

Tackling the weathering with low ranks

Handling the complex near surface of land seismic data with low-rank-based methods

Alfaraj, Ali

DOI

[10.4233/uuid:2c4f8ca0-f44a-485a-99d6-d7952c902fa2](https://doi.org/10.4233/uuid:2c4f8ca0-f44a-485a-99d6-d7952c902fa2)

Publication date

2024

Document Version

Final published version

Citation (APA)

Alfaraj, A. (2024). *Tackling the weathering with low ranks: Handling the complex near surface of land seismic data with low-rank-based methods*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:2c4f8ca0-f44a-485a-99d6-d7952c902fa2>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

The background of the cover is a dark blue gradient with numerous orange and yellow splatters and streaks, creating a dynamic, textured effect. The text is centered and rendered in a bold, serif font.

TACKLING THE WEATHERING WITH LOW RANKS

Handling the complex near surface of land seismic
data with low-rank-based methods

ALI M. ALFARAJ

TACKLING THE WEATHERING WITH LOW RANKS
HANDLING THE COMPLEX NEAR SURFACE OF LAND SEISMIC DATA WITH
LOW-RANK-BASED METHODS

TACKLING THE WEATHERING WITH LOW RANKS
HANDLING THE COMPLEX NEAR SURFACE OF LAND SEISMIC DATA WITH
LOW-RANK-BASED METHODS

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen
chair of the Board for Doctorates
to be defended publicly on
Wednesday, 6 March 2024 at 15:00 o'clock

by

Ali Mohammed A. ALFARAJ

Master of Science in Applied Geophysics
IDEA League Joint Master Program
TU Delft, The Netherlands
ETH Zurich, Switzerland
RWTH Aachen, Germany
born in Dammam, Saudi Arabia

This dissertation has been approved by the promoters.

Composition of the doctoral committee:

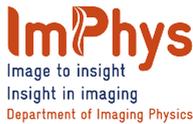
Rector Magnificus,
Dr.ir. D. J. Verschuur,
Prof.dr.ir. F. J. Herrmann,

chairperson
Delft University of Technology, promotor
Georgia Institute of Technology, promotor

Independent members:

Prof.dr. B. Rieger
Prof.dr.ir. M.B. van Gijzen
Prof.dr. M. van der Baan
Dr.ir. D.S. Draganov
Dr. C.V. Kostov

Delft University of Technology
Delft University of Technology
University of Alberta, Canada
Delft University of Technology
Consultant, France



Delphi Consortium

The work in this dissertation was conducted at the department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology within the Delphi Consortium, and was financially supported by Saudi Aramco.

Printed by: Proefschriftspecialist

Copyright © 2023 by A. M. A. Alfaraj

ISBN 978-94-93330-65-8

An electronic version of this dissertation is available at

<http://repository.tudelft.nl/>.

*To my grandma, who I dearly miss.
To Leen & Yasmeen, the joy and happiness of my life.*

CONTENTS

Summary	xi
1 Introduction	1
1.1 Relevance to society.	1
1.2 Beginner's guide to seismic waves.	2
1.3 The weathering effect.	3
1.4 The acquisition requirement	5
1.5 Undoing the weathering effects	6
1.5.1 The Long-wavelength component	8
1.5.2 The short-wavelength component	10
1.6 A summary of existing challenges.	13
1.7 Research questions	14
1.8 Objectives and Contributions	14
1.9 Dissertation Outline	17
References	17
2 2D model-independent near-surface estimation and correction	23
2.1 Introduction	25
2.1.1 Contributions	26
2.1.2 Outline.	27
2.2 Rank-based processing principles.	27
2.2.1 Simulated data.	27
2.2.2 Low-rank structure.	28
2.2.3 Structure destruction	31
2.2.4 Structure promotion	32
2.3 LR-ReS estimation and correction.	33
2.4 Results	37
2.4.1 Synthetic data example	38
2.4.2 Field data example.	46
2.5 Discussion	51
2.5.1 Practical aspects	52
2.5.2 Mitigation of the noise effect.	53
2.5.3 Further extensions and applications	56
2.6 Conclusions.	57
2.7 Acknowledgments	58
References	58

3	Near-surface pre-processing for RTM and LS-RTM	63
3.1	Introduction	64
3.2	RTM and LS-RTM	64
3.3	LR-ReS estimation and correction.	65
3.4	Results	65
3.5	Conclusions.	69
3.6	Acknowledgements	69
	References	70
4	Reconstruction of subsampled data	73
4.1	Introduction	75
4.1.1	Contributions	76
4.1.2	Outline.	76
4.2	The singular values decay.	76
4.3	Methodology	77
4.3.1	LR-based near surface correction	77
4.3.2	Rank-minimization-based interpolation.	79
4.3.3	Joint near-surface correction and interpolation	81
4.4	Results	84
4.4.1	Synthetic data	84
4.4.2	Field data	85
4.5	Discussion	91
4.5.1	Data windowing	93
4.5.2	Rank selection	94
4.5.3	Computational efficiency	95
4.5.4	Limitations and extensions	96
4.6	Conclusions.	97
4.7	Acknowledgments	97
	References	97
5	3D model-independent near-surface estimation and correction	101
5.1	Introduction	103
5.1.1	Contributions	103
5.2	Matricization domain.	104
5.2.1	Simulated data.	105
5.3	Methodology	109
5.4	Results	111
5.4.1	Synthetic data example	112
5.4.2	Field data example.	113
5.5	Discussion	113
5.5.1	Advantages.	118
5.5.2	Sensitivity to rank selection	120
5.5.3	Limitations.	120
5.6	Conclusions.	121
5.7	Acknowledgments	121
	References	121

6	Near- and sub-surface model estimation with joint migration inversion	125
6.1	Introduction	126
6.2	Methodology	127
6.2.1	Rank selection	128
6.3	Results and discussion	128
6.4	Conclusions.	129
6.5	Acknowledgements	130
	References	130
7	Conclusions and Recommendations	135
7.1	Conclusions.	136
7.1.1	Bypassing the weathering	136
7.1.2	Data reconstruction	137
7.1.3	Near- and sub-surface model estimation	137
7.2	Recommendations	138
7.2.1	3D reconstruction	138
7.2.2	Adding back the near-surface effects.	138
7.2.3	Near- and sub-surface model estimation	138
7.2.4	A target in the near-surface	138
7.2.5	Accurate modeling.	139
7.2.6	Changing the workflow	139
7.2.7	Marine data	139
	References	140
A	Full waveform inversion of one-way wavefields at the ocean bottom	141
A.1	Introduction	142
A.2	Wavefield decomposition	143
A.3	Waveform inversion using one-way wavefields	143
A.4	Numerical results and discussion	144
A.5	Conclusions.	147
A.6	Acknowledgements	148
	References	148
	Acknowledgements	151
	Curriculum Vitæ	159
	List of Publications	161

SUMMARY

Imaging and inversion with seismic data recorded with sources and receivers at the surface are powerful tools to infer knowledge about the subsurface. However, creating an image with seismic data is unfortunately not as easy as taking a picture with a smartphone. The estimated subsurface models in many situations are far from ideal due to the low quality nature of the data. One of the reasons can be weathering of the near-surface geology that generates unconsolidated material characterized by slow velocity with rapidly varying, heterogeneous and season-dependent nature. Acquiring seismic data on such near-surface leads to complex wave propagation, posing challenges to imaging and inversion. In this dissertation, we tackle the weathering effects during seismic data processing, imaging and inversion with low-rank-based methods.

One approach to tackle the weathering effects on seismic data is removing them during seismic data processing. To do so for 2D data, we propose a model-independent low-rank-based near-surface estimation and correction in the midpoint-offset-frequency domain. In this domain, ideal data exhibit low rank structures, which get destroyed due to the influence of the weathering layers. Accordingly, the method makes use of the redundant nature of seismic data that allows for accurate approximation by low-rank matrices. To estimate the time shifts that compensate for the weathering effects, we cross-correlate a data set influenced by the near-surface weathering layers with its low-rank approximated version. Since we estimate time shifts (commonly referred to as statics) and no longer the directly low-rank approximated data, we avoid losses of the amplitude information. To improve the estimated statics and to alleviate the need for accurate rank selection for low-rank approximation, we implement the method in an iterative and multi-scale fashion. Since the low-rank approximation deteriorates at high frequencies, we utilize its better performance at low frequencies and exploit the common statics amongst different frequency bands. Using synthetic and field data, we demonstrate the performance of the proposed method, which requires no knowledge of the subsurface model, demands minimal data pre-processing, and provides accurate solutions with high computational efficiency compared to existing techniques.

When seismic data acquired on complex near-surface are additionally subsampled for economical reasons, such as monitoring of sequestered carbon dioxide and hydrogen, the problem is further exacerbated. Both the weathering layers and randomized subsampling render coherent energy incoherent. Therefore, they both contribute to destruction of the low-rank structure commonly associated with statics-free densely-sampled data. Frugal data acquisition in complex near-surface regimes makes separation of the distinct sampling and weathering effects on the rank structure difficult, which as a result lead to poor reconstruction. To overcome that, we propose to reconstruct the data with joint rank-reduction-based near-surface correction and interpolation. The method simultaneously accounts for the weathering and subsampling effects to provide accurate reconstruction. Since low-rank approximation is used for near-surface correction,

we also utilize it in rank-minimization interpolation as a cost-free initial solution to the optimization problem. As both near-surface correction and interpolation operate in the midpoint-offset domain, we avoid the cost of transformations back and forth from the source-receiver to midpoint-offset transform domain. Consequently, the proposed reconstruction, which shows its potential on synthetic and field data, additionally increases the computational efficiency.

While the aforementioned near-surface correction deals with 2D data, the Earth is a 3D object that requires acquisition of 5D data for proper subsurface model estimation. For 5D data, the limitations and challenges of conventional near-surface correction methods are magnified. To avoid them, we propose a 5D model-independent low-rank-based near-surface correction. To compute the singular value decomposition of 5D data volumes with 1 temporal and 4 spatial dimensions, which is necessary for low-rank approximation, we need to perform matricization of the 5D data, i.e. organization of the 5D data into matrices. At the same time, it is essential that the chosen organization domain reveals the underlying low-rank structure. Therefore, we first analyze different matricization domains that can be used to organize the 5D data. Similar to the 2D case, we show that — in the potential domain — the near-surface weathering layers render coherent energy incoherent, which results in slowly decaying singular values compared to the statics-free data that are of low-rank nature. The proposed method, which we show on synthetic and field data, enjoys the same benefits of the proposed method for 2D data, in addition to being able to capture the 3D nature of the Earth.

Due to the complex nature of the near-surface and due to its impact on the subsurface model, the near-surface model gets treated separately from the subsurface model. However, the optimal goal is not to remove the near-surface effects with data processing, but to accurately estimate near- and sub-surface models simultaneously. To do so, we use the inherent scale separation of joint migration inversion that estimates a low-wavenumber velocity and high-wavenumber reflectivity. Since rapid variations in surface elevation and near-surface model result in high wavenumber effects, they end up affecting the reflectivity model. At the same time, the estimated reflectivity influences velocity estimation. Consequently, JMI provides erroneous subsurface models in the presence of complex weathering layers. To mitigate that, we use multi-scale low-rank updates in the reflectivity domain. The proposed method reduces the near-surface effects at the initial iterations, but it allows more details of the near-surface model to enter the solution at later iterations. In the end, we estimate accurate near- and sub-surface models simultaneously without the need to bypass the weathering layers.

1

INTRODUCTION

1.1. RELEVANCE TO SOCIETY

Continuous access to energy has become an essential element of modern life. It is necessary to be able to get access to drinking water, healthcare, sanitation, heating, cooling, lighting, housing, manufacturing, telecommunication and transportation. Without them, life can be challenging and development may come to an end as the focus shifts to surviving. According to the United Nations (UN), 733 million people had no electricity in 2020 [1], which means that they had to bear the hardship of a difficult life. To avoid that, it is important that the world unites and collaborates to provide energy to those who need it.

As humanity continues to grow and develop, the demand for energy is also projected to increase [2]. While the existing affordable, reliable and sustainable energy resources are unable to meet the current energy demand, it is essential to continue to explore for more energy resources that can meet current and future demands. In addition to fossil fuels that we heavily rely on at the moment, alternative resources of energy such as geothermal, hydro, wind and solar energy are expected to play vital roles in reaching energy security [2]. To avoid periods of low energy availability, energy storage, e.g. in a form of natural gas or hydrogen in the subsurface, is necessary. At the same time, there is an urgent call to curb global warming to 1.5° compared to pre-industrial levels, where reduction of the carbon footprint is important [3]. To do so, we need to replace energy resources that produce high amounts of carbon with low carbon or carbon-free emitters. However, this transition can not take place now since the low carbon emission resources are not yet available at a scale to cover the current demand. An alternative way to reduce the carbon footprint is to capture the CO₂ that is usually emitted in the atmosphere, utilize it, and store it in the subsurface. If successful, we can continue to responsibly use existing resources to generate energy and gradually replace heavy carbon emitters with those that have low or even zero carbon footprint.

The subsurface, both shallow and deep can be a common denominator for energy security and meeting global warming targets. It is where geothermal energy and fossil fuels are trapped in reservoirs. By extracting them, we can produce different forms of

energy. The subsurface can also host wind turbines to produce energy from wind. It is also where we can store energy in a form of natural gas or hydrogen in depleted reservoirs or salt caverns. These reservoirs can also be used to store the captured carbon. Therefore, understanding the subsurface is essential to find energy resources and storage reservoirs to supply the world with the means to meet modern life demands and reduce emissions. It is also important to monitor the reservoirs to find out the status of the resources and stored emissions to prevent a leak out of the reservoir. Moreover, engineers need to measure certain properties of the subsurface for construction and drilling purposes, e.g. placement of wind farms. With advanced technology, including seismic imaging technology, we can achieve the aforementioned objectives.

1.2. BEGINNER'S GUIDE TO SEISMIC WAVES

When we speak, air is pushed between the vocal cords, which vibrate to produce energy in a form of sound waves. The sound waves propagate to carry information to the recipient during verbal communication. When they encounter a wall, the waves reflect back to create an echo. By measuring the time it takes for the sound to travel from the source to the wall and back to the speaker, we can calculate the distance the waves travel, given the speed of sound in the propagation medium. Based on similar principles, seismic waves, a type of sound waves, can be used to create a model of the subsurface and identify its properties, e.g. an image of the subsurface. This allows us to get information about the subsurface remotely without the need to install sensors inside the medium. The process can be thought of in a similar way to using ultrasound waves to image a fetus during pregnancy or to inspect body organs [4].

Seismic waves exist naturally in the Earth due to its continuous movement. They are magnified during earthquakes or volcanic eruptions, which can be measured to infer information about the Earth's activities (passive seismic). Passive seismic data can also be used to image the Earth's interior, e.g. crust, mantle and core, at a global scale [5]. Since only the low frequency component of existing seismic waves reach the recording stations at the surface, the estimated global models with the recorded wavefields have limited resolution. For detailed model estimation such as imaging a shallow and small part of the Earth's crust, e.g. where energy resources and storage reservoirs exist, seismic waves with higher frequencies can be created. With higher frequencies, higher resolution models can be created, which come with a price compared to the freely existing seismic waves used in the field of global seismology.

To acquire seismic data, sources at the Earth's surface are used to send seismic waves. One difference between seismic and ultrasound waves is that seismic sources emit waves with much lower frequency content compared to ultrasound sources. For typical ultrasound imaging, the used frequencies range between 1 and 15 MHz to result in millimetres model resolution [6]. On the other hand, the frequencies used in active seismic imaging are typically between 1 and 120 Hz to provide metres resolution. For detailed near-surface imaging, higher frequencies in the range of 1000 Hz are used to provide sub-metre resolution. While higher frequencies provide higher resolution, their propagation to deeper targets, i.e. penetration depth, is limited due to energy decay of the wavefields caused by attenuation at the shallow layers. That is why passive seismic data below 1 Hz are used to image the Earth's interior to reach 100s or 1000s of kilometres deep.

Onshore, vibrating trucks are used to generate seismic waves (recall the vocal cords) that are usually recorded by many active receivers called geophones (geo-phones) distributed spatially at the surface (Figure 1.1). The emitted waves, usually a sweep from low to high frequencies, propagate through the subsurface (Figure 1.1(b)). Since the subsurface is composed of rocks with different characteristics such as different density and wave propagation velocity, the seismic wavefields reflect (Figure 1.1(c), like the echo), refract and transmit (Figure 1.1(b)) at interfaces where the rock properties change. When the wavefields reach the surface, they get recorded by the geophones (Figure 1.1(d)). The recorded data contain the travel time it takes for the waves to travel from the source to the receivers. Each trace correspond to a source-receiver pair, while a gather (Figure 1.1(d)) is a collection of traces plotted next to each others. For a medium with horizontal layers, these reflection curves are well-approximated by hyperbolas [7]. The apexes of the reflections correspond to the normal incidence, where the wavefields travel along the shortest paths. The longest traveled paths correspond to the farthest source-receiver distances. Therefore, they arrive at later times (Figure 1.1(d)). The recorded data also contain amplitude information that indicates the intensity of the wavefields. The higher the contrast at the interface, the stronger the amplitude. Using the recorded data at the surface, we can create images and models of the subsurface [8–10]. These models are essential for finding energy resources, monitoring reservoirs, construction and engineering projects. Therefore, they need to be of high accuracy as critical and costly decisions are made based on them. However, that is challenging in the presence of complex weathering layers near the surface.

1.3. THE WEATHERING EFFECT

Creating an image with seismic data is unfortunately not as easy as taking a picture with a smartphone. It is also not as easy as using the simulated data displayed in Figure 1.1(d) to estimate the model shown in Figure 1.1(a). The estimated subsurface models in many situations are far from ideal due to the nature of the data that can be of low quality. This can lead to difficulty and inaccuracy when making critical decisions. Such situation is common for land seismic data when acquired on top of unconsolidated weathering layers. In contrast, offshore seismic data are acquired by sources and receivers placed near the water surface. Due to the presence of the water layer, the data are not influenced by the weathering layers that affect land data. However, if the overburden velocity model, i.e. water velocity, of marine data is not accurate, full waveform inversion (FWI) produces erroneous subsurface models as demonstrated in the appendix (Appendix A).

By the weathering layers, we refer to the loosely compacted material near the Earth's surface that can be heterogeneous, rapidly changing, season-dependent and composed of low-velocity material between 250 and 1000 m/s [11]. These materials are naturally created, transported and deposited by the weathering process. In comparison, the deeper subsurface is characterized by consolidated and compacted rocks as they have been exposed to higher pressure by the rocks above them. Even though the deeper targets are far from the near-surface, their reflection responses in the seismic data are still influenced by the weathering layers if not well-resolved.

Figure 1.2 shows a real data example composed of five gathers. The data is much more complicated compared to the simulated data (Figure 1.1(d)). Due to the imprint of the the

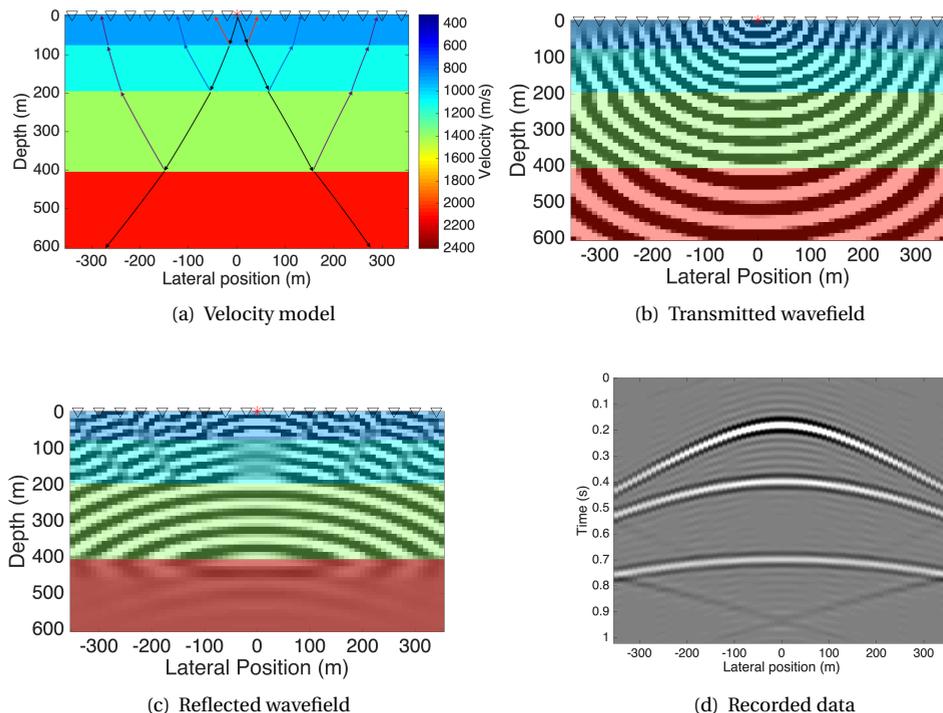


Figure 1.1: Illustration of seismic data acquisition with data simulation. (a) Simple model depicting the subsurface composed of layered rocks with different sound propagation velocities. Triangles at the surface indicate receivers and the asterisk indicates the source position. Seismic waves emitted by the source propagate to the subsurface downwards (black arrows). When the medium properties change, the velocity in this case, the wavefields reflect as indicated by the red, blue and magenta arrows and transmit with different propagation angles. The (b) transmitted and (c) reflected wavefields in the subsurface for one frequency (26 Hz) overlaid over the velocity model. Note the changes at the interfaces. (d) The recorded reflected wavefield by the spatially distributed receivers at the surface over time. The gather shows three primary reflections (hyperbolas) corresponding to the three interfaces that generate the three reflections. For the sake of simplicity, multiple reflections are not modelled in this figure.

near-surface weathering layers, the reflections do not resemble hyperbolas. They are also noisy and non-continuous. Therefore, to image the deeper targets, accurate knowledge of the weathering layers is essential. Otherwise, we may estimate erroneous subsurface models that may lead to poor decisions. [12] illustrate that if the weathering layers are not resolved, the resultant images become distorted, as demonstrated by Figures 1.3(a) and 1.3(b), which show the distortion due to sand dunes and karsts in the near-surface.

