utilizando métodos de previsão de séries temporais. *V e-Science Workshop - XXXI Congresso da Sociedade Brasileira de Computação*, Natal, Rio Grande do Norte, Brasil, p. 967–974, 2011.
- COSTA, E. F. *Inferência de porosidade e saturação aplicada no monitoramento sísmico de reservatório*. Tese (Doutorado) — Universidade Estadual do Norte Fluminense (UENF), 2005.
- CUNHA, E. S. *Identificação de litofácies de poços de petróleo utilizando um método baseado em redes neurais artificiais*. Dissertação (Mestrado) — Universidade Federal de Campina Grande (UFCG), 2002.
- CURBELO, F. D. S. *Recuperação avançada de petróleo utilizando tensoativos*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte (UFRN), Departamento de Engenharia Química, 2006.
- DIX, C. H. Seismic velocities from surface measurements. *Geophysics*, v. 20, n. 1, p. 68–86, 1955.
- ENDERS, W. *Applied econometric time series*. John Wiley and Sons, 2004.
- FIGUEIREDO, A. M. *Mapeamento Automático de Horizontes e Falhas em Dados Sísmicos 3D baseado no algoritmo de Gás Neural Evolutivo*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2007.
- FOMEL, S. Local seismic attributes. *Geophysics*, v. 72, n. 3, p. A29–A33, 2007a.
- FOMEL, S. Shaping regularization in geophysical-estimation problems. *Geophysics*, v. 72, p. R29–R36, 2007b.
- FOMEL, S. *Revisiting SEP tour with RSF and SCons*. 2009. Disponível em: <http://www.reproducibility.org/wiki/Revisiting_SEP_tour_with_Madagascar_and_SCons>. Acesso em: 08 dez. 2009.

- FOMEL, S.; HENNENFENT, G. Reproducible computational experiments using scon. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, v. 4, p. 1257–1260, 2007.
- FOMEL, S.; JIN, L. Time-lapse image registration using the local similarity attribute. *77th Annual International Meeting, SEG, Expanded Abstracts*, p. 2979–2983, 2007.
- FONSECA, F. J. Métodos sísmicos: ondas, refração e reflexão sísmicas. LPGA, Curitiba - Paraná, 2002.
- GIRALDO, L. A.; CHEN, H.-Y.; TEUFEL, L. W. Spe 60297 field case study of geomechanical impact of pressure depletion in the low-permeability cupiagua gas-condensate reservoir. *Well Testing*, Society of Petroleum Engineers, 2000.
- HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, v. 5, n. 6, Novembro 1994.
- HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA data mining software: an update. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, n. 1, p. 10–18, nov. 2009. ISSN 1931-0145. Disponível em: <<http://dx.doi.org/10.1145/1656274.1656278>>.
- HOLDEN, M.; HILL, D. L. G.; DENTON, E. R. E.; JAROSZ, J. M.; COX, T. C. S.; ROHLFING, T.; GOODEY, J.; HAWKES, D. J. Voxel similarity measures for 3D serial MR brain image registration. *IEEE Transactions on Medical Imaging*, v. 19, p. 94–102, 2000.
- KIM, H. Y. *ProEikon - Rotinas e Programas em C++ para Processamento de Imagens e Visão Computacional*. 2010. Disponível em: <<http://www.lps.usp.br/hae/software>>. Acesso em: 23 nov. 2010.
- KISUKA, A. L.; ANDRADE, T. P. da S. Efeito dos ruídos na análise de avo. *Trabalho de Conclusão de Curso, Universidade Agostinho Neto*, 2006.
- LAFET, Y.; ROURE, B.; DOYEN, P. M.; BURAN, H.; TAYLOR, R. Time-lapse fluid monitoring with 4d global inversion. *PESA News Resources*, p. 48–51, Junho/Julho, 2009.
- LEITE, F. E. A. *Análise estatística de padrões sísmicos: decomposição em multiescala*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte (UFRN), Departamento de Física Teórica e Experimental, 2007.
- LINDEN, R. *Algoritmos Genéticos*. 1. ed. Rio de Janeiro: Editora Brasport, 2006.
- LUIZ, J. F.; SILVA, L. M. C. Geofísica de prospecção. Editora CEJUP Ltda., Belém - PA, p. 311, 1995.
- LUMLEY, D. E. *Seismic Time-Lapse Monitoring of Subsurface Fluid Flow*. Tese (Doutorado) — Stanford University, 1995a.
- LUMLEY, D. E. 4D seismic monitoring of an active steamflood. *65th Ann. Internat. Mtg, Soc. of Expl. Geophys.*, p. 203–206, 1995b.
- LUMLEY, D. E. Business and technology challenges for 4D seismic reservoir monitoring. *The Leading Edge*, v. 23, n. 11, p. 1166–1168, 2004.