To observe the effect of a complex near-surface on the modelled data displayed in Figure 1.1, we modify the near-surface model from a constant velocity to a complex one (Fig-

ure 1.4(b)). It resembles a real near-surface that exhibits a rapidly varying nature (Figure 1.4(a)). Due to that, the wavefields in the subsurface become asymmetric (Figures 1.4(c) and 1.4(d)) compared to those in Figures 1.1(b) and 1.1(c). The energy transmitted and received also becomes lower. As a result, the recorded data at the surface 1.4(e) are no longer hyperbolic with less continuity and varying amplitudes. To estimate a subsurface model, we can iteratively solve an inverse problem that matches the recorded data with modelled data. Joint migration inversion (JMI) is such method that estimates speed of sound profiles (velocity) and reflectivity (analogous to the image) of the subsurface [13, 14]. It uses data modelling to create synthetic data, which relies on speed of sound profiles of the near- and sub-surface (initial models). The modelled data are then matched with the observed data, where the residual is used to update the subsurface models. Since the starting models do not resemble the near-surface features found in the true model – as they need to be estimated–, the estimated subsurface velocity and reflectivity become erroneous (Figure 1.5). Even though the subsurface is simple, the complexity of the near-surface leads the inversion to a local minimum, which can result in errors when making critical decisions. However, the estimation of accurate near-surface weathering layers is usually challenging.

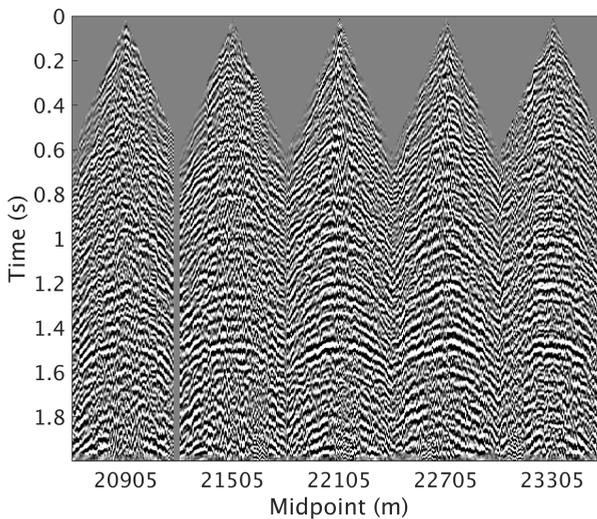


Figure 1.2: Field data example showing the weathering layers effect on subsurface reflections.

1.4. THE ACQUISITION REQUIREMENT

In the previous section, we demonstrate the effect of the weathering layers. Even though the target lies in the subsurface beneath the weathered material, accurate estimation of the near-surface model is essential as it influences the subsurface model. However, the near-surface can be heterogeneous, rapidly changing in space and season-dependent.

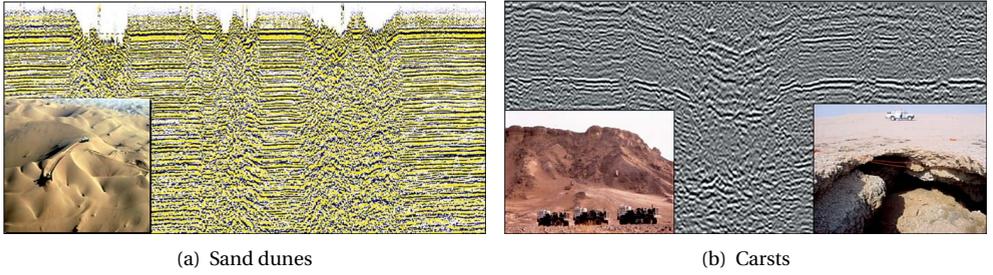


Figure 1.3: (a) The effect of sand dunes and (b) karsts on the estimated images from [12].

The heterogeneity and rapid variations are evident from Figure 1.4(a), which require sub-metre resolution. To capture the near-surface details with high resolution near-surface models, dense data sampling becomes essential. Moreover, the near-surface is composed of low velocity material usually between 250 and 1000 m/s. According to the Nyquist-Shannon sampling criterion, it is required to measure two samples per minimum wavelength λ_{\min} [16, 17]:

$$\lambda_{\min} = v_{\min} / f_{\max}, \quad (1.1)$$

where v_{\min} and f_{\max} are the minimum wave propagation velocity in the medium and maximum frequency, respectively. To image a medium with a velocity of 250 m/s, 1.25 m source and receiver sampling is necessary to avoid aliasing and obtain a 2.5 m wavelength with 100 Hz maximum frequency. Since the deeper subsurface is composed of rocks with higher velocities, it does not require as dense sampling. Therefore, the near-surface in many situations dictate how to acquire seismic data [12]. However, fine sampling can be prohibitively expensive.

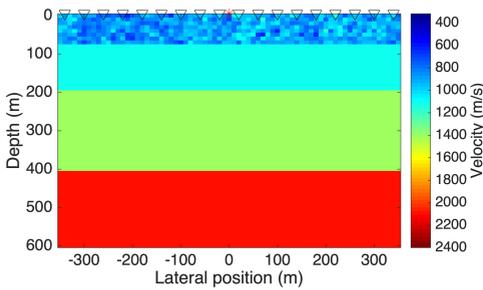
An alternative approach is to reduce the acquisition costs by avoiding the Nyquist-Shannon sampling criterion. However, periodic data subsampling results in coherent artifacts (aliased energy) that can be challenging to remove. Another way to reduce the acquisition costs is to acquire the data with randomized subsampling as suggested by the field of compressive sensing (CS) [18–20]. Such acquisition scenario also results in artifacts that need to be removed. However, they are easier to handle with data interpolation as they can be shown incoherent in the appropriate domain. But for land data, the weathering layers additionally add their influence, which exacerbates the complexity of the situation. As a result, interpolation of randomly subsampled data influenced by the weathering layers results in noisy and low resolution reconstruction (Figure 1.6).

1.5. UNDOING THE WEATHERING EFFECTS

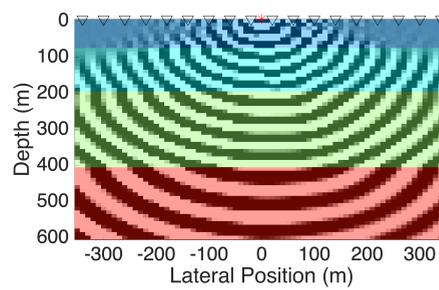
Since resolving the near-surface with dense source and receiver sampling can be prohibitively expensive, many methods are proposed to approximate and remove the weathering layers effects on the subsurface. We refer to that as near-surface estimation and correction. The weathering layers introduce undesired wave propagation effects on subsurface reflectors. They mostly manifest themselves as time-shifts, which are clear when



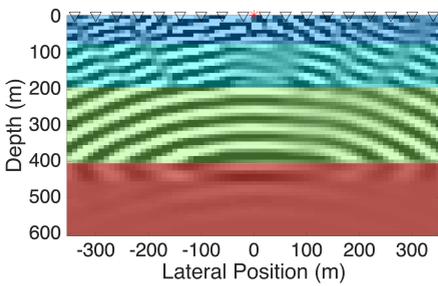
(a) Real complex near-surface



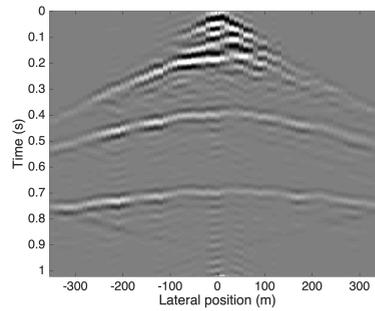
(b) Velocity model



(c) Transmitted wavefield



(d) Reflected wavefield



(e) Recorded data

Figure 1.4: Illustration of the influence of a complex near-surface. (a) A real near-surface that shows its rapidly varying nature, extracted from [15]. (b) Complex near-surface model that depicts the real one in (a) overlaying a simple subsurface model. The (c) transmitted and (d) reflected wavefields in the subsurface for one frequency (26 Hz) overlaid over the velocity model. (e) The recorded reflected wavefield. Even though the subsurface model is similar to that in Figure 1.1(a), except for the near-surface layer, the resultant wavefields are drastically different compared to those in Figure 1.1.

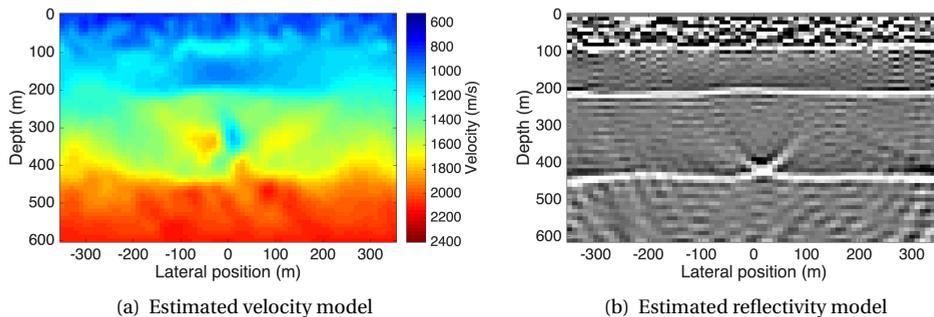


Figure 1.5: Joint migration inversion of data influenced by a complex near-surface (Figure 1.4(e)) to estimate a (a) velocity and (b) reflectivity models. Due to the poor initial near-surface model, the inversion results in wrong structures, which can lead to errors when making critical decisions.

comparing the second and third reflectors of Figures 1.1(d) and 1.4(e). These time shifts are known as ‘statics’ [7, 21]. To overcome the acquisition limitations, near-surface (statics) estimation to approximate the weathering layers effect, followed by near-surface (statics) correction are essential steps for land seismic data processing.

Near-surface correction is an approximation of a complex problem [21, 22]. The commonly used term ‘statics correction’ refers to constant time-shifts applied to the whole trace, which is contrary to the ‘dynamic-correction’ that is related to the subsurface such as velocity changes. While statics estimation should be optimally replaced by the dynamic velocity analysis of the near-surface [23, 24], it is still a convenient way to overcome the limitations of data acquisition, velocity and image estimation algorithms.

An abundance of statics estimation methods already exist. These can be generally divided into two categories based on their resolution capabilities: (i) long-wavelength and (ii) short-wavelength statics estimation methods [7, 24].

1.5.1. THE LONG-WAVELENGTH COMPONENT

Long-wavelength statics correction methods are used to redatum seismic data to a certain depth level below the near-surface weathering layers to overcome their effects. In practice, estimation of long-wavelength statics is commonly achieved by model-based methods that depend on a velocity model of the near-surface weathering. The velocity model needs to delineate the low-wavenumber features as they are necessary to derive the long-wavelength component of the statics. Based on a velocity model, time-shifts (statics) to redatum the data are computed. Methods in this category include, but are not limited to, uphole surveys, refraction traveltime tomography [25], waveform tomography [26], multi-physics inversion [27], image-based modelling [28] and smart distributed acoustic sensing [29]. Redatuming can also be implemented without the need of a near-surface velocity model using for example the common focus point technology [30, 31] and virtual source imaging with buried receivers [32].

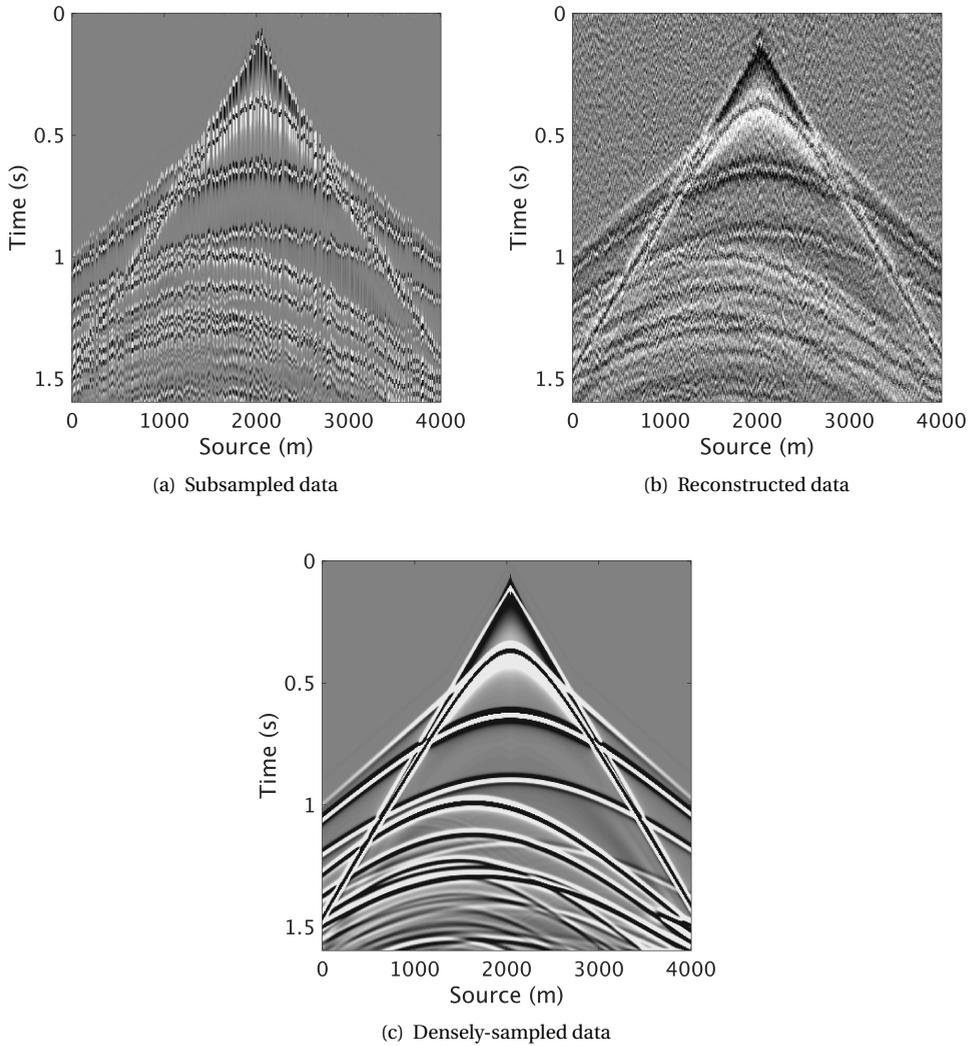


Figure 1.6: Illustration of reconstruction of (a) randomly subsampled data influenced by a complex near-surface. (b) The resultant reconstruction, which is noisy and of low resolution due to the incoherency of the data. (c) The desired reconstruction.

When the estimated redatuming operators or when the estimated near-surface velocity contains errors, some statics will remain uncorrected for. Since model-based statics correction methods depend on a smooth near-surface model, the derived statics to compensate for the near-surface effect are also expected not to be detailed. As a result, these methods do not account for rapid changes in surface elevation, base of the weathering layers and the weathering layers' velocity, which can result in distortion of the propagating wavefronts' phases and amplitudes [7, 21]. Therefore, correcting for short-wavelength statics, which are also referred to as residual statics, becomes necessary.

1.5.2. THE SHORT-WAVELENGTH COMPONENT

Figure 1.4(a) shows the rapid variations in properties of the weathering layers that a simple velocity model can not describe. They are behind the high-resolution near-surface model requirement, which consequently relies on a dense acquisition grid. Since only a smooth model is used for long-wavelength statics correction, short-wavelength statics correction becomes necessary to capture more details of the near-surface. That is why short-wavelength statics estimation and correction are routinely applied on land seismic data. Despite the advancements in subsurface model-building technology, short-wavelength statics estimation and correction remain irreparable for land seismic data processing and are required prior to full waveform inversion of land data [33, 34]. Methods in this category are mainly data-driven, where the statics are estimated using the data directly without the necessity to estimate a near-surface velocity model. Some of the commonly used methods include linear traveltime inversion [35], stack power maximization [23] and non-linear inversion with global Monte-Carlo techniques [36].

VELOCITY-WEATHERING CHALLENGE

Even though a near-surface velocity model is not necessary for short-wavelength statics estimation, normal moveout (NMO) or migration velocity are required to obtain NMO-corrected or migrated gathers [37]. Migration or NMO correction, which is commonly applied, reduces the dynamic variations to the level of residual moveout [21]. If the migrated or NMO-corrected gathers are not aligned, then short-wavelength statics are the main cause. However, short-wavelength statics influence velocity estimation. Otherwise, their estimation becomes irrelevant for subsurface model estimation. Due to the complex changes in the near-surface layer, the estimated subsurface velocity model displayed in Figure 1.5(a) contains erroneous structures. At the same, using an inaccurate velocity can lead to erroneous statics as parts of the dynamic variations may end up in the statics solution. To avoid that in practice, velocity and statics estimation are applied iteratively to improve each others performance [7], which can be efforts- and time-consuming.

MULTIPLES AND DATA WINDOWING

It follows from the previous point that the data must only contain primary reflections as multiple reflections are unaligned after migration or NMO correction. Therefore, multiple removal [38] becomes essential. That can be substituted with selection of a window that only contains primaries, which may not always be possible as primaries and multiples can be overlapping. If the selected window contains multiples, the estimated statics can be wrong. Consequently, unfocused or erroneous structures can be created. This can be

seen from the stack after residual statics correction with stack power maximization (SPM) [23] using a suboptimal window (Figure 1.7(b)). As a result, new false structures that do not exist in the original stack (Figure 1.7(a)) are created, which can lead to errors when making critical decisions. Despite using an optimal window for residual statics estimation with stack power maximization (Figure 1.7(c)), the resolution and stack power are much lower compared to the reference statics-free stack displayed in Figure 1.7(d). That is related to the surface-consistency assumption, which we introduce next.

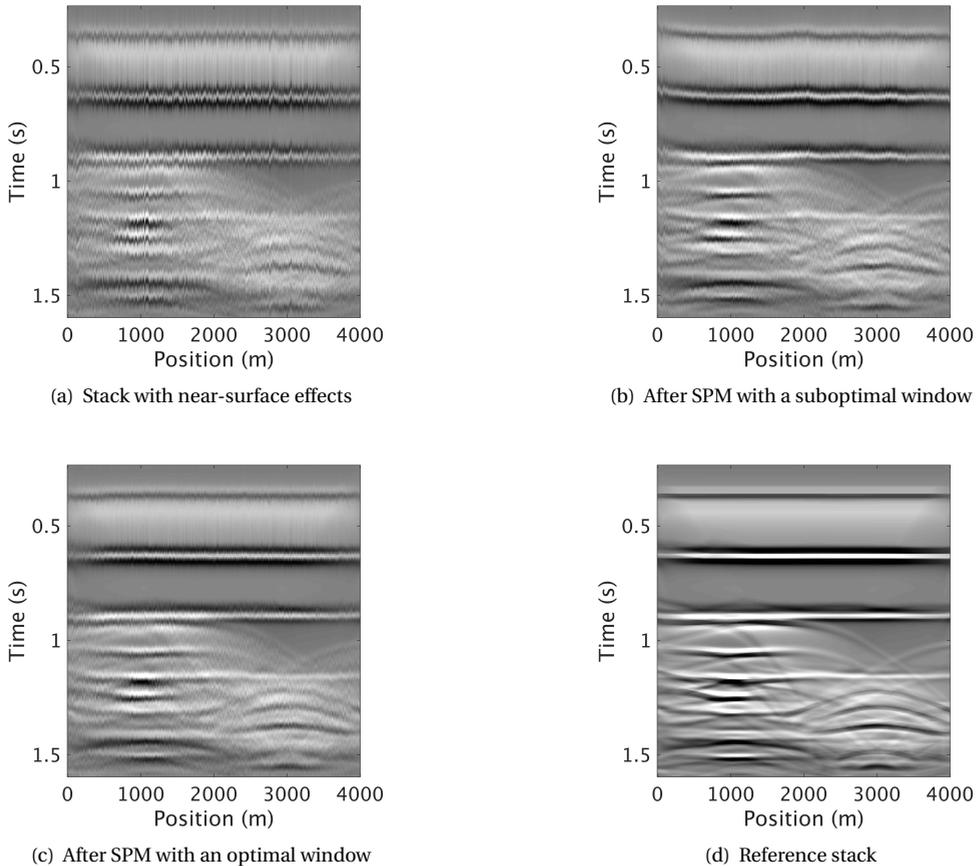


Figure 1.7: Stack sections of (a) data with near-surface effects and data after SPM with (b) a suboptimal and (c) an optimal window. (d) Reference stack of statics-free data.

SURFACE-CONSISTENCY

Conventional residual statics estimation methods assume that the near-surface weathering layers result in surface-consistent effects [7, 21, 24, 39]. With this assumption, ray-paths in the near-surface are considered vertical making them independent of the source-

receiver offsets. Therefore, the static t_{s-r} related to a pair of source s and receiver r depends only on their positions at the surface. Despite still being in the INTRODUCTION chapter, we include a bit of simple mathematics to fully grasp the concept, which can also be skipped for the sake of simplicity. According to the surface-consistency assumption, conventional residual statics estimation uses the following model [7, 21, 35]:

$$t_{s-r} = t_s + t_r + t_m + t_{\text{rmo}}, \quad (1.2)$$

where t_s and t_r are the surface-consistent source and receiver static, respectively. t_m is referred to as the structural term, which accounts for the time-shift compared to a reference common midpoint (CMP). t_{rmo} is a residual moveout term that is assumed parabolic. To estimate t_{s-r} , it is required to construct a pilot trace, commonly by stacking a number of traces together, followed by its cross-correlation with the rest of the data. After that, t_{s-r} is decomposed into the four terms in equation 1.2. For statics correction, only the terms t_s and t_r are applied in a surface-consistent manner. In principle, t_{s-r} should lead to perfectly aligned primary reflection after migration or NMO-correction, assuming that the velocity model is accurate. However, the model described with equation 1.2 does not consider the time-shift associated with the source-receiver offset. Therefore, the estimated residual statics can be inaccurate when the surface-consistency assumption is violated. Since the near-surface effects are composed of surface- and non-surface-consistent statics (Figure 1.7(a)), the stack after stack power maximization residual statics correction (Figure 1.7(c)) shows lower resolution and power compared to the statics-free stack (Figure 1.7(d)). This situation represents the reality, where surface-consistency is not strictly satisfied, which calls for an additional non-surface-consistent statics estimation and correction step.

NON-SURFACE-CONSISTENCY

Near-surface correction tries to remove the weathering layers effect on subsurface targets by simplifying a more complex problem. Adding the surface-consistency assumption further removes details from the already simplified near-surface correction model. [40] argues that surface-consistency is not strictly correct. [24] further analyzes the assumption and concludes that it works in practice in many situations as surface-consistent statics are the equivalent of a simple near-surface model. However, they are not valid in situations where the near-surface layers are thick, the velocity contrast between the near-surface and immediate subsurface is low or the dip of the reflector is steep. Examples where it is known to fail are in the presence of a permafrost near-surface layer of high velocity followed by lower-velocity material or when the overburden is complex. Therefore, unresolved statics can still remain in the data if surface-consistency is not strictly satisfied.

To improve the near-surface solution, non-surface-consistent methods can be necessary. [41] propose a framework that utilizes the common-angle domain based on ray-path interferometry. The method requires horizons picking and pilot trace construction, which are not trivial tasks when the events picked are not continuous or in the presence of noise. Another type of non-surface-consistent statics correction is trim statics that is based on cross-correlation of individual traces with a model trace usually formed by stacking a number of CMP gathers [24]. Trim statics are usually small relative to the surface-consistent statics. They are commonly used to flatten NMO-corrected reflectors

by brute-force to remove any remaining statics without surface-consistent decomposition (equation 1.2). However, in the presence of noise, the method can result in erroneous time-shifts. [42] demonstrate its danger in aligning noise as signal and describes it by "playing with fire". More recently, [43] use deep learning to estimate aligned events from unaligned ones by assuming that trim statics are the cause of unalignment. However, this approach requires accurate velocity as well as accurate knowledge of primaries to avoid aligning multiples as primaries. Surface-consistent statics estimation methods can also be evaluated over multiple offset windows. This option, however, adds to the computational costs and decreases the number of traces in each offset window, which may lead to errors as a result of low signal-to-noise ratio.

1.6. A SUMMARY OF EXISTING CHALLENGES

The following list provides a summary of existing challenges related to the near-surface weathering layers:

- (i) Despite the advancements in subsurface model-building technology, dealing with different scales at the near-surface and deeper subsurface remains challenging. Therefore, short-wavelength statics estimation and correction is essential for data processing and prior to full waveform inversion of land data. However, to reach to the short-wavelength statics estimation stage, a long processing workflow can be necessary. Additionally, knowledge of the subsurface such as velocity estimation and multiples attenuation can be needed. Moreover, multiple short-wavelength surface- and non-surface-consistent near-surface corrections are usually essential to overcome the complexity of the weathering layers. Since short-wavelength statics correction focuses on enhancing primaries, multiples can inevitably get destroyed, which limits imaging and inversion of primaries and multiples. These impediments make short-wavelength statics estimation and correction inefficient, efforts- and time-consuming for both processing workflows and prior to full waveform inversion.
- (ii) To estimate accurate subsurface models, dense data sampling according to the Nyquist-Shannon sampling criterion is essential. However, that can be prohibitively expensive. Reduction of the acquisition costs can be achieved by randomized subsampling. While data reconstruction techniques are successful in the marine environment, they suffer in the presence of complex weathering layer as the data exhibit low coherency. To improve the coherency of the data and enable their reconstruction, near-surface correction becomes essential, which also faces the previously mentioned challenges (Challenge (i)). When the subsampling ratio and near-surface effects are large, data reconstruction can fail as prior near-surface estimation and correction become difficult.
- (iii) Seismic data processing, imaging and inversion all share the same objective, which is to create accurate subsurface models. However, data processing can be disadvantageous as it requires human interaction to perform many steps that can be efforts- and time-consuming. While it is currently feasible to directly invert for subsurface models from raw marine data without prior data processing, it is still challenging

to do so from land data in the presence of complex near-surface weathering layers. When the near-surface model is not accurate, the subsurface model also becomes inaccurate. As a result, model building with land data still relies on data processing to remove the short-wavelength near-surface effects prior to the inversion [33]. On the one hand, data processing improves the inversion, but on the other hand, it reduces the efficiency. When the target lies in the shallow near-surface, e.g. for engineering and construction purposes, estimation of the near-surface models become essential.

1.7. RESEARCH QUESTIONS

To address the aforementioned challenges, we need to answer the following key questions:

1. Can we correct for the near-surface effects (surface- and non-surface-consistent) on 2D and 3D data with a model-independent approach? (Chapters 2, 3 and 5).
2. Can we reconstruct land data acquired on complex near-surface regimes with a compressive sensing acquisition design? (Chapter 4).
3. Is it feasible to estimate the near- and sub-surface velocity and reflectivity models simultaneously with joint migration inversion from data affected by weathering layers? (Chapter 6).

1.8. OBJECTIVES AND CONTRIBUTIONS

The main objective of our research is to obtain accurate subsurface properties from seismic data collected onshore in the presence of complex near-surface weathering layers. These models need to be of high accuracy and resolution to allow for making critical decisions. Accordingly, we derive the ensuing key objectives to answer the aforementioned research questions, and summarize the enabling contributions:

1. **Efficient and accurate estimation of short-wavelength surface- and non-surface-consistent near-surface effects with a model-independent low-rank-based approach (Chapter 2).**

To overcome the challenges of existing short-wavelength near-surface correction methods (Challenge (i)), we propose a novel model-independent low-rank-based near-surface estimation and correction in the midpoint-offset-frequency domain. In this domain, 2D seismic data exhibit low rank structures, which get destroyed due to the influence of the weathering layers. Accordingly, the method makes use of the redundant nature of seismic data that allows for accurate approximation by low-rank matrices. To estimate the time shifts that compensate for the weathering effects, we cross-correlate a data set influenced by the near-surface weathering layers with its low-rank-approximated version. Since we estimate time shifts (statics) and no longer the directly low-rank approximated data (as shown by [44, 45]), we preserve the amplitude versus offset response. To mitigate the poor performance of low-rank approximation at high frequencies, we exploit the common statics amongst different frequency bands; we first utilize the low frequency bands, where low-rank approximation performs

better than high frequency bands, to estimate the statics followed by near-surface correction of the full-band data. Since these statics are not accurate enough for the total bandwidth, they get updated during statics estimation when including the high frequencies. To improve the estimated statics and to alleviate the need for accurate rank selection for low-rank approximation, we implement the method in an iterative and multi-scale fashion.

The proposed method relaxes the surface-consistency assumption to estimate more accurate surface- and non-surface-consistent statics at once. In practice, multiple steps are performed, where surface-consistent residual statics are first estimated followed by a brute-force non-surface-consistent step, to account for both components. Moreover, the proposed method does not require knowledge of the subsurface for near-surface correction. Therefore, statics estimation becomes independent of errors in the subsurface velocity model. Consequently, we can reduce the efforts- and time-consuming multiple iterations of near-surface correction and velocity estimation commonly needed for conventional methods to improve their performance. Additionally, near-surface correction of the total wavefield without windowing, e.g. to select aligned primaries or avoid noise, becomes feasible. We demonstrate that the proposed method preserves the seismic response of the structure (both kinematic and dynamic), while estimating accurate statics in a computationally efficient manner. The aforementioned benefits make the proposed method favorable compared to existing techniques as it overcomes their limitations.

2. **Preservation of the subsurface model while removing the near-surface effects prior to reverse time migration with a low-rank-based approach (Chapter 3).**

There can be a concern that low-rank-based methods remove subtle details from the subsurface model. Despite the success of the low-rank-based near-surface correction (Chapter 2), we demonstrate that it also preserves subtle subsurface structures and improves the resolution of images obtained with reverse time migration.

3. **To reconstruct densely-sampled land data from cost-efficient randomly subsampled data influenced by the weathering layers (Chapter 4).**

Our contributions are towards rank-reduction-based reconstruction of randomly subsampled land data influenced by the near-surface weathering layers. Such problem requires solving for the near-surface and subsampling effects as both contribute to the low-rank structure destruction (both effects render coherent energy incoherent that breaks the low-rank structure typically associated with statics-free densely sampled data). The low-rank structure destruction means that the singular values no longer decay fast, while in the appropriate domain, seismic data can be well-approximated by low-rank matrices. Although the low-rank-based near-surface estimation and correction show its potential for periodically- and densely-sampled data (Chapter 2), its application to randomly- (non-periodically)-sampled data on the grid is yet to be examined.

When collecting economical data in complex near-surface regimes, separation between the subsampling and weathering effects becomes difficult. Consequently,

low-rank-based model-independent near-surface correction followed by interpolation, or vice-versa performs poorly. This is because interpolation and near-surface correction require that data incoherency comes only from subsampling or statics, respectively, but not both. To obtain accurate densely-sampled data, we propose joint reconstruction with rank-reduction-based near-surface correction and interpolation. With this approach, near-surface correction improves interpolation, and vice versa. Since we need to compute low-rank approximation for near-surface correction, we utilize it as a cost-free initial solution for the rank-minimization optimization problem. We also lower the number of required transformations between the acquisition and transform domain as both interpolation and near-surface correction operate in the midpoint-offset domain. Therefore, the proposed reconstruction additionally improves the computational efficiency. To demonstrate its potential, we apply it to synthetic data and field data affected by complex weathering layers and noise.

4. **3D Model-independent low-rank-based near-surface correction (Chapter 5).**

While Objective 1 deals with 2D data, the Earth is a 3D object that requires acquisition of 5D data for proper subsurface model estimation. For 5D data, the limitations and challenges of conventional near-surface correction (Challenge (i)) are magnified. To avoid them, we propose a 5D Model-independent low-rank-based near-surface correction. To compute the singular value decomposition of 5D data volumes with 1 temporal and 4 spatial dimensions, which is necessary for low-rank approximation, we need to perform matricization of the 5D data, i.e. organization of the 5D data into matrices. At the same time, it is essential that the chosen organization domain reveals the underlying low-rank structure. Therefore, we first analyze different matricization domains that can be used to organize the 5D data. We show that — in the potential domain — the near-surface weathering layers render coherent energy incoherent, which results in slowly decaying singular values compared to the statics-free data that are of low-rank nature. According to the findings, we describe the details of our proposed algorithm. It enjoys the same benefits as those described in Objective 1, in addition to being applicable to 5D data. We demonstrate its performance on synthetic and field data.

5. **To simultaneously estimate near- and sub-surface velocity and reflectivity models with joint migration inversion (Chapter 6).**

The near-surface model gets treated separately from the subsurface model due to its complex nature and its effect on the subsurface model. However, the optimal goal is not to remove the near-surface effects with data processing (Challenge (iii)), but to accurately estimate near- and sub-surface models simultaneously. To do so, we use the inherent scale separation of joint migration inversion that estimates a low-wavenumber velocity and high-wavenumber reflectivity. Since rapid variations in surface elevation and near-surface model result in high wavenumber effects, they end up affecting the reflectivity model. At the same time, the estimated reflectivity influences velocity estimation. Consequently, JMI provides erroneous subsurface models in the presence of complex weathering layers. To mitigate that, we use multi-scale low-rank updates in the reflectivity domain. The proposed method reduces the near-surface effects at the initial iterations, but it allows more details of the near-surface model to enter the solution at later iterations.

In the end, we estimate accurate near- and sub-surface models simultaneously without the need to bypass the weathering layers. In the next section we provide the dissertation's outline, which follows the same order of the aforementioned objectives.

1.9. DISSERTATION OUTLINE

The dissertation is organized as follows. **Chapter 2** provides the details of our proposed low-rank-based near-surface correction in the midpoint-offset domain, which we apply to 2D synthetic data and field data affected by complex weathering layers and noise. Even though **Chapter 2** demonstrates the success of the proposed method, we show in **Chapter 3** that it is also capable of preserving subtle structures and improving the resolution of reverse time migrated images. While **Chapters 2 and 3** demonstrate their potential on periodically and densely sampled data, real data is not always well-sampled due to the associated acquisition costs. To deal with that, we propose in **Chapter 4** a joint reconstruction method with rank-reduction-based near-surface correction and interpolation. To estimate 3D subsurface models from land data influenced by the weathering layers, we find a proper matricization domain of 5D data volumes that satisfies the requirements of the rank-based near-surface correction. Accordingly, we extend the model-independent low-rank-based near-surface correction to 3D data (**Chapter 5**). In **Chapter 6**, we use multi-scale low-rank image updates to allow for estimation of the near- and sub-surface models simultaneously with joint migration inversion. After that, we provide general conclusions and future recommendations in **Chapter 7**.

In **Appendix A**, we deviate from land to marine data, which is beyond the scope of this dissertation. Therefore, the near-surface is no longer influenced by the weathering layers. However, the inversion of marine data is sensitive to the water velocity (the overburden in the case of ocean bottom receivers) as we show in **Appendix A**. That is analogous to the weathering layers, where errors in the overburden model lead to erroneous subsurface models as shown in **Chapter 6**.

REFERENCES

- [1] U. N. D. of Economic and S. Affairs, *The Sustainable Development Goals: Report 2022* (UN, 2022).
- [2] IEA, *World energy outlook 2022*, (IEA Paris, France, 2022).
- [3] V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, *et al.*, *Global Warming of 1.5 C: IPCC special report on impacts of global warming of 1.5 C above pre-industrial levels in context of strengthening response to climate change, sustainable development, and efforts to eradicate poverty* (Cambridge University Press, 2022).
- [4] T. G. Leighton, *What is ultrasound?* Progress in biophysics and molecular biology **93** (2007), pp. 3–83, <https://doi.org/10.1016/j.pbiomolbio.2006.07.026>.
- [5] A. Fichtner, B. L. Kennett, H. Igel, and H.-P. Bunge, *Theoretical background for continental-and global-scale full-waveform inversion in the time-*

- frequency domain*, *Geophysical Journal International* **175** (2008), pp. 665–685, <https://doi.org/10.1111/j.1365-246X.2008.03923.x>.
- [6] T. L. Szabo, *Diagnostic ultrasound imaging: inside out* (Academic press, 2004) <https://doi.org/10.1016/C2011-0-07261-7>.
- [7] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <http://dx.doi.org/10.1190/1.9781560801580>.
- [8] J. F. Claerbout, *Imaging the Earth's interior*, Vol. 1 (Blackwell scientific publications Oxford, 1985).
- [9] K. Aki and P. Richards, *Quantitative Seismology*, Geology (University Science Books).: Seismology (University Science Books, 2002).
- [10] A. Berkhout, *Seismic Migration: Imaging of Acoustic Energy by Wave Field Extrapolation*, Developments in solid earth geophysics (Elsevier Science, 2012).
- [11] R. E. Sheriff and L. P. Geldart, *Exploration seismology* (Cambridge university press, 1995).
- [12] T. H. Kebo and P. G. Kelamis, *Focus on land seismic technology: The near-surface challenge*, *The Leading Edge* **31** (2012), pp. 62–68, <https://doi.org/10.1190/1.3679329>.
- [13] A. Berkhout, *An outlook on the future of seismic imaging, Part III: Joint migration inversion*, *Geophysical Prospecting* **62** (2014), pp. 950–971, <https://doi.org/10.1111/1365-2478.12158>.
- [14] D. Verschuur, X. Staal, and A. Berkhout, *Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering*, *The Leading Edge* **35** (2016), pp. 1037–1046, <https://doi.org/10.1190/tle35121037.1>.
- [15] C. Crook, A. Crook, C. Sandanayake, and P. Stephenson, *Land seismic acquisition testing strategies and results — Southern Chad, Africa 2013-2015, first break* **34** (2016), <https://doi.org/10.3997/1365-2397.2016008>.
- [16] H. Nyquist, *Certain topics in telegraph transmission theory*, *Transactions of the American Institute of Electrical Engineers* **47** (1928), pp. 617–644.
- [17] C. E. Shannon, *Communication in the presence of noise*, *Proceedings of the IRE* **37** (1949), pp. 10–21.
- [18] D. L. Donoho, *Compressed sensing*, *IEEE Transactions on information theory* **52** (2006), pp. 1289–1306, <https://doi.org/10.1109/TIT.2006.871582>.
- [19] E. J. Candès and M. B. Wakin, *An Introduction To Compressive Sampling*, *IEEE signal processing magazine* **25** (2008), pp. 21–30, <http://dx.doi.org/10.1109/MSP.2007.914731>.

- [20] F. J. Herrmann, M. P. Friedlander, and O. Yilmaz, *Fighting the curse of dimensionality: compressive sensing in exploration seismology*, *Signal Processing Magazine, IEEE* **29** (2012), pp. 88–100, <http://dx.doi.org/10.1109/MSP.2012.2185859> .
- [21] M. T. Taner, F. Koehler, and K. A. Alhilali, *Estimation and correction of near-surface time anomalies*, *Geophysics* **39** (1974), pp. 441–463, <https://doi.org/10.1190/1.1440441> .
- [22] D. Marsden, *Static corrections—A review, Part III*, *The leading edge* **12** (1993), pp. 210–216.
- [23] J. Ronen and J. F. Claerbout, *Surface-consistent residual statics estimation by stack-power maximization*, *Geophysics* **50** (1985), pp. 2759–2767, <https://doi.org/10.1190/1.1441896> .
- [24] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999) <https://doi.org/10.1190/1.9781560801818> .
- [25] X. Zhu, D. P. Sixta, and B. G. Angstman, *Tomostatics: Turning-ray tomography + static corrections*, *The Leading Edge* **11** (1992), pp. 15–23, <https://doi.org/10.1190/1.1436864> .
- [26] J. Sheng, A. Leeds, M. Buddensiek, and G. T. Schuster, *Early arrival waveform tomography on near-surface refraction data*, *Geophysics* **71** (2006), pp. U47–U57, <https://doi.org/10.1190/1.2210969> .
- [27] D. Colombo, M. Mantovani, M. Sfolciaghi, P. Van Mastrigt, A. Al-Dulaijan, and T. Nafie, *Near surface solutions in South Rub Al-Khali, Saudi Arabia applying seismic-gravity joint inversion and redatuming*, *First Break* **28** (2010), <https://doi.org/10.3997/1365-2397.28.2.38216> .
- [28] O. Yilmaz, *An image-based effective-medium modeling of near-surface anomalies*, *The Leading Edge* **32** (2013), pp. 394–401, <https://doi.org/10.1190/tle32040394.1> .
- [29] A. Bakulin, P. Golikov, R. Smith, K. Erickson, I. Silvestrov, and M. Al-Ali, *Smart das upholes for simultaneous land near-surface characterization and subsurface imaging*, *The Leading Edge* **36** (2017), pp. 1001–1008, <https://doi.org/10.1190/tle36121001.1> .
- [30] P. G. Kelamis, K. E. Erickson, D. J. Verschuur, and A. J. Berkhou, *Velocity-independent redatuming: A new approach to the near-surface problem in land seismic data processing*, *The Leading Edge* **21** (2002), pp. 730–735, <https://doi.org/10.1190/1.1503185> .
- [31] M. Al-Ali and D. Verschuur, *An integrated method for resolving the seismic complex near-surface problem*, *Geophysical Prospecting* **54** (2006), pp. 739–750, <https://doi.org/10.1111/j.1365-2478.2006.00575.x> .
- [32] A. Bakulin and R. Calvert, *Virtual source: new method for imaging and 4D below complex overburden*, in *SEG Technical Program Expanded Abstracts 2004* (Society of Exploration Geophysicists, 2004) pp. 2477–2480, <https://doi.org/10.1190/1.1845233> .