- LUMLEY, D. E. 4D seismic monitoring of CO₂ sequestration. *The Leading Edge*, v. 29, n. 2, p. 150–155, 2010.
- LYONS, P. K. Preliminary outline for power plant educational module. Disponível em: <http://www.iit.edu/research/services/doc_ret_2004/I_ve_got_the_Power.doc>. Acesso em 12 de Dezembro de 2010. 2010.
- MACKAY, D. J. C. Introduction to gaussian processes. *C. M. Bishop, editor, Neural Networks and Machine Learning*, Springer, v. 168 of NATO ASI Series, p. 133–165, 2009.
- MADAGASCAR. 2011. Disponível em: <<http://www.reproducibility.org>>. Acesso em: 12 jun. 2011.
- MEDEIROS, S. C. D. *Inversão de Parâmetros em Dados Sísmicos por Algoritmos Genéticos*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2005.
- MEDIAWIKI. 2010. Disponível em: <<http://www.mediawiki.org>>. Acesso em: 03 jan. 2011.
- MORETIN, P. A.; TOLOI, M. C. *Análises de Séries Temporais*. 1. ed. São Paulo: Edgard Blücher, 2004.
- NETO, A. R. Kalman filtering stochastic optimal estimation algorithm and usual backpropagation in neural nets training. *Proc. of Second Brazilian Cong. in Neural Net.*, Curitiba, Paraná, Brasil, p. 139–144, 1995.
- NUR, A. Four-dimensional seismology and (true) direct detection of hydrocarbons: the petrophysical basis. *The Leading Edge*, v. 8, n. 9, p. 30–36, 1989.
- OSTRANDER, W. J. Plane-wave reflection coefficients for gas sand at non normal angles of incidence. *Geophysics*, v. 49, p. 1637–1648, 1984.
- QUEIROLO, C. C.; BELLON, O. R. P.; SILVA, L. Reconhecimento de faces a partir do registro de imagens 3d. *REIC*, v. 7, p. 1, 2007.
- RISSO, V. F.; SCHIOZER, D. J. Utilização de sísmica 4D e de mapas de saturação no ajuste de histórico de reservatórios petrolíferos. *Revista Brasileira de Geociências*, v. 38, n. 1, p. 172–187, 2008.
- ROBINSON, E. A.; TREITEL, S. *Geophysical Signal Analysis*. Englewood Cliffs: Prentice-Hall, 1980.
- RODRIGUES, E. P. *Avaliação de métricas para o corregristo não rígido de imagens médicas*. Tese (Doutorado) — Universidade de São Paulo (USP), 2010.
- SANCEVERO, S. S.; REMACRE, A. Z.; PORTUGAL, R. de S. O papel da inversão para a impedância acústica no processo de caracterização sísmica de reservatórios. *Revista Brasileira de Geofísica*, v. 24, p. 495–512, 12 2006. ISSN 0102-261X.
- SEEGER, M. Gaussian processes for machine learning. *International Journal of Neural Systems*, v. 14, p. 2004, 2004.
- SHERIFF, R. E.; GELDART, L. P. *Exploration Seismology*. 2. ed. Welwyn, United Kingdom: Cambridge University Press, 1995.

TANG, L.; HAMARNEH, G.; CELLER, A. Co-registration and fusion of ct and spect images for attenuation correction and localization. *IEEE Symposium on Signal Processing and Information Technology (ISSPIT)*, 2006.

THOMAS, J. E. *Fundamentos da Engenharia de Petróleo*. 1. ed. Rio de Janeiro: Editora Interciência, 2001.

van den Elsen, P. A.; PAUL, E. J. D.; VIERGEVER, M. A. Medical image matching – a review with classification. *IEEE Engineering in Medicine and Biology*, v. 12(1), p. 26–39, Março 1993.

van Gestel, J. P.; KOMMEDAL, J. H.; BARKVED, O. I.; MUNDAL, I.; R. Bakke, K. D. B. Continuous seismic surveillance of Valhall Field. *The Leading Edge*, v. 27, p. 1616–1621, 2008.

WIENER, N. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. 1964. ISBN 0262730057.

WILLIAMS, C.; RASMUSSEN, C. *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

ZACHARIAS, S.; HEATWOLE, C.; CAKLEY, C. Robust quantitative techniques for validating pesticide transport models. *Transactions of the American Society of Agricultural Engineers*, v. 39, p. 47–54, 1996.