- [33] Ö. Yilmaz, K. Gao, M. Delic, J. Xia, L. Huang, H. Jodeiri, and A. Pugin, *A reality check on full-wave inversion applied to land seismic data for near-surface modeling*, *The Leading Edge* **41** (2022), pp. 40–46, <https://doi.org/10.1190/tle41010040.1>.
- [34] J. Sheng, C. Reta-Tang, F. Liu, A. Vázquez Cantú, and A. Cabrales Vargas, *Full-waveform inversion and FWI imaging for land data*, in *Second International Meeting for Applied Geoscience & Energy* (2022) pp. 827–831, [10.1190/image2022-3749887.1](https://doi.org/10.1190/image2022-3749887.1).
- [35] R. A. Wiggins, K. L. Larner, and R. D. Wisecup, *Residual statics analysis as a general linear inverse problem*, *Geophysics* **41** (1976), pp. 922–938, <https://doi.org/10.1190/1.1440672>.
- [36] D. H. Rothman, *Nonlinear inversion, statistical mechanics, and residual statics estimation*, *Geophysics* **50** (1985), pp. 2784–2796, <https://doi.org/10.1190/1.1441899>.
- [37] J. Xu, H. Zhang, and J. Zhang, *A robust surface-consistent residual phase correction method based on migrated gathers*, *Exploration Geophysics* **49** (2018), pp. 336–344, <https://doi.org/10.1071/EG17017>.
- [38] D. J. Verschuur, *Seismic multiple removal techniques: past, present and future (Revised edition)* (EAGE publications Houten, Netherlands, 2013) <https://doi.org/10.3997/9789073834965>.
- [39] D. Marsden, *Static corrections—A review, Part II*, *The Leading Edge* **12** (1993), pp. 115–120, <https://doi.org/10.1190/1.1436936>.
- [40] R. E. Sheriff, *Encyclopedic Dictionary of Applied Geophysics, Fourth edition* (Society of Exploration Geophysicists, 2002) <https://library.seg.org/doi/pdf/10.1190/1.9781560802969>.
- [41] D. C. Henley, *Interferometric application of static corrections*, *Geophysics* **77** (2012), pp. Q1–Q13, <https://doi.org/10.1190/geo2011-0082.1>.
- [42] C. P. Ursenbach and J. C. Bancroft, *Playing with fire: Noise alignment in trim and residual statics*, in *SEG Technical Program Expanded Abstracts 2001* (Society of Exploration Geophysicists, 2001) pp. 1973–1976, <https://library.seg.org/doi/pdf/10.1190/1.1816525>.
- [43] A. Breuer, N. Ettrich, and P. Habelitz, *Deep learning in seismic processing: Trim statics and demultiple*, in *SEG Technical Program Expanded Abstracts 2020* (Society of Exploration Geophysicists, 2020) pp. 3199–3203, <https://library.seg.org/doi/pdf/10.1190/segam2020-3427887.1>.
- [44] A. M. Alfaraj, R. Kumar, and F. J. Herrmann, *Automatic statics and residual statics correction with low-rank approximation*, in *80th EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2018) pp. 1–5, <https://doi.org/10.3997/2214-4609.201801107>.

- [45] A. Alfaraj, M. Almubarak, and E. Herrmann, *Correcting for short-wavelength statics with low rank approximation*, in *81st EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2019) pp. 1-5, <https://doi.org/10.3997/2214-4609.201901205>.

2

2D MODEL-INDEPENDENT NEAR-SURFACE ESTIMATION AND CORRECTION

Since rapid variations of properties of the weathering layers are challenging to estimate during velocity model building, short-wavelength near-surface correction becomes an essential step to obtain accurate subsurface models. Surface-consistency forms the basis for short-wavelength near-surface correction. When raypaths in the near-surface diverge from normal-incidence or when the migration or normal moveout (NMO) velocity are inaccurate, surface-consistent methods may fail to estimate accurate statics. Existing non-surface-consistent techniques can be prone to errors due to the need to construct pilot traces or pick horizons, while imposing additional computational costs. To overcome these limitations and correct for the surface- and non-surface-consistent statics, we propose a low-rank-based residual statics (LR-ReS) estimation and correction framework. The method makes use of the redundant-nature of seismic data by utilizing its low-rank structure in the midpoint-offset-frequency domain. Due to the near-surface effect, the low-rank structure gets destroyed. Therefore, we estimate the statics by means of low-rank approximation and cross-correlation. To alleviate the need for accurate rank selection for low-rank approximation and for improved statics estimation, we implement the method in an iterative and multi-scale fashion. Since the low-rank approximation deteriorates at high frequencies, we utilize its better performance at low frequencies and exploit the common statics amongst the different frequency bands. The LR-ReS estimation and correction can be applied to data without migration or NMO correction, which makes statics estimation independent of the migration or NMO velocity errors. Consequently, it can reduce the multiple iterations of velocity estimation and short-wavelength statics correction commonly needed for conventional methods to improve their performance. Moreover, the LR-ReS estimation does not

This chapter is a modified version of the paper "A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, Low-rank-based residual statics estimation and correction, *Geophysics* 88 (2023), pp. V215–V231."

require windowing of a noise-free area containing aligned primaries nor mute to avoid the NMO stretch effect, which enables statics correction of the wavefield of all offsets. To evaluate the performance of our proposed method, we apply it to simulated data and a challenging field data set affected by complex weathering layers and noise, which show substantial improvement compared to conventional short-wavelength statics correction.

2.1. INTRODUCTION

Land seismic data sets are often challenged by the near-surface weathering layers. They can lead to wave propagation effects that manifest themselves mostly as undesired time-shifts, commonly called 'statics' [1, 2]. While statics correction should be replaced by dynamic velocity analysis of the near-surface as has been suggested by [3] and [4], it is still a convenient way to overcome the limitations in acquisition, velocity and image estimation engines.

Due to rapid changes in surface-elevation, base of the weathering layers and the weathering layers' velocity, land data get affected by short-wavelength (residual) statics [2]. Conventional short-wavelength statics correction methods rely on the surface-consistency assumption, where raypaths in the near-surface are assumed vertical [1–5]. In this situation, the statics depend only on the locations of sources and receivers at the surface without considering their offsets. [6] argues that the surface-consistency assumption is not strictly correct. [4] further analyzes the assumption and concluded that it works in practice in most cases as surface-consistent statics are approximately equivalent to a simple near-surface model. The majority of surface-consistent methods estimate residual statics using normal moveout- (NMO)-corrected common midpoint (CMP) gathers [2]. NMO correction should ideally remove the dynamic component related to the velocity such that the residual static component becomes more clear. In other words, surface-consistent methods assume that traces are aligned after NMO correction, and if they are not, then residual statics or residual NMO are the main causes. Moreover, it becomes easier to obtain a reference trace with less near-surface imprint from the presumably aligned traces after NMO correction. Other than NMO correction, surface-consistent residual statics correction methods require windowing of a noise-free area containing primaries because they are the aligned reflections. However, short-wavelength statics can also lead to ambiguity during NMO velocity estimation. Therefore, multiple iterations of NMO velocity estimation and short-wavelength statics correction are usually carried out to improve their performance [2], which can be efforts- and time-consuming. Using migrated gathers, [7] estimate surface-consistent statics. The method also needs accurate velocity for migration that can be computationally expensive. When the surface-consistency assumption is violated, errors in the estimated statics may arise, which calls for a non-surface-consistent near-surface correction. Throughout the paper, the term 'residual' statics is not only limited to surface-consistent statics, but also non-surface-consistent ones.

[8] proposes a framework that uses the common-angle domain based on raypath interferometry to correct for non-surface-consistent statics. The method requires horizon picking and pilot trace construction, which are not trivial tasks when the events picked are not continuous or in the presence of noise. Another type of non-surface-consistent statics correction is trim statics that is based on cross-correlation of individual traces with a model trace usually formed by stacking a number of CMP gathers [4]. [9] demonstrate its danger in aligning noise as signal and described it by "playing with fire". More recently, [10] use deep learning to estimate aligned events from unaligned ones by assuming that trim statics are the cause of unalignment. However, this approach may require accurate velocity as well as accurate knowledge of primaries to avoid aligning multiples as primaries. Surface-consistent statics estimation methods can also be evaluated over

multiple offset windows. This option, however, adds to the computational costs and decreases the number of traces in each offset window, which may lead to errors as a result of low signal-to-noise ratio. The above methods still need access to an NMO velocity model, which can influence the estimated statics. In an alternative approach, one can process a seismic data set by exploiting its redundant nature that allows for accurate low-rank approximation. When near-surface weathering layers influence the data's coherency, a rank-based solution can be used for statics correction.

Rank-based methods have been applied in different areas of seismic data processing, which date back to the 1980s [11]. More recent examples include denoising [12–17], interpolation [14, 18–21], deblending [22–24] and residual statics correction [25, 26]. The common ground of these methods is that ideal seismic data in a transform domain, e.g. when organized as a matrix in midpoint-offset domain, can be approximated by a low-rank matrix, while non-ideal data, e.g. noisy or subsampled data, exhibit slowly decaying singular values. Alternative to rank is to use sparsity to compensate for short-wavelength statics. [27] applies sparsity maximization in the Fourier domain on synthetic data and in the curvelet domain on field data due to the latter's better performance. However, the curvelet transform [28] can be computationally demanding. [29] modify projection onto convex sets to compensate for residual statics during interpolation in the Fourier domain. Similarly, [30] uses phase retrieval to interpolate data affected by residual statics, but using only the amplitude spectrum and sparsity promoting regularization. To ensure sparsity, these methods may require data windowing that may affect the estimated statics. [25, 26] corrects for the statics by estimation of low-rank approximated data. Nonetheless, that method may suffer from amplitude losses, particularly at high frequencies, where low-rank approximation is more challenging, leading to erroneous amplitude versus offset (AVO) responses. To circumvent the amplitude losses and preserve the AVO response, we utilize low-rank approximation as an intermediate step in statics estimation and correction.

2.1.1. CONTRIBUTIONS

In a step towards more accurate short-wavelength statics correction, we diverge from the vertical raypath assumption in the near-surface. We propose a novel low-rank-based residual statics (LR-ReS) estimation and correction framework. The method makes use of the redundant nature of seismic data that allows for accurate approximation by low-rank matrices in the midpoint-offset domain. To estimate the statics, we cross-correlate a data set influenced by the near-surface weathering layers with its low-rank-approximated version. Since we estimate the statics and no longer the directly low-rank approximated data (as shown by [25, 26]), we preserve the AVO response. To mitigate the poor performance of low-rank approximation at high frequencies, we exploit the common statics amongst different frequency bands; we first utilize the low frequency bands, where low-rank approximation performs better than high frequency bands, to estimate the statics followed by statics correction of the full-band data. Since these statics are not accurate enough for the total bandwidth, they get updated during statics estimation when including the high frequencies. To improve the estimated statics and to alleviate the need for accurate rank selection for low-rank approximation, we implement the method in an iterative and multi-scale fashion.

The proposed method relaxes the surface-consistency assumption to estimate more accurate surface- and non-surface-consistent statics at once. In practice, multiple steps are performed, where surface-consistent residual statics are first estimated followed by a brute-force non-surface-consistent step, to account for both components. Moreover, the proposed method does not require NMO correction for short-wavelength statics estimation. Therefore, statics estimation becomes independent of errors in the NMO velocity model. Consequently, we can reduce the efforts- and time-consuming multiple passes of short-wavelength statics correction and NMO velocity estimation commonly needed for conventional methods to improve their performance. Additionally, statics estimation and correction of the total wavefield without windowing to select aligned primaries or to avoid the NMO stretch effect becomes feasible. We demonstrate that the proposed method preserves the underlying structure's kinematic and dynamic properties, while estimating accurate statics in a computationally efficient manner.

2.1.2. OUTLINE

We first provide the reader with the requirements for the success of short-wavelength statics correction with a rank-based approach. We then describe the details of our proposed method, which we apply to synthetic data and a noisy field data set affected by complex weathering layers. We show the uplift we obtain compared with conventional methods on the CMP gathers, NMO velocity semblance, stack, AVO analysis and automatic horizon picking. After that, we discuss the results within the context of NMO velocity estimation and data windowing, the method's practical aspects, its computational efficiency, improvement by mitigation of the noise effect, and further extensions and applications.

2.2. RANK-BASED PROCESSING PRINCIPLES

Rapid variations in surface-elevation, base of the weathering layers and the weathering layers' velocity affect the data with short-wavelength statics [2]. As a result, the coherency of the data will decrease to result in slowly decaying singular values compared to the statics-free situation. To exploit the data's redundancy and correct for the weathering layers' effect, we need to estimate a low-rank matrix given that the desired statics-free data are of low-rank nature. With further elaboration, we define the main principles required to correct for short-wavelength statics with a low-rank-based approach.

2.2.1. SIMULATED DATA

For the demonstration, we use synthetic data modelled with acoustic finite difference modelling [31]. The data's source and receiver intervals are 10 m and its maximum offset is 4 km. Figure 4.1 shows the velocity model we use to simulate the data displayed in the time and frequency domains in Figures 2.2(a, c, e) and 2.3(a – d), respectively. To add short-wavelength statics, we shift each trace in each shot and receiver gather by up to ± 52 ms, which is considerably larger than the usual residual statics (Figures 2.2(b, d, f) and 2.4(a, b)). To mimic a realistic scenario, we design the statics to contain surface-consistent elements, where all traces recorded at the same shot or receiver location are assigned the same statics (up to ± 40 ms), and non-surface-consistent elements (up to ± 20 ms).

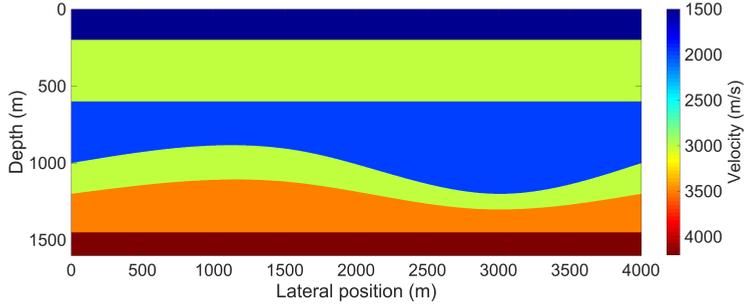


Figure 2.1: The velocity model used for synthetic data simulation.

2.2.2. LOW-RANK STRUCTURE

Seismic data exhibit redundancy that can be exploited through their rank structure in the midpoint-offset (m-h) domain. Each midpoint m and offset h can be calculated from the source s and receiver r coordinates by:

$$m = \frac{s + r}{2} \quad (2.1a)$$

and

$$h = s - r, \quad (2.1b)$$

respectively. In the midpoint-offset domain, multiple raypaths sample the same CMP with different offsets (Figures 2.2(d), 2.2(e), 2.3(b) and 2.3(d)). Moreover, it rotates the strong energy along the diagonal in the source-receiver (s-r) domain to the near-offset columns (Figure 2.3). Therefore, we expect the structure of statics-free data in the midpoint-offset domain to be of low-rank nature, i.e. low-rank structure corresponds to matrices that can be approximated by a low-rank matrix. The approximation can be achieved when the singular values decay rapidly, which enables their truncation. Given a constant frequency slice $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$ with complex entries of midpoints and offsets, where n_m and n_h correspond to the number of midpoints and offsets, respectively, we can compute its orthogonal decomposition by singular value decomposition (SVD) [32]:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \quad (2.2)$$

where H denotes the Hermitian transpose, $\mathbf{U} \in \mathbb{C}^{n_m \times k}$ and $\mathbf{V} \in \mathbb{C}^{n_h \times k}$ are the orthogonal matrices holding the left and right singular vectors, respectively. The block diagonal matrix $\mathbf{S} \in \mathbb{R}^{k \times k}$ contains the non-negative real-valued singular values, such that $\mathbf{S} = \text{diag}(s_1, s_2, s_3, \dots, s_k)$, where $s_1 \geq s_2 \geq s_3 \geq \dots \geq s_k \geq 0$ and $k = \min\{n_m, n_h\}$ is the rank of the frequency slice. For the matrix to exhibit a low-rank structure, the condition $k \ll \min\{n_m, n_h\}$ must be satisfied such that the singular values decay rapidly so they can be truncated for small k .

The low-rank structure can be exploited in different domains, where the domain of choice and subsequently the performance of the method depend on how rapid the singular values of statics-free data decay. In addition to the potential midpoint-offset transform

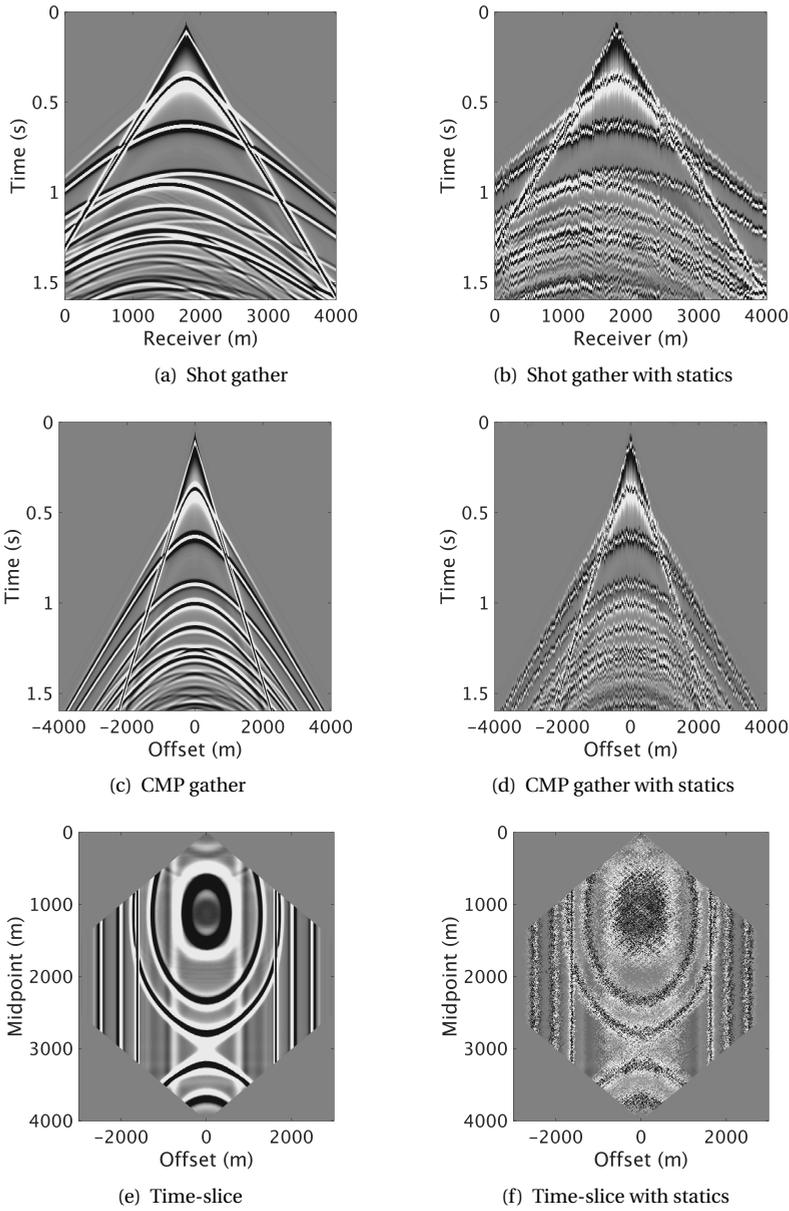


Figure 2.2: Simulated (a, c, e) statics-free data and (b, d, f) data after application of up to ± 52 ms of surface- and non-surface consistent statics. The Figures contain (a, b) shot gathers, (c, d) CMP gathers and (e, f) time-slices at 1.0 s in the midpoint-offset domain.

domain, we examine the acquisition (source-receiver) domain. We consider monochro-

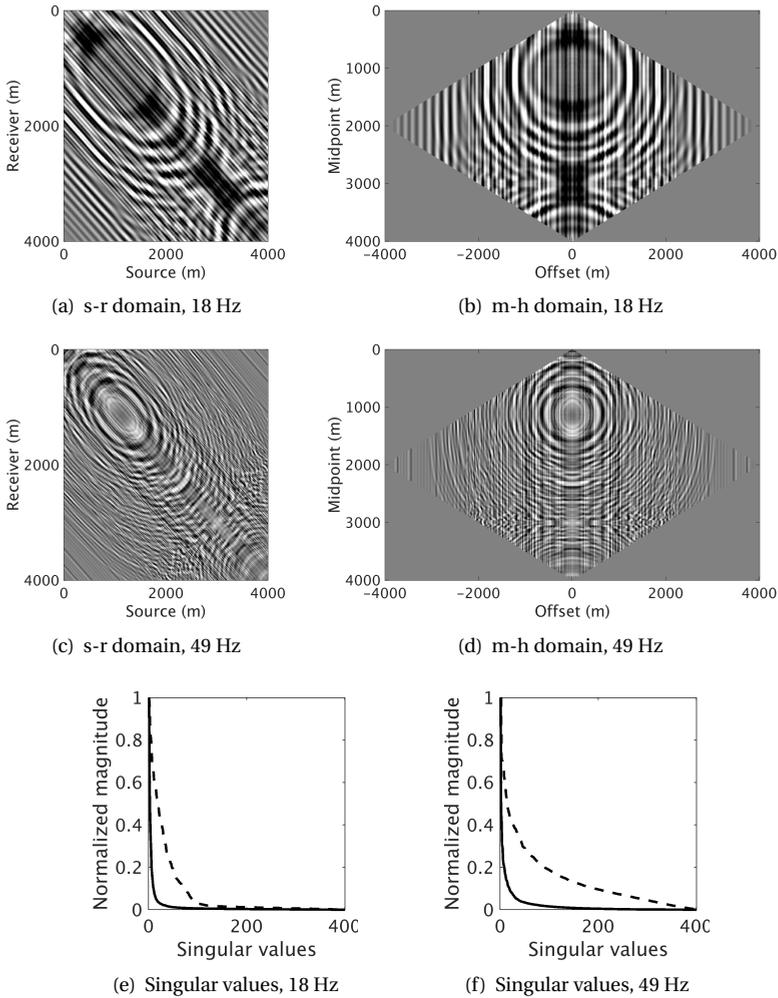


Figure 2.3: Statics-free frequency slices (real part) and their singular values decay curves extracted from statics-free data at (a, b, e) 18 Hz and (c, d, f) 49 Hz in the (a, c) source-receiver domain and (b, d) midpoint-offset domain. Dashed lines in (e, f) the singular values decay curves indicate the source-receiver domain, while solid lines correspond to the midpoint-offset domain.

matic frequency slices rather than time slices. The reason is that time-slices contain more variability as they encompass all the frequency ranges, which makes their low-rank approximation more difficult. On the contrary, low-rank approximation of low frequency slices can be achieved with high accuracy. From the modelled statics-free data, we select frequency slices spanning relatively low- to high-frequencies and examine the decay of their singular values in the source-receiver and midpoint-offset domains (Figure 2.3). As

expected, we observe that the singular values in the midpoint-offset domain decay more rapidly compared to those in the acquisition domain. This makes the former a better potential transform domain as it better satisfies the requirements of the first principle. Additional analysis on the singular values behaviour of data affected by the near-surface weathering layers is required to confirm the potential of this transform domain.

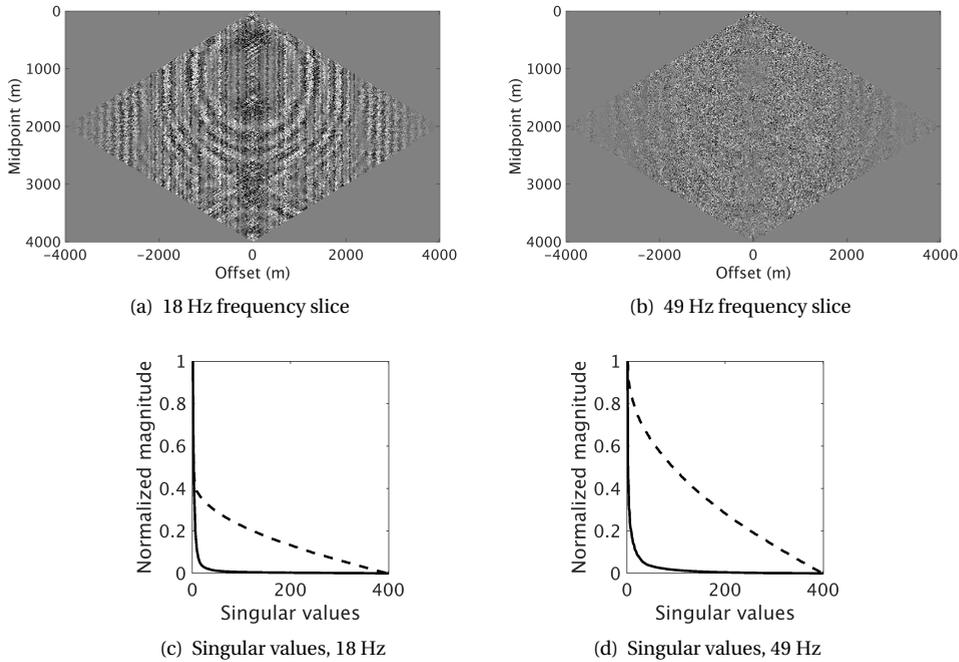


Figure 2.4: Frequency slices of synthetic data affected by up to ± 52 ms of short-wavelength statics at (a) 18 Hz and (b) 49 Hz in the midpoint-offset domain along with their singular values decay curves (c) and (d), respectively, in dashed lines. They are slowly decaying compared to those of statics-free data (Figures 2.3(b) and 2.3(d)) repeated here in solid lines.

2.2.3. STRUCTURE DESTRUCTION

We rely on rapid variations in the near-surface weathering layers to destroy the low-rank structure. Due to short-wavelength statics, the data become less coherent, which leads to slowly decaying singular values. To analyze the resultant rank structure of data affected by the statics, we transform the data from the source-receiver domain to the midpoint-offset domain, where we can exploit the data's redundancy and compute the singular values of two frequency slices at relatively low and high frequencies (Figure 2.4). We observe that they are slower than those of statics-free data, which is more visible at higher frequencies as short-wavelength statics affect them more than the low frequencies. Since we satisfy the first and second principles using the midpoint-offset domain, we proceed with low-rank promotion to correct for short-wavelength statics.

2.2.4. STRUCTURE PROMOTION

Due to the near-surface effects, the singular values become slowly decaying as it renders coherent energy incoherent. Owing to the inherent redundancy of seismic data in the midpoint-offset-frequency domain, the largest singular values preserve the coherent energy, while the lower ones are related to the incoherency. Therefore, low-rank approximation is one way to promote the low-rank structure to obtain a coherent signal without the near-surface imprint. Given a midpoint-offset frequency slice $\mathbf{Y} \in \mathbb{C}^{n_m \times n_h}$ selected from observed data, we can obtain a low-rank approximated matrix $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$ by solving the following rank-minimization problem:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X} - \mathbf{Y}\|_F \quad \text{subject to} \quad \text{rank}(\mathbf{X}) \leq k, \quad (2.3a)$$

where

$$\|\mathbf{X}\|_F = \sqrt{\sum_{i=1}^{n_m} \sum_{j=1}^{n_h} \mathbf{X}^{(ij)^2}} \quad (2.3b)$$

is the Frobenius norm that is equivalent to the ℓ_2 norm of a vector and $\mathbf{X}^{(ij)}$ are the ij -elements of \mathbf{X} . By imposing the constraint $k \ll \min\{n_m, n_h\}$, we restrict the solution \mathbf{X} to be of low-rank nature, which can be found through the SVD [33]. [13] suppress random noise with eigenimage and Cadzow filtering, where the former is more applicable to 3D stacked traces [34]. [15] uses SVD to attenuate swell noise. The method proposed by [25, 26], which requires NMO velocity estimation suggests that by selecting a small number of singular vectors corresponding to the few largest singular values, one can obtain an estimate of \mathbf{Y} without the imprint of short-wavelength statics. Since we have shown that the rank-based processing principles can be satisfied without NMO correction, we modify that method to correct for residual statics with low-rank approximation (Algorithm 2.1).

Algorithm 2.1: Low-rank approximation

Input: $\mathbf{D}_{\text{sr}} \in \mathbb{R}^{n_t \times n_r \times n_s}$ and $\mathbf{k} \in \mathbb{R}^{n_f}$

Output: $\hat{\mathbf{D}}_{\text{lr}} \in \mathbb{R}^{n_t \times n_h \times n_m}$

- 1 Transform \mathbf{D}_{sr} to $\mathbf{D}_{\text{mh}} \in \mathbb{R}^{n_t \times n_h \times n_m}$ with equation 2.1
 - 2 fft \mathbf{D}_{mh} to $\tilde{\mathbf{D}}_{\text{mh}} \in \mathbb{C}^{n_m \times n_h \times n_f}$
 - 3 **for** $f \leftarrow 1$ **to** n_f **do**
 - 4 Calculate SVD for $\tilde{\mathbf{D}}_{\text{mh}}^{(f)}$ with equation 2.2
 - 5 $\tilde{\mathbf{D}}_{\text{lr}}^{(f)} \leftarrow \sum_{j=1}^{k^{(f)}} s^{(j)} \mathbf{u}^{(j)} \mathbf{v}^{(j)H}$
 - 6 iff $\tilde{\mathbf{D}}_{\text{lr}}^{(f)}$ to $\hat{\mathbf{D}}_{\text{lr}}$
-

The input to Algorithm 2.1 is data \mathbf{D}_{sr} in the time-source-receiver domain, where n_t , n_s and n_r correspond to the number of time-, source- and receiver-samples, (Figure 2.2(b)) and rank k for each frequency slice. Throughout the algorithm, the subscripts and superscripts sr , mh , lr , $\tilde{}$ and $\hat{}$ indicate source-receiver domain, midpoint-offset domain, low-rank approximated data, frequency domain and output estimated data, respectively. We first transform the data \mathbf{D}_{sr} from the source-receiver domain to the midpoint-offset domain, where we can exploit the data's redundancy to obtain \mathbf{D}_{mh} (Figure 2.2(d)). We then Fourier

transform the data to the frequency domain to obtain $\tilde{\mathbf{D}}_{\text{mh}}$ (Figures 2.4(a) and 2.4(b)) as indicated by the second step of the algorithm. After that, there is a loop over frequencies (step 3), where the singular values of frequency slices are calculated (step 4). In the fifth step, where $\mathbf{u}^{(j)}$ and $\mathbf{v}^{(j)}$ correspond to the j^{th} columns of the matrices \mathbf{U} and \mathbf{V} of equation 2.2, respectively, data below the rank threshold level k are neglected (Figures 2.5(c) and 2.5(d)). $\tilde{\mathbf{D}}_{\text{lr}} \in \mathbb{C}^{n_m \times n_h \times n_f}$ is the estimated low-rank data from all the frequencies in the midpoint-offset domain, which can be inverse Fourier transformed (ifft) to obtain the low-rank approximated data in the time domain $\hat{\mathbf{D}}_{\text{lr}}$ as indicated by the sixth step (Figures 2.5(a) and 2.5(b)).

At low frequencies, low-rank approximation can reduce the effect of statics as shown by the time- and low-frequency-slice in Figures 2.5(b) and 2.5(c), respectively, compared to those of data with statics (Figures 2.2(f) and 2.4(a)). However, there is dimming in the data's amplitude after low-rank approximation. The second hurdle is that at high frequencies, the data contain more variability while the influence of the near-surface can be stronger (Figure 2.4). Consequently, the performance of low-rank approximation deteriorates, resulting in noisy data (Figures 2.5(a) and 2.5(d)). To avoid low-rank approximation errors at low and high frequencies, we propose the following iterative and multi-scale framework.

2.3. LR-RES ESTIMATION AND CORRECTION

Our proposed method utilizes the properties of the midpoint-offset domain, where short-wavelength statics lead to slowly decaying singular values, while statics-free frequency slices are of low-rank nature (Figures 2.4(c) and 2.4(d)). By imposing the low-rank constraint, we can obtain frequency slices with less statics imprint. Although they may contain amplitude losses due to inaccurate low-rank approximation, we can use the low-rank approximation as an intermediate step to estimate the statics. To tackle the poor performance of low-rank approximation at high frequencies, we utilize its better performance at low frequencies and exploit the similarity in the statics' influence amongst multiple frequency bands. For improved statics estimation and to alleviate the need for accurate rank selection, which is required for low-rank approximation, we implement the framework in an iterative and multi-scale fashion. The proposed method uses parts of Algorithm 2.1 as building blocks for short-wavelength statics estimation and correction as detailed in Algorithm 2.2.

The LR-ReS estimation and correction algorithm contains the same pre-processing steps of Algorithm 2.1 (steps 1 and 2). Similarly, Algorithm 2.2 includes a loop over frequency slices, where f_{min} and f_{max} are, respectively, the minimum and maximum frequencies we loop over (step 4). As it's the case for Algorithm 2.1, Algorithm 2.2 involves calculation of the singular values and approximation of low-rank data as indicated by steps 5 and 6, respectively. Figures 2.6(b) and 2.6(e) show that low-rank approximation at low frequencies is capable of reducing the imprint of statics compared to the low-pass filtered data affected by statics (Figures 2.6(a) and 2.6(d)). At the same time, the figures demonstrate that low-rank approximation results in reduction of the data's amplitude due to neglecting data of importance. To preserve the AVO response, we avoid the use of the low-rank approximated data as the final solution. We rather estimate the statics \mathbf{T}_{mh} by cross-correlation of data with statics \mathbf{D}_{mh} (Figures 2.6(a) and 2.6(d)) and low-rank approx-

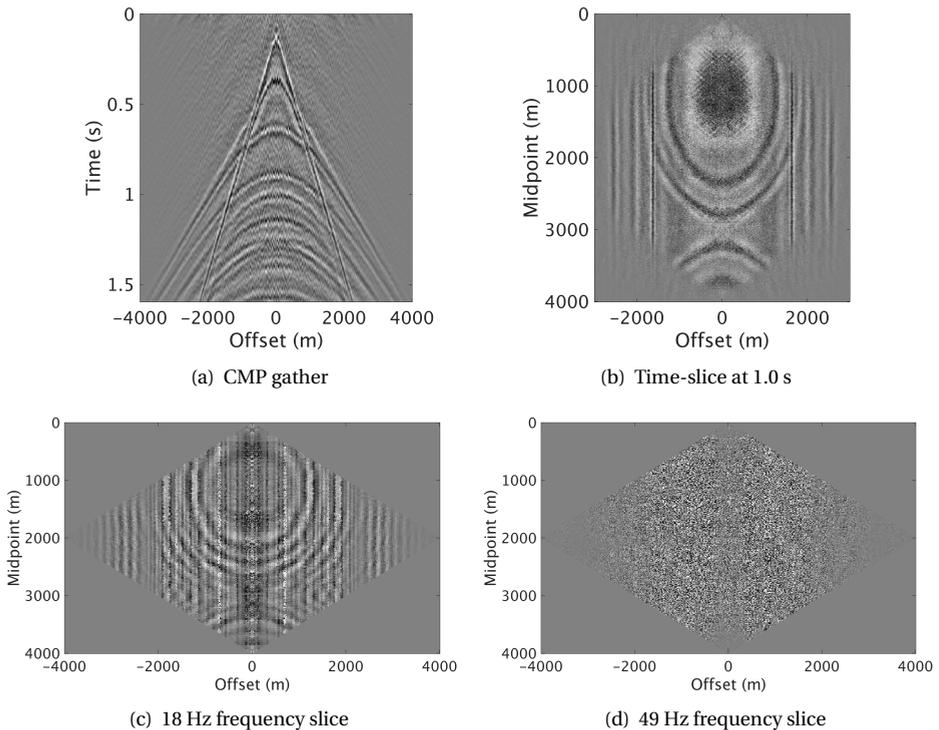


Figure 2.5: Residual statics correction with low-rank approximation (Algorithm 2.1): (a) CMP gather, (b) time-slice at 1.0 s and frequency slices at (c) 18 and (d) 49 Hz in the midpoint-offset domain.

imated data \mathbf{D}_r (Figures 2.6(b) and 2.6(e)) along the time dimension in the midpoint-offset domain (steps 7–10). Note that $\underline{\mathbf{D}}_{mh}$ (Figures 2.6(a) and 2.6(d)) is a band-pass filtered version of the data with statics \mathbf{D}_{mh} (Figures 2.2(d) and 2.2(f)). The estimated statics can then be used for statics correction of the full-band data \mathbf{D}_{mh} to obtain statics-corrected data $\hat{\mathbf{D}}_{mh}$ (step 11). Figures 2.6(c) and 2.6(f) display the data after partial statics correction with statics estimated from the low frequencies, which is already an improvement compared to the input (Figures 2.2(d) and 2.2(f)).

While low-rank approximation at low frequencies is accurate enough for statics estimation, its performance deteriorates at high frequencies (Figure 2.5). To mitigate the poor low-rank approximation performance at high frequencies, we exploit the common statics amongst multiple frequency bands \mathbf{f}_b , i.e we utilize the similarity in the statics' influence on multiple frequencies. To do so, we first estimate the statics from low frequency bands (Figures 2.6(a), 2.6(d), 2.6(b) and 2.6(e)) followed by statics correction of the full-band data (Figures 2.6(c) and 2.6(f)). By applying the statics estimated at lower frequencies to the full-band data, we already increase the redundancy of higher frequency slices since they share common statics with lower frequencies. Consequently, low-rank approx-

Algorithm 2.2: LR-Res estimation and correction

Input: $\mathbf{D}_{\text{sr}} \in \mathbb{R}^{n_t \times n_r \times n_s}$, $\mathbf{K} \in \mathbb{R}^{n_f \times n_l}$, $\mathbf{f}_b \in \mathbb{R}^{n_{f_b}}$, $f_{\min} \in \mathbb{R}$, $f_{\max} \in \mathbb{R}$
Output: $\hat{\mathbf{D}}_{\text{mh}} \in \mathbb{R}^{n_t \times n_h \times n_m \times n_l}$, $\mathbf{T}_{\text{mh}} \in \mathbb{R}^{n_h \times n_m \times n_l \times n_{f_b}}$, $\hat{\mathbf{D}}_{\text{sr}} \in \mathbb{R}^{n_t \times n_r \times n_s}$ and $\mathbf{T}_{\text{sr}} \in \mathbb{R}^{n_r \times n_s}$

- 1 Transform \mathbf{D}_{sr} to $\mathbf{D}_{\text{mh}} \in \mathbb{R}^{n_t \times n_h \times n_m}$ with equation 2.1
- 2 fft \mathbf{D}_{mh} to $\tilde{\mathbf{D}}_{\text{mh}} \in \mathbb{C}^{n_m \times n_h \times n_f}$
- 3 **for** $l \leftarrow 1$ **to** n_l **do**
- 4 **for** $f \leftarrow f_{\min}$ **to** f_{\max} **do**
- 5 Calculate SVD for $\tilde{\mathbf{D}}_{\text{mh}}^{(f)}$ with equation 2.2
- 6 $\tilde{\mathbf{D}}_{\text{lr}}^{(f)} \leftarrow \sum_{j=1}^{\mathbf{K}^{(f,l)}} s^{(j)} \mathbf{u}^{(j)} \mathbf{v}^{(j)H}$
- 7 **if** f is contained within \mathbf{f}_b **then**
- 8 $i \leftarrow$ frequency band index
- 9 ifft $\tilde{\mathbf{D}}_{\text{lr}}$ to \mathbf{D}_{lr} , ifft $\tilde{\mathbf{D}}_{\text{mh}}$ to $\underline{\mathbf{D}}_{\text{mh}}$
- 10 $\mathbf{T}_{\text{mh}}^{(l,i)} \leftarrow \mathcal{C}(\underline{\mathbf{D}}_{\text{mh}}, \mathbf{D}_{\text{lr}})$
- 11 $\hat{\mathbf{D}}_{\text{mh}}^{(l)} \leftarrow \tau(\mathbf{D}_{\text{mh}}, \mathbf{T}_{\text{mh}}^{(l,i)})$
- 12 $\mathbf{D}_{\text{mh}} \leftarrow \hat{\mathbf{D}}_{\text{mh}}^{(l)}$, fft \mathbf{D}_{mh} to $\tilde{\mathbf{D}}_{\text{mh}}$
- 13 Transform $\sum_{l=1}^{n_l} \sum_{i=1}^{n_{f_b}} \mathbf{T}_{\text{mh}}^{(l,i)}$ to \mathbf{T}_{sr}
- 14 $\hat{\mathbf{D}}_{\text{sr}} \leftarrow \tau(\mathbf{D}_{\text{sr}}, \mathbf{T}_{\text{sr}})$

imation of higher frequencies gets improved, which can be observed when comparing low-rank approximation after processing all the frequencies (Figures 2.5(a) and 2.5(b)) to that obtained after statics estimation and correction using two frequency-bands starting from low frequencies (Figures 2.7(b) and 2.7(e)). The latter figures show less statics imprint, less noise and higher frequency content when compared to the former. Since the statics estimated from low- or mid-frequency bands are not accurate enough for the high frequencies (Figures 2.6(c) and 2.6(f)), they get updated when estimating the statics at high frequency bands (Figures 2.7(c) and 2.7(f)). Compared to statics estimation and correction after processing each frequency slice, our proposed approach makes the algorithm more stable by avoiding spurious statics due to low-rank approximation errors or due to missing or corrupted frequencies, e.g. in the case of low signal-to-noise ratio.

Therefore, when the loop over frequencies reaches the desired frequency band for statics estimation (step 7), we inverse Fourier transform the data with statics $\tilde{\mathbf{D}}_{\text{mh}}$ and low-rank approximated data $\tilde{\mathbf{D}}_{\text{lr}}$ (step 9) to the time domain (Figures 2.6(a), 2.6(d) and 2.6(b), 2.6(e), respectively). Note that \mathbf{f}_b contains the maximum frequency of each frequency band. To estimate the statics, we then perform cross-correlation along the time dimension to find the lag corresponding to the largest cross-correlation coefficient, which is indicated by the operator \mathcal{C} (step 10). Let $\underline{\mathbf{d}}_{\text{mh}}$ and $\mathbf{d}_{\text{lr}} \in \mathbb{R}^{n_t}$ be traces extracted from data with statics $\underline{\mathbf{D}}_{\text{mh}}$ and low-rank approximated data \mathbf{D}_{lr} , respectively, the estimated statics

t_{mh} from these two traces can be found by:

$$t_{\text{mh}} = \underset{t}{\operatorname{argmax}}(\underline{\mathbf{d}}_{\text{mh}}(t) \star \mathbf{d}_{\text{lr}}(t)), \quad (2.4)$$

where \star indicates cross-correlation. After processing each frequency band, we apply equation 2.4 to all the traces to obtain the statics in the midpoint-offset domain $\mathbf{T}_{\text{mh}} \in \mathbb{R}^{n_m \times n_h}$. To correct for the statics, we simply shift each trace as indicated by the operator τ (step 11):

$$\hat{\mathbf{d}}_{\text{mh}} = \mathbf{d}_{\text{mh}}(t + t_{\text{mh}}), \quad (2.5)$$

where \mathbf{d}_{mh} is a trace extracted from the full-band data with statics \mathbf{D}_{mh} . Figure 2.8(a) displays the estimated statics \mathbf{T}_{mh} from all the frequency bands at the first iteration, which we use for statics-correction to obtain the estimated data (Figures 2.7(c) and 2.7(f)). The improvement can also be noticed when comparing the low and high frequency slices after LR-ReS estimation and correction at this stage (Figures 2.9(a) and 2.9(b)) to the input frequencies (Figures 2.4(a) and 2.4(b)). However, there are still unresolved statics.

To improve the statics estimation and correction, the proposed algorithm contains an iterative loop over rank scales n_l (step 3), i.e. multiple ranks for the same frequency slice, which are required for low-rank approximation (step 6). As a result, the input rank $\mathbf{K} \in \mathbb{R}^{n_f \times n_l}$ of Algorithm 2.2 becomes not only a function of frequency f as it was the case for Algorithm 2.1, but also rank-scale l . By using a multi rank-scale approach, whereby we start with a relatively high rank and reduce it further with the number of iterations, we gradually extract multi-scale time-shifts. The reason we lower the rank is that the singular values decay faster as iterations progress, which we further elaborate on in the next paragraph. Figure 2.8(a) shows that the majority of the statics are estimated at the first rank-scale iteration, which are further fine-tuned at the second and third iterations of Algorithm 2.2 (Figures 2.8(b) and 2.8(c), respectively). Even though the first-scale eliminates the bulk of the statics, which leads to considerable improvements, it is still insufficient to account for all the statics. After two multi-scale iterations, the frequency slices (Figures 2.9(c), 2.9(d), 2.9(e) and 2.9(f)), their singular values decay (Figures 2.10(b) and 2.10(d)) and time-domain data (Figures 2.11(e) and 2.11(f)) become similar to the statics-free ones (Figures 2.3(b), 2.3(d), 2.2(c) and 2.2(e)). An additional benefit of using a multi-rank-scale approach is to reduce the need of accurate low-rank approximation, i.e. if the chosen rank is inadequate to compensate for all the statics at the first iteration, fine tuning can be carried out at later iterations to improve the estimated statics and data.

We choose \mathbf{K} according to the singular-value decay relations described in the RANK-BASED PROCESSING PRINCIPLES section. Since frequency slices with short-wavelength statics exhibit low coherency, their singular values are characterized by slow decay (Figures 2.4(c) and 2.4(d)). Statics correction makes low-rank approximation more accurate as it maps incoherent energy in the tail of the singular values to coherent energy captured by the first singular values. In other words, the singular-value decay of the same frequency slice becomes more rapid after partial statics correction due to the improved coherency, which can be seen from the singular values decay curves after the first rank-scale iteration compared to those prior to statics correction (Figures 2.10(a) and 2.10(c)). Since the singular values decay rapidly after applying partial statics correction, we can use a lower rank at the next iteration to capture the coherent energy and neglect the incoherent one.

The further the iterations progress, the faster the singular values will decay, which can be seen through the comparison of Figures 2.10(a) and 2.10(c) to Figures 2.10(b) and 2.10(d). Accordingly, we start the first rank-scale iteration with a relatively higher-rank approximation and decrease it further at later iterations. This is the opposite of other rank-based methods used for example in data interpolation, which needs to fill-in missing gaps with accurate amplitudes. In that case, the smaller singular values need to be preserved as they also contain important information to obtain accurate data. On the other hand, the LR-ReS estimation and correction is after coherent energy, which can be obtained from low-rank approximated data. While a low-rank version of the data can be inaccurate (Algorithm 2.1), we avoid using it as the final solution. We mitigate the inaccuracy of low-rank approximation with statics estimation (Algorithm 2.2), i.e. we implicitly use low-rank approximation to estimate the statics (time-shifts). Therefore, we preserve the amplitudes.

When the largest singular values represent the signal of interest, low-rank approximation preserves the predominant features of the signal. As a result, cross-correlating a low-rank version of the data with its full-rank one can be used to estimate the statics that enhance the signal. On the other hand, when the data are influenced by undesired events, e.g. coherent noise or residual coherent noise, that span the largest singular values, initial rank-scale iterations can provide better statics estimation because they still preserve the signal. Since the algorithm outputs the multi-scale statics \mathbf{T}_{mh} and data $\hat{\mathbf{D}}_{\text{mh}}$, the user can quality control (QC) the results.

At the end of the iterations over frequencies and rank scales, the total midpoint-offset dependent statics \mathbf{T}_{mh} (Figure 2.8(d)) can be transformed to the source-receiver domain to obtain \mathbf{T}_{sr} (Figure 2.8(e)) as indicated by the thirteenth step of Algorithm 2.2. When compared with the actual non-surface-consistent statics originally applied on the data (Figure 2.8(f)), the error of the LR-ReS displayed in Figure 2.8(g) is minimal. The estimated statics for one shot gather is shown in Figure 2.11(h), where the error compared to the actual non-surface-consistent statics is low. By applying \mathbf{T}_{sr} on the input data \mathbf{D}_{sr} (step 14), we estimate statics-corrected data in the source-receiver domain $\hat{\mathbf{D}}_{\text{sr}}$ (Figure 2.11(d)) that are of minimal discrepancy compared to the statics-free data (Figure 2.2(a)). Since the only difference between the input and output (Figures 2.2(b) and 2.11(d), respectively) is statics correction, we are certain to preserve the AVO response. We note that if only surface-consistent statics are desired, e.g. when raypaths in the near-surface are vertical to satisfy the surface-consistency assumption, they can be obtained by averaging along rows and columns of the estimated LR-ReS in the source-receiver domain \mathbf{T}_{sr} .

With our approach, we are assuming that the estimated time-shifts are solely due to the near-surface effect. Since the dynamic component, e.g. data moveout, tends to be preserved in the largest singular values, it should not be affected. This can be illustrated when comparing, respectively, the input, output and statics-free data displayed in the source-receiver domain (Figures 2.2(b), 2.11(d) and 2.2(a)) and midpoint-offset domain (Figures 2.2(d), 2.11(e) and 2.2(e)).

2.4. RESULTS

To further illustrate the potential of our proposed LR-ReS estimation and correction, we evaluate its performance on enhancing the data, stack, NMO velocity semblance, AVO analysis and automatic horizons picking. We first show additional results of the synthetic

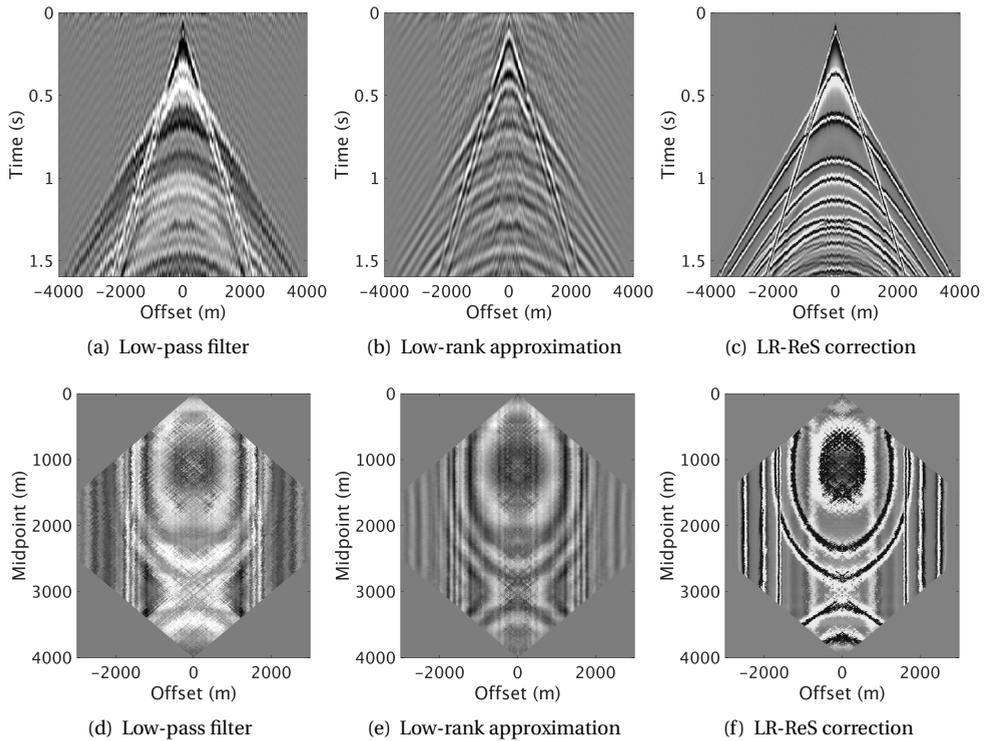


Figure 2.6: (a, b, c) CMP gathers and (d, e, f) time-slices at 1.0 s extracted from (a, d) low-pass filtered data with statics \mathbf{D}_{mh} , (b, e) low-rank approximated data \mathbf{D}_{lr} and (c, f) statics-corrected data at the first frequency-band and first rank-scale iteration.

data, followed by application to field data. The results will be compared with conventional residual statics correction.

2.4.1. SYNTHETIC DATA EXAMPLE

One of the popular short-wavelength statics correction methods is residuals statics correction with stack power maximization [3], which requires NMO velocity estimation, windowing over noise-free area containing aligned primaries and windowing to avoid the NMO stretch effect. Recently, [35] use quantum annealing to improve the convergence of stack power maximization to a global optimum. When applying the residual statics estimated with stack power maximization to the synthetic data displayed in Figures 2.2(b), 2.2(d) and 2.2(f), it provides suboptimal results because it only accounts for the surface-consistent statics (Figures 2.11(a), 2.11(b) and 2.11(c)). Figure 2.11(g) displays the estimated statics with stack power maximization and their error compared to the total non-surface-consistent statics, which is high, and to only the surface-consistent component, which is lower. This can be seen from the midpoint-offset domain time-slice (Figure 2.11(c)), where the surface-consistent statics on the diagonals at 45 degrees have been

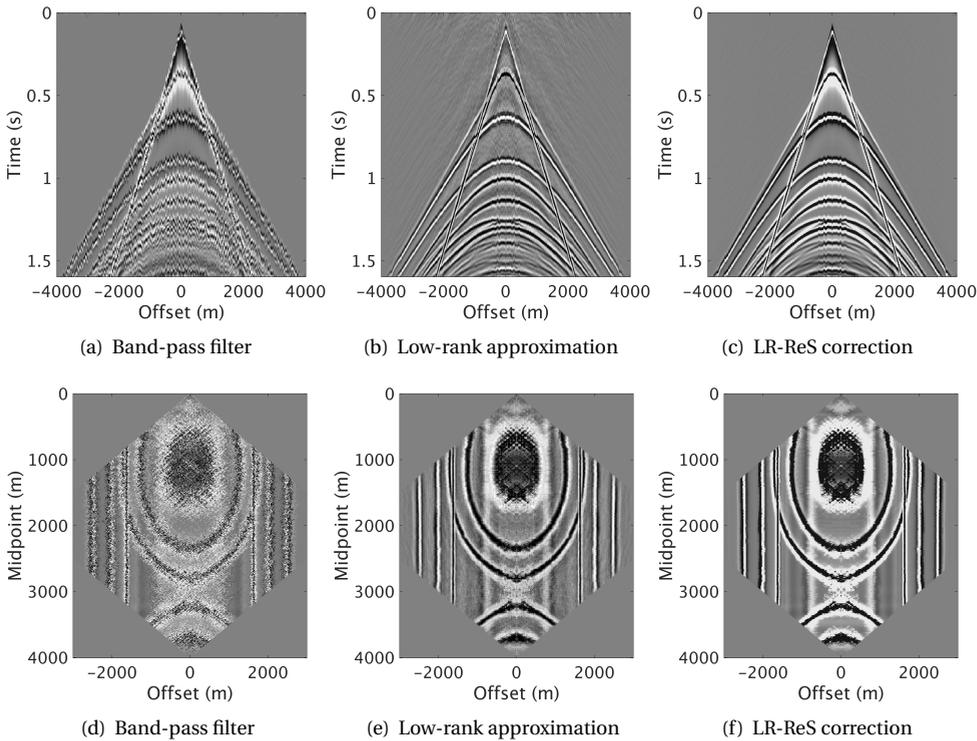


Figure 2.7: (a, b, c) CMP gathers and (d, e, f) time-slices at 1.0 s extracted from (a, d) band-pass filtered data with statics $\underline{\mathbf{D}}_{mh}$, (b, e) low-rank approximated data \mathbf{D}_{lr} and (c, f) statics-corrected data at the third frequency-band and first rank-scale iteration.

minimized compared to the time-slice of data with statics (Figure 2.2(f)). In this case and in similar situations in-practice, where surface-consistent short-wavelength statics correction methods fail to resolve all the statics, an additional non-surface-consistent solution is required. On the contrary, LR-ReS estimation and correction (Algorithm 2.2), which does not require neither NMO velocity estimation nor data windowing, is capable of accounting for the surface- and non-surface-consistent statics to provide improved results (Figures 2.11(d), 2.11(e) and 2.11(f)) that resemble the statics-free data (Figures 2.2(a), 2.2(c) and 2.2(e)).

In the presence of short-wavelength statics, NMO-corrected CMP gathers result in distorted and low-resolution subsurface structures as they stack out of phase (Figure 2.12(b)), contrary to the statics-free data (Figure 2.12(a)). Residual statics correction with stack power maximization increases the stack's resolution compared with the one of data with statics (Figure 2.12(d)). However, since it fails to fully correct for short-wavelength statics, the stack's resolution is not optimal as is evident from the computed average amplitude spectrum (Figure 2.12(e)). Using our proposed method, we obtain undistorted and higher resolution a stack that is similar to the one obtained from the statics-free data (Figure

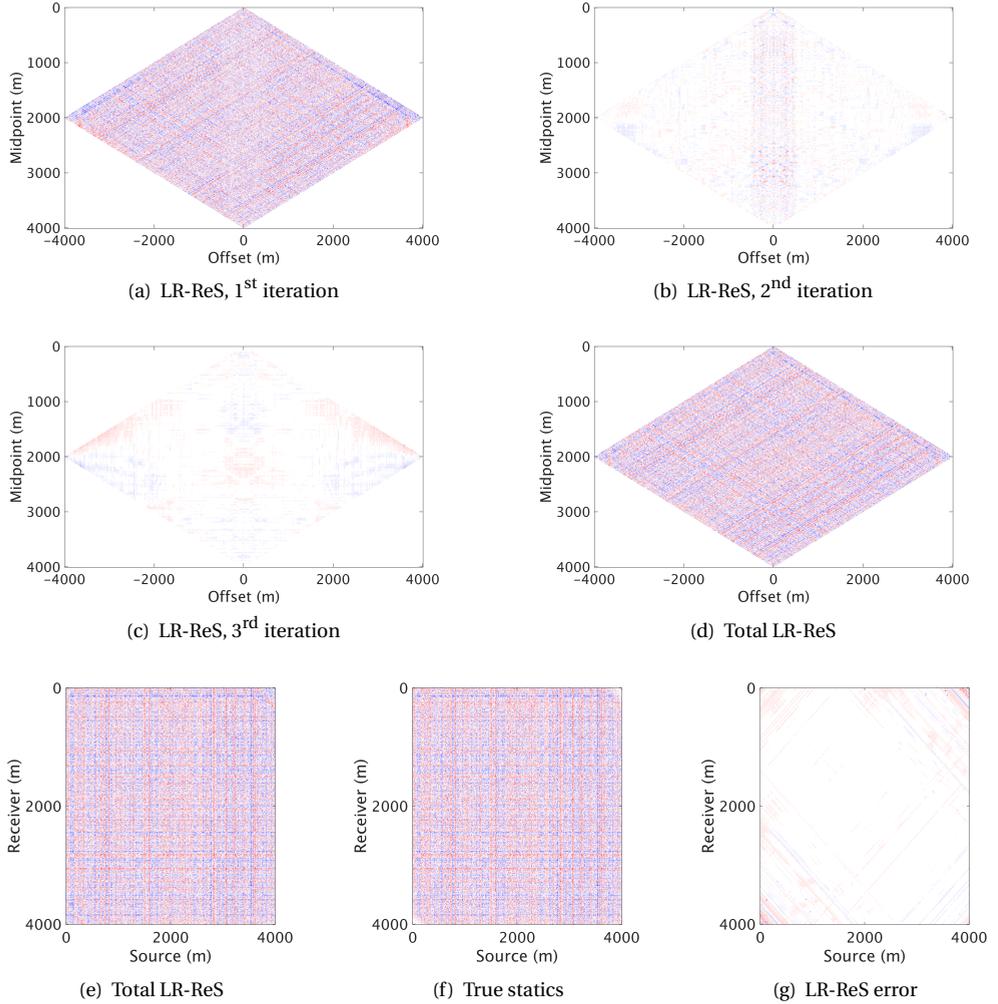


Figure 2.8: The estimated statics in the midpoint-offset domain at the (a) first, (b) second and (c) third rank-scale iterations. The total estimated statics in the (d) midpoint-offset domain and (e) source-receiver domain. (f) The true statics and (g) the LR-ReS estimation error. The Figures are clipped to ± 52 ms.

2.12(c)), which is also confirmed by the amplitude spectrum (Figure 2.12(e)). For additional quantitative analysis, we compute the average stack power p given by:

$$p = \frac{1}{n_t} \sum_{t=1}^{n_t} |s(t)|^2, \quad (2.6)$$

which corresponds to the average sum of absolute squares of a stacked trace $s(t)$ at each

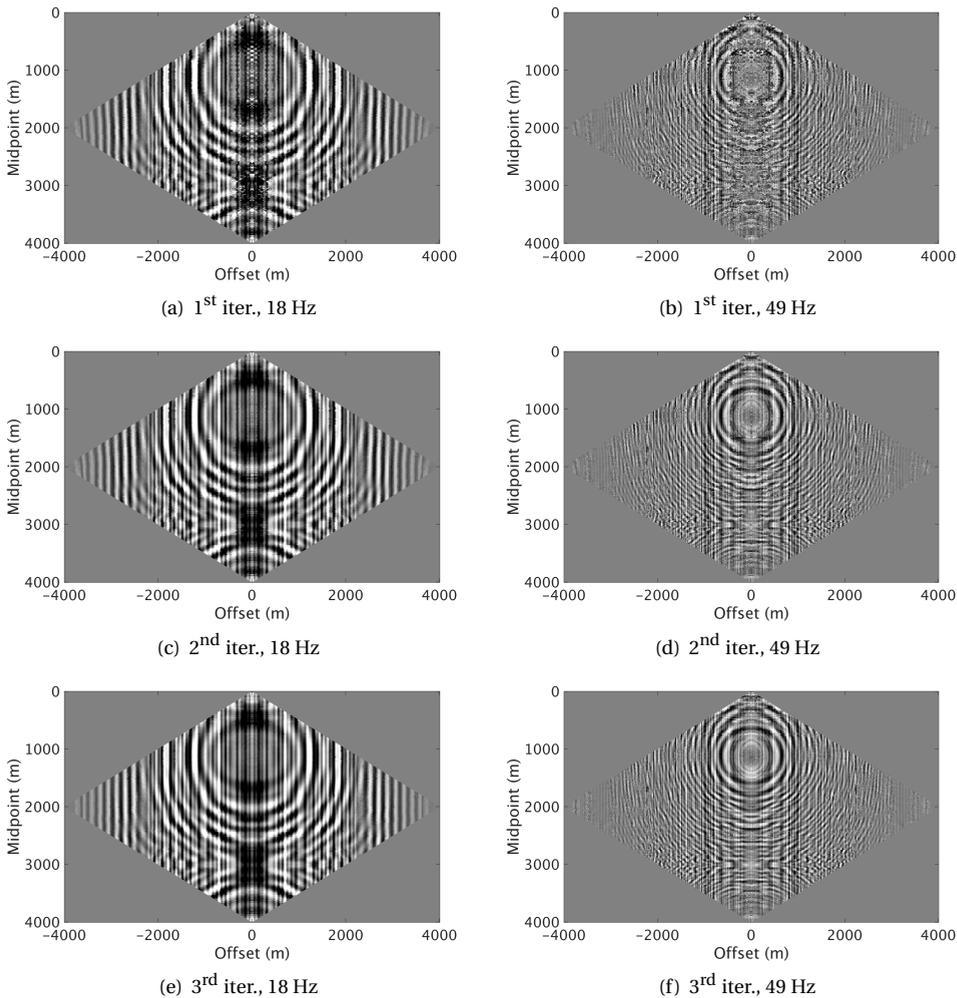


Figure 2.9: Frequency slices after LR-ReS estimation and correction at the (a, b) first, (c, d) second and (e, f) third rank-scale iterations of (a, c, e) 18 and (b, d, f) 49 Hz.

common midpoint. Relative to the statics-free stack, the average stack power of all the common midpoints of data with statics is 0.27. After residual statics correction with stack power maximization, it becomes 0.45, while it is 0.96 after our proposed method.

Quality control of residual statics can be performed by analysis of the common midpoints and their stacks, which showed that our proposed method preserves the structure and its amplitudes. Additional analyses are usually preformed on the NMO velocity semblance (see the next paragraph) and stack of the receiver or shot gathers. The latter is also used to ensure that the original structure does not change. Figure 2.13 shows the results of shot-gather stacks. The conclusions are similar to those inferred from the CMP stacks,

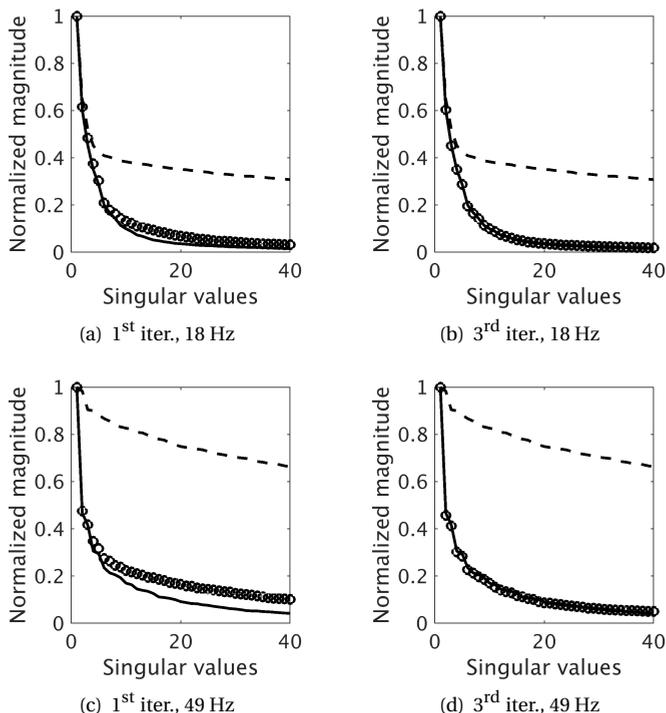


Figure 2.10: The largest singular values of (a, b) 18 and (c, d) 49 Hz frequency slices of data with statics (Figures 2.4(a) and 2.4(b)) plotted in dashed curves, statics-free data (Figures 2.3(b) and 2.3(d)) plotted in solid curves and after LR-ReS correction at the (a, c) first (Figures 2.9(a) and 2.9(b)) and (b, d) third rank-scale iterations (Figures 2.9(e) and 2.9(f)) plotted in circle markers.

where the LR-ReS do not introduce erroneous structures while accounting for the majority of the statics (Figure 2.13(c)), when compared to the statics-free case (Figure 2.13(a)). On the other hand, the surface-consistent statics estimated by stack power maximization require an additional non-surface-consistent statics correction step (Figure 2.13(d)). If the used window partly includes unaligned multiples, stack power maximization introduces erroneous structures (Figure 2.13(e)), which demonstrates its dependence on optimal window selection and accurate NMO velocity estimation.

Another important role for short-wavelength statics correction is improving the NMO velocity estimation, where both steps are usually applied in a flip-flop mode to improve their performance [2]. The semblance after LR-ReS estimation and correction, which is independent of the NMO velocity model, leads to less ambiguity and more confidence in picking the velocity similar to the statics-free situation (Figures 2.14(a) and 2.14(c), respectively). Due to the influence of the statics, the resolution of the semblances obtained from data with statics and after stack power maximization residual statics correction are suboptimal, which can result in ambiguity during the velocity picking process (Figures 2.14(b)

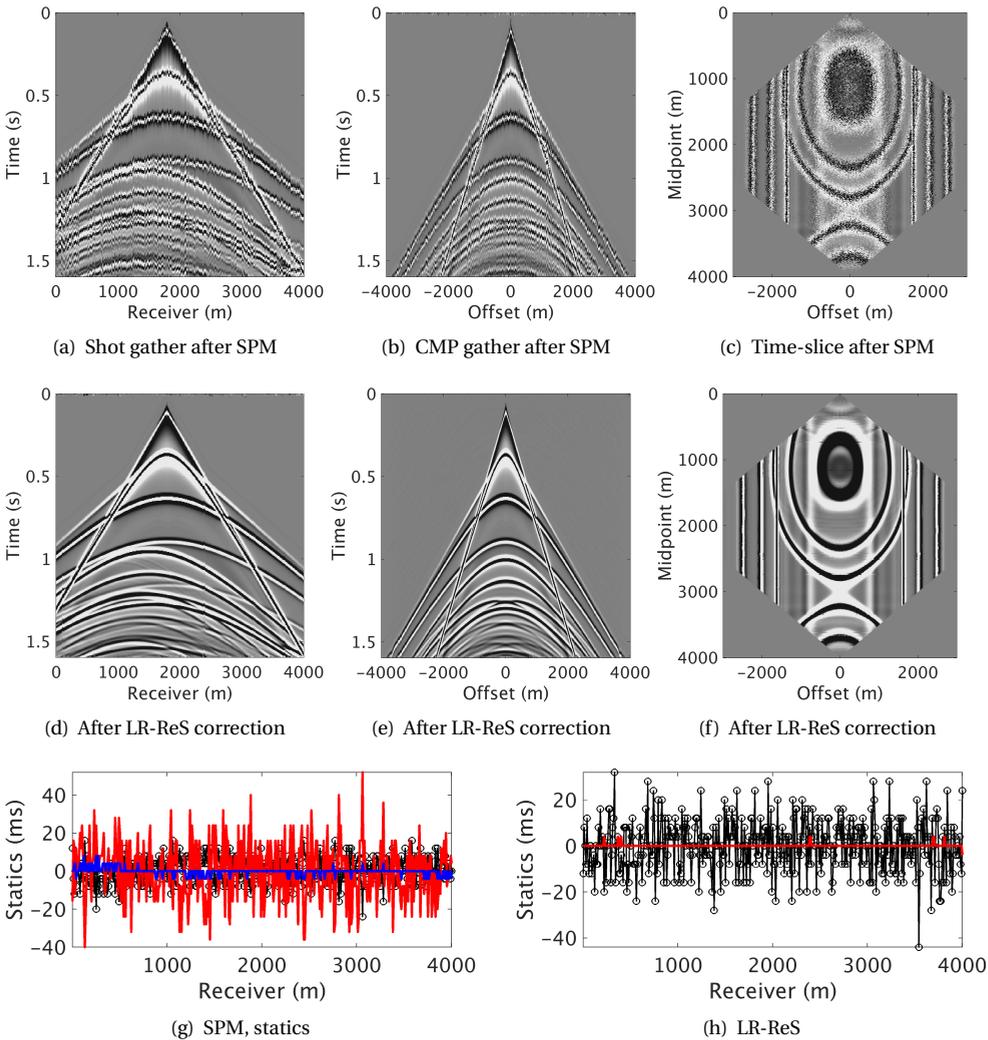


Figure 2.11: The estimated (a, d) shot gathers, (b, e) CMP gathers, (c, f) time-slices in the midpoint-offset domain and (g, h) source statics of (a, d) plotted in black along with their error in red after (a, b, c, g) stack power maximization (SPM) residual statics correction and (d, e, f, h) our proposed method. The blue line represent the error of SPM with respect to only the surface-consistent statics.

and 2.14(d)).

Due to the importance of amplitude fidelity, we examine whether the LR-ReS estimation and correction can preserve the AVO response. Since the only difference between the output and the input is statics correction (Algorithm 2.2), we do not expect to change the amplitude values. However, the AVO analysis can be influenced when traces are unbal-

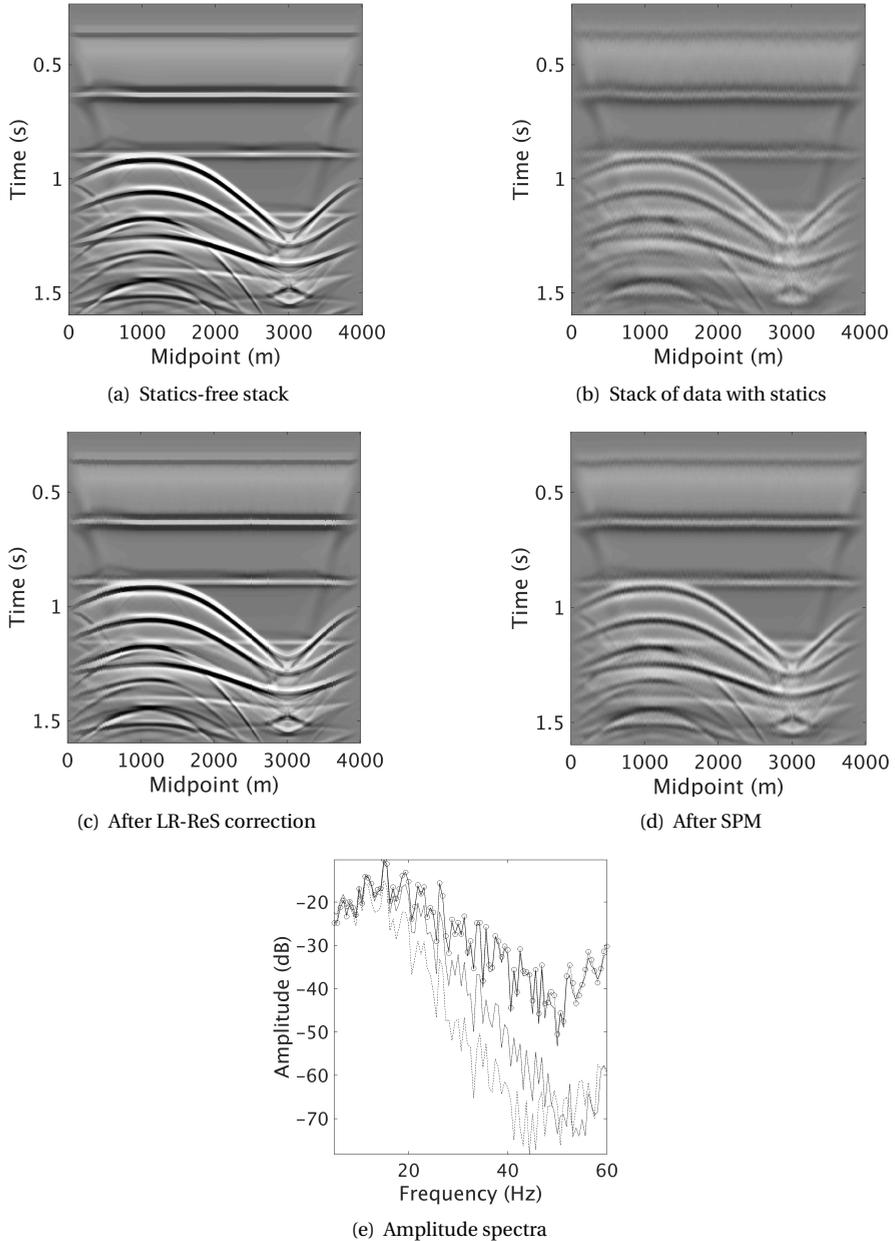


Figure 2.12: CMP stacks of (a) statics-free data, (b) data affected by short-wavelength statics and data after statics correction with (c) LR-ReS and (d) stack power maximization. (e) The average amplitude spectra computed from (a), (b), (c) and (d) drawn in dotted line with circle markers, dashed line, solid line and dash-dot line, respectively.

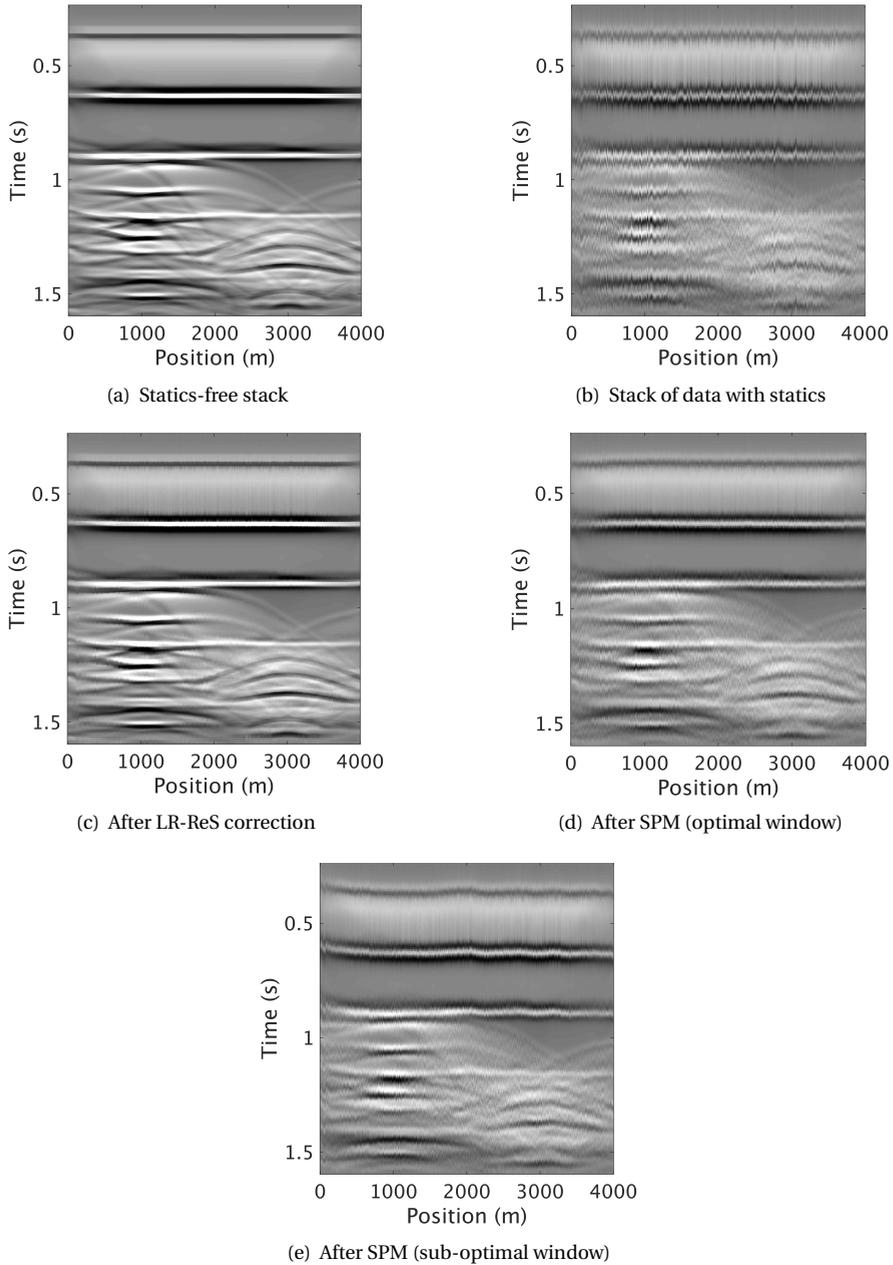


Figure 2.13: Shot gathers stacks of (a) statics-free data, (b) data affected by short-wavelength statics and data after statics correction with (c) LR-ReS and stack power maximization using (d) optimal and (e) sub-optimal window.

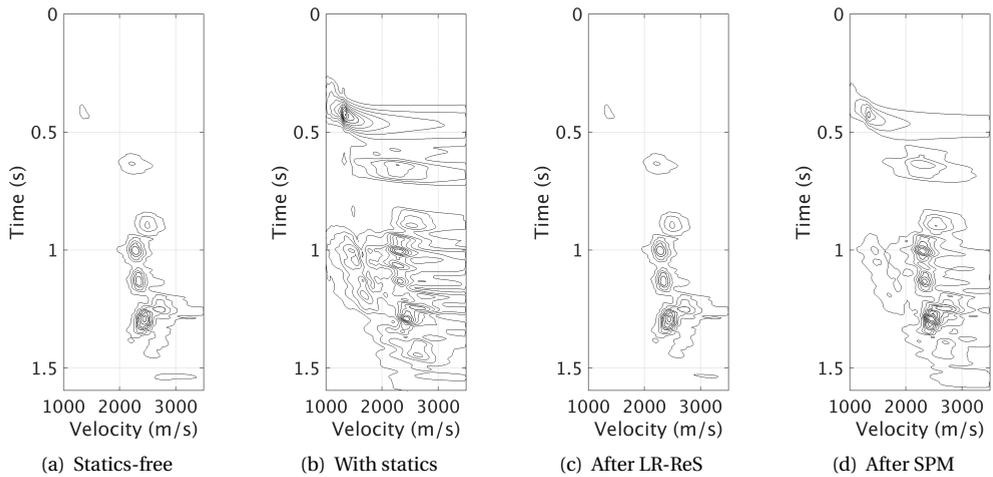


Figure 2.14: NMO velocity semblances computed from (a) statics-free data, (b) data with statics and data after statics correction with (c) our proposed method and (d) stack power maximization.

anced in the presence of residual statics [36]. This can be seen from Figure 2.15, where the auto-picked amplitudes of the horizon at 0.65 s from CMP gathers with statics and after stack power maximization suffer from erroneous picks. In contrast, the auto-picked amplitudes on the CMP gather after LR-ReS estimation and correction resembles those of statics-free data. When computing the AVO intercepts and gradients from all the CMP gathers at the same horizon, the statics-free data and the data after our proposed method provide highly similar trend confined to the same area (Figure 2.15(f)). However, the intercepts and gradients computed from CMP gathers with statics show a highly scattered trend. It becomes slightly improved after residual statics correction with stack power maximization, but still unsatisfactory due to the effect of the unresolved statics. To this extent, the LR-ReS estimation and correction framework has shown its potential in correcting for the surface- and non-surface-consistent short-wavelength statics on simulated data. To further evaluate its performance, we apply it to field data.

2.4.2. FIELD DATA EXAMPLE

The challenging field data set we consider for the demonstration is affected by complex near-surface weathering layers with rapidly varying velocities ranging between 600 and 5000 m/s [37]. The data's source- and receiver-intervals are 30 m and the maximum offset is 3,585 m. Figure 2.18(d) shows the surface elevation variation across the section. The topography also varies widely and includes gravel, loose sand, fast carbonates and karsts. These conditions lead to the low-quality data and stack shown in Figures 2.16(a) and 2.18(a), respectively. We apply minimal pre-processing on the data, which includes elevation statics correction and frequency-wavenumber (FK) filter for ground roll attenuation. Figures 2.16(a) and 2.17(a) display five of the CMP gathers after pre-processing

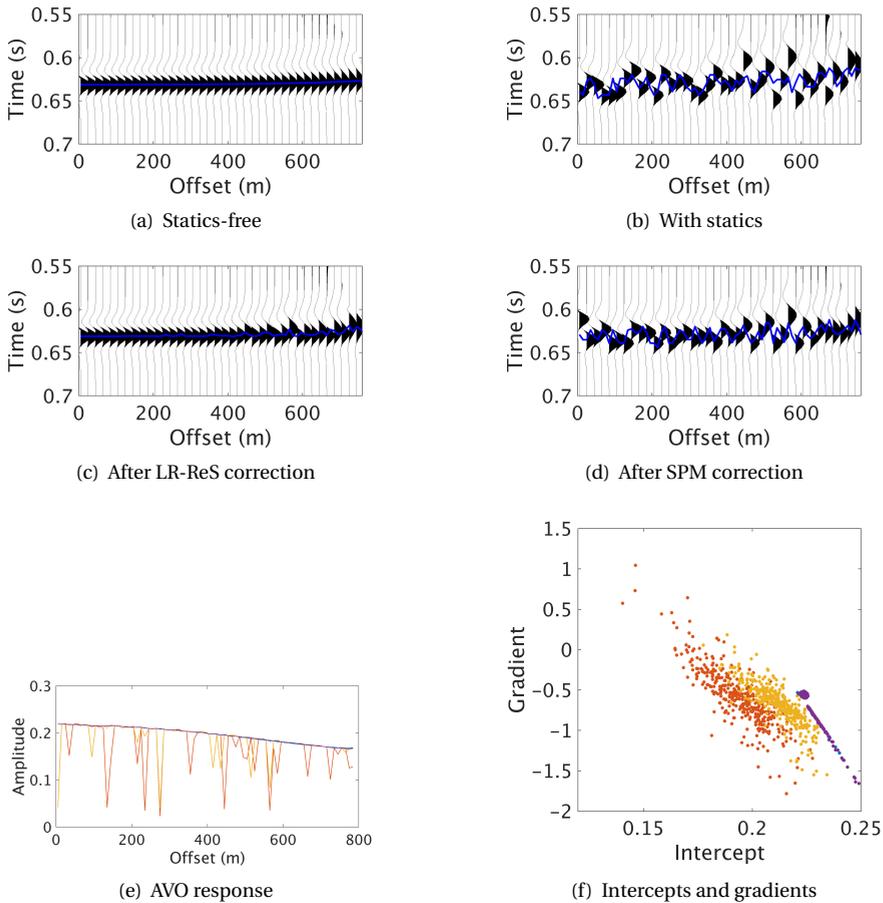


Figure 2.15: Auto-picked AVO responses in blue lines corresponding to the horizon at 0.65 s for one NMO corrected CMP gather: (a) statics-free, (b) with statics and after statics correction with (c) LR-ReS and (d) stack power maximization. (e) The AVO responses from the four CMP gathers and (f) intercepts and gradients computed from all CMP gathers at the same horizon that correspond to statics-free data (blue), data with statics (red) and data after statics correction with LR-ReS (magenta) and stack power maximization (yellow).

and after NMO correction, respectively, which clearly show the near-surface effect on the noncontinuous reflections. The data also contain near-offset noise, random noise and residual ground roll. Unfortunately, residual statics correction with stack power maximization results in minimal improvement (Figure 2.17(c)). The limited performance can be attributed to violation of the surface-consistency assumption due to the complexity of the near-surface weathering layers that can lead to non-vertical raypaths in the near-surface. To overcome the limitations of conventional residual statics correction, we apply our proposed method to this challenging field dataset.

By using the proposed LR-ReS estimation and correction framework, we avoid the assumption that requires raypaths in the near-surface to be vertical. This makes the method suitable for this data given the complexity of the near-surface. As before, we use three-rank scale iterations and three frequency bands within the available frequency content (10.5 to 58 Hz). Figure 2.16(b) displays the CMP gathers after LR-ReS estimation and correction, which show improved reflections with less near-surface imprint compared to the input data (Figure 2.16(a)). The improvement is also noticeable after NMO correction of the CMP gathers obtained after LR-ReS estimation and correction (Figure 2.17(b)) when compared to the CMP gathers with statics (Figure 2.17(a)) and after residual statics correction with stack power maximization (Figure 2.17(c)). Note that our proposed method leads to more continuous events at the near- and far-offsets in the shallow and deep parts of the CMP gathers as it accounts for the non-surface-consistent statics. The estimated LR-ReS in the midpoint-offset domain (Figure 2.18(e)) show similar pattern to the ones obtained from the synthetic data (Figure 2.8). They are composed of surface-consistent (diagonals at 45 degrees) and non-surface-consistent statics. The nature of the statics across the line also correlates well with the variations of the elevation profile (Figure 2.18(d)).

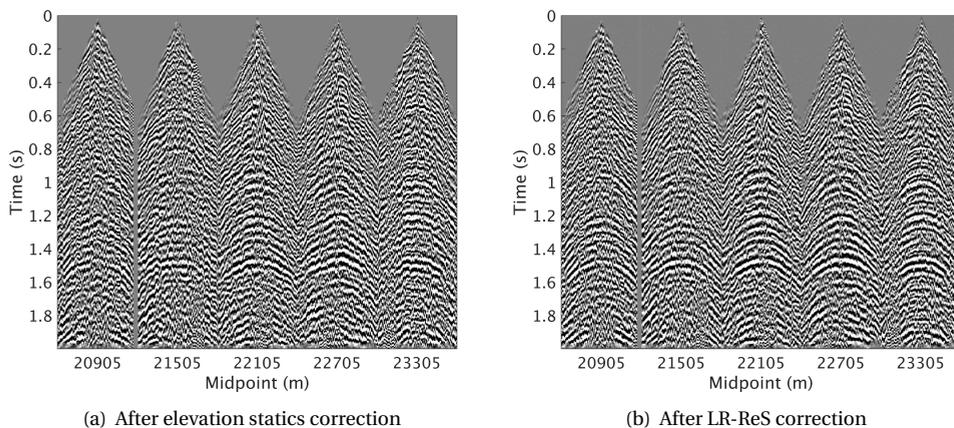


Figure 2.16: Part of the field data's CMP gathers after (a) elevation statics correction and (b) LR-ReS correction.

The complexity of the near-surface weathering layers still lead to a distorted stack after elevation statics correction (Figure 2.18(a)). It gets minor improvement after stack power maximization residual statics correction as it only accounts for surface-consistent statics (Figure 2.18(c)). Using the LR-ReS estimation and correction, we obtain a stack section with improved continuity and higher power compared to the other two sections (Figure 2.18(b)). The improvement can also be quantitatively assessed using the average stack power (equation 2.6). The power of the stack after residual statics correction with stack power maximization is 6% higher than that with elevation statics correction. In contrast, the stack power after our proposed method implemented on data without NMO correction becomes 17% higher relative to that after elevation statics correction.

Zooming in on the highlighted areas of Figure 2.18 and auto-picking the horizon be-

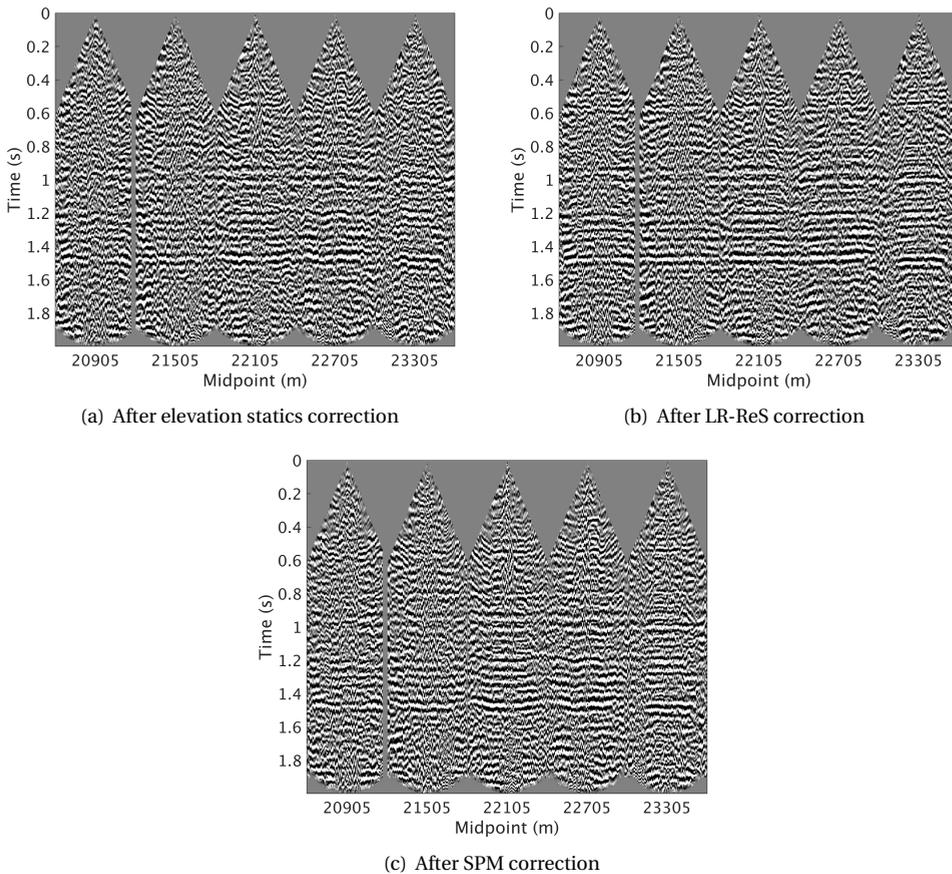


Figure 2.17: Part of the field data's NMO-corrected CMP gathers after (a) elevation statics correction, (b) LR-ReS correction and (c) stack power maximization residual statics correction.

tween 0.55 and 0.6 s confirms the improvement attained with the LR-ReS estimation and correction framework (Figure 2.19). The auto-picked horizons on the stack after elevation and conventional residual statics corrections suffer due to the imprint of residual statics that lead to erroneous picks. On the contrary, the auto-picking process was able to easily pick the horizon after the application of our proposed method, which can provide upgraded interpretation capabilities. Similar to the synthetic data, the estimated LR-ReS can enhance the NMO velocity semblance (Figure 2.20(b)), which allows for more confident velocity picking compared with the semblances computed after elevation- and conventional residual statics correction shown in Figures 2.20(a) and 2.20(c), respectively. Therefore, by estimating non-surface-consistent LR-ReS, we can obtain more accurate near-surface correction that can lead to substantial improvements.

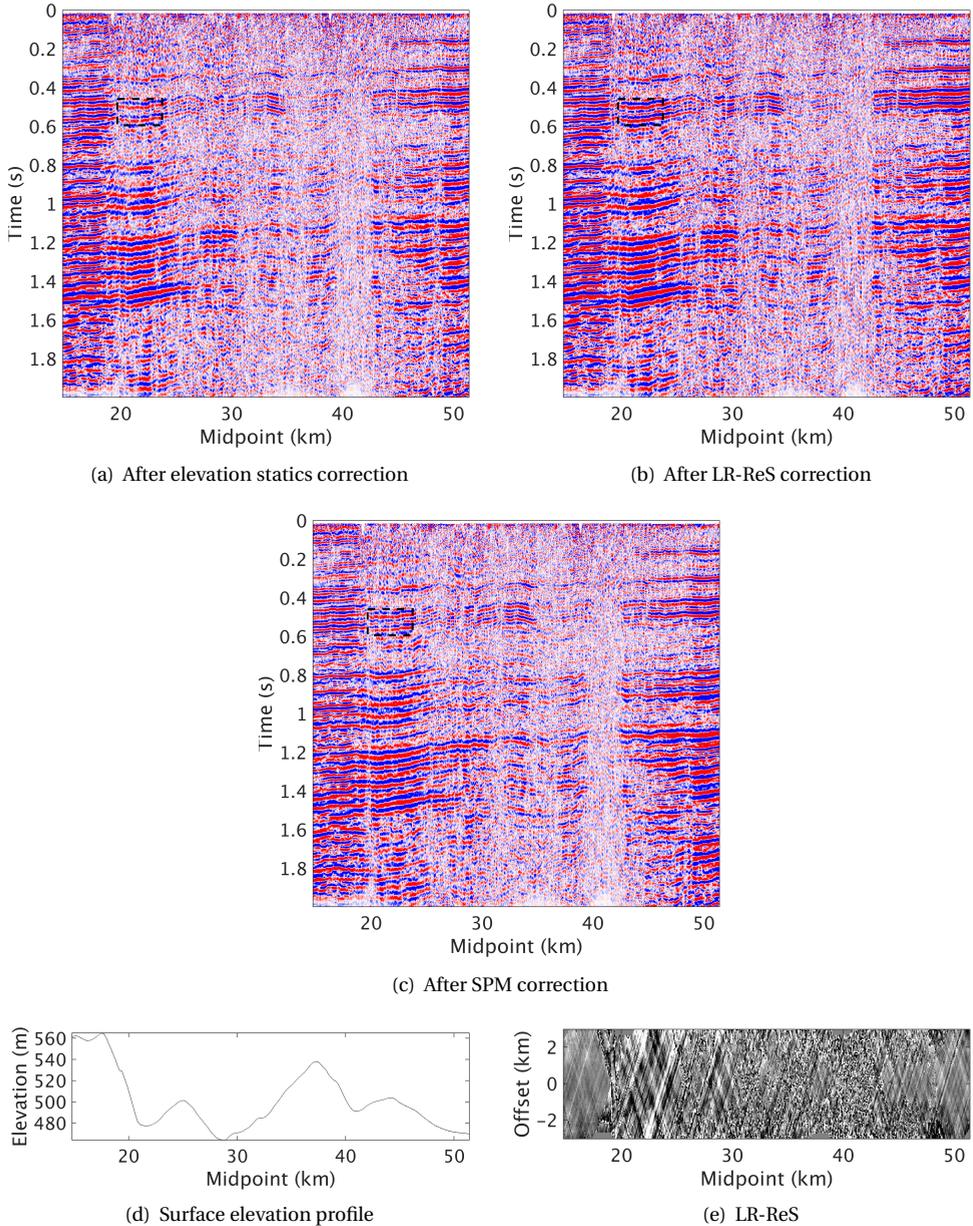


Figure 2.18: Stack sections of the field data after (a) elevation statics correction, (b) LR-ReS estimation and correction and (c) residual statics correction with stack power maximization. (d) The average surface elevation profile and (e) the total LR-ReS in the midpoint-offset domain after three rank-scale iterations clipped to ± 20 ms.

2.5. DISCUSSION

Using the LR-ReS estimation and correction, we were able to correct for the surface- and non-surface-consistent statics typically associated with rapid variations of surface-elevation, weathering layers' velocity and thickness. Removing the surface-consistency assumption can be considered dangerous for existing non-surface-consistent techniques such as trim statics, see [9]. In contrast, our proposed method captures the intrinsic relationships between the different midpoints and offsets since it operates on the whole line, i.e. we do not low-rank approximate single gathers. Moreover, LR-ReS estimation and correction does not perform any sort of temporal or spatial windowing to the data, which prevents the statics from getting biased by certain events. This is the opposite of what other conventional surface- and non-surface-consistent techniques usually do, i.e. they are dependent on data windowing, which can make them strongly biased (see Figure 2.13).

The proposed method does not require NMO corrected gathers for statics estimation. Therefore, short-wavelength statics correction becomes independent of errors in the NMO velocity model. In contrast, conventional residual statics correction can be affected by errors in the NMO velocity model, which can lead to confusion of whether statics correction, velocity refinement or both are necessary. The events at around 1.2 and 1.5 s of the field data's CMP gathers after elevation and conventional residual statics correction (Figures 2.17(a) and 2.17(b), respectively) may indicate that updating the velocity model is necessary. This can also be notice from the semblances on Figures 2.20(a) and 2.20(c). On the other hand, the CMP gathers after our proposed method, which is not biased by the NMO velocity model as it was applied to data without NMO correction, are much less affected by the surface- and non-surface-consistent statics (Figure 2.17(c)). As a result, the NMO velocity model appears to be accurate enough. Moreover, the velocity semblance after LR-ReS correction is much less affected by the weathering layers compared to that after conventional residual statics correction (Figure 2.20). Additionally, in the presence of a high-velocity layer followed by lower one, NMO velocity picking of primaries can be challenged by multiples. In certain situations, NMO velocity estimation may need to be guided by an interpreter acquainted with the geology. When the velocity semblance is affected by multiples and short-wavelength statics, NMO velocity estimation can be one of the most efforts- and time-consuming processes in land seismic data processing. On the contrary, the LR-ReS estimation is independent of the NMO velocity model, which can reduce the multiple iterations of NMO velocity estimation and short-wavelength statics correction. Therefore, it can increase the overall efficiency of data processing.

Unlike conventional methods that need access to multiple- and noise-free data with aligned primaries after NMO correction [2, 4], which may not always be feasible, the proposed method can be carried out using the total wavefield. If multiples are present in the data while the selected window for conventional residual statics correction only contains primaries, multiples may inadvertently be removed, which can limit imaging and inversion using primaries and multiples [38]. Similarly, if the selected window partly contains unaligned multiples, erroneous structures may get introduced (Figure 2.13(e)). With the LR-ReS estimation and correction, both primaries and multiples can be improved as demonstrated from the stacks in the Examples section. Moreover, the proposed method does not require horizon picking or pilot trace construction that can be necessary for non-

surface-consistent methods. In the presence of noise or non-continuous events, these tasks become nontrivial (see Figures 2.19(a) and 2.19(c)). In contrast, LR-ReS estimation and correction showed it can correct for short-wavelength statics on field data in the presence of noise. Nevertheless, Algorithm 2.2 requires selection of two main parameters, which are the ranks \mathbf{K} and frequency bands \mathbf{f}_b that we further elaborate on below.

2

2.5.1. PRACTICAL ASPECTS

We demonstrate the benefits of frequency-band-dependent and multi-rank-scale statics estimation and correction in the LR-RES ESTIMATION AND CORRECTION section. We further clarify some of the practical aspects. For the synthetic and real data, it is sufficient to estimate the statics at three frequency bands spanning low to high-frequencies such that we cover the total frequency bandwidth. One way to decide on the frequency bands is to estimate the statics after processing each third of the frequency slices. In practice, the frequency bands at which the statics are estimated will depend on the frequency content of the data set and its quality. For example, the signal to noise ratio of the field data below 10.5 Hz and above 58 Hz is low. Therefore, we estimate the statics at 25, 40 and 58 Hz.

To select the rank for both Algorithms 2.1 and 2.2, we linearly increase it with increasing frequency content. This is because higher frequencies contain more variability and therefore require higher rank for better low-rank approximation compared with the low-frequencies (Figure 2.4). We note that the rank selection also depends on the complexity of the data, which requires the user's analysis. In Algorithm 2.2, the role of low-rank approximation is to mitigate the near-surface effect, which we then utilize in cross-correlation for statics estimation. Since we estimate the statics after processing each frequency band rather than each frequency slice, we avoid errors due to low-rank approximation or poor signal-to-noise ratio. Therefore, an error due to the rank selection has less influence on the LR-ReS estimation compared to other rank-based methods that use the rank for explicit output estimation, e.g. Algorithm 2.1 or rank-minimization problems (equation 2.3a) usually used for denoising or interpolation. Moreover, using the proposed multi-rank-scale approach alleviates the requirements for accurate low-rank approximation. The strategy of starting with high-rank approximation followed by lower-rank approximation at later iterations results in improved statics estimation as explained in the LR-RES ESTIMATION AND CORRECTION section, which we demonstrate by means of synthetic and field data.

Using more rank scales to estimate the statics at additional frequency bands can add further improvements, e.g. 1 – 2 dB at certain frequencies of the stack's amplitude spectrum on the shown examples. However, the computational efficiency also depends on these two factors, namely the number of rank scales and frequency bands. The number of rank scales determines the number of SVD computes, while the frequency bands determine the number of cross-correlations. Therefore, the computational complexity of the method is determined by the SVD $O(\min\{n_m^2 \times n_h, n_m \times n_h^2\})$, which becomes $O(n_m \times n_h^2)$ as n_h is almost always less than n_m , and cross-correlations $O(n_t \times n_g)$, where n_g corresponds to the number of cross-correlation lags. For both field and synthetic data, we only compute the SVD three times for each frequency slice. We also estimate the statics at three frequency bands, which requires three cross-correlations at each rank-scale iteration. The proposed method turned out to be more computationally efficient compared to residual statics correction with stack power maximization, while at the same time providing better

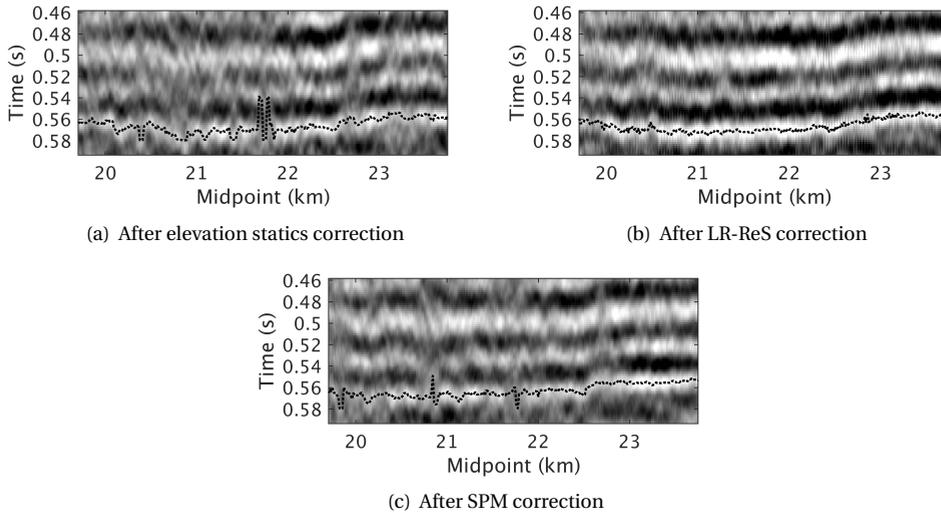


Figure 2.19: Auto picked horizons (dotted lines) on zoom of the field data's stacks at the highlighted areas of Figure 2.18: (a) before residual statics correction, (b) after LR-ReS correction and (c) after stack power maximization residual statics correction.

results. Moreover, the efforts- and time-consuming NMO velocity estimation required for conventional methods makes our proposed method even more attractive. The computational efficiency of Algorithm 2.2 can be increased by running it on multiple processors along frequency slices. For large scale matrices, e.g. in the case of 3D seismic data, SVD can be computationally demanding. Therefore, computation of only a subset of singular vectors along with their corresponding singular values may increase the computational efficiency. To further improve the performance of the proposed framework, we mitigate the effect of noise.

2.5.2. MITIGATION OF THE NOISE EFFECT

We assume that the low-rank structure destruction is due to the effect of the weathering layers, while the largest singular values preserve the signal of interest. In reality, more factors may influence the singular values such as the residual ground roll, near-offset noise and random noise of the field data (Figures 2.16(a) and 2.17(a)). The singular values of these undesired events become muddled with the larger singular values of the reflections, which influence the performance of low-rank approximation. Consequently, parts of the estimated statics contain the imprint of noise as can be seen from Figure 2.18(e). In this case, there is a need for improved pre-processing for ground roll- and noise-attenuation, especially that we only apply simple FK filter and elevation statics correction as pre-processing steps to this challenging field data.

An alternative approach to mitigate the noise effect is to apply NMO correction, which [13] use to minimize the number of dipping events and consequently use a smaller rank during denoising with Cadzow filtering. After NMO correction, aligned reflections in the

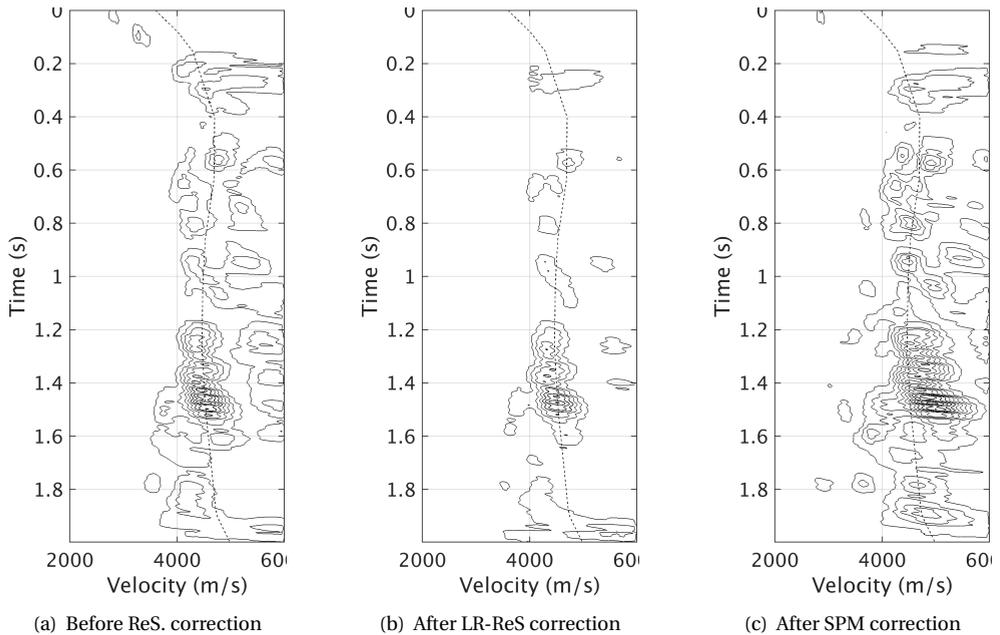


Figure 2.20: NMO velocity semblances computed from CMP gathers of the field data (a) before residual statics (ReS) correction, (b) after LR-ReS correction and (c) after stack power maximization residual statics correction. Dashed lines show the originally picked NMO velocity.

CMP gathers form the majority of the events (Figure 2.17(a)), which leads the undesired events to be less coherent and consequently exhibit the smaller singular values space. Therefore, the singular values become more rapidly decaying after NMO correction compared to those without NMO correction (Figure 2.22(a)). As a result, application of our proposed method on NMO-corrected data allows for improved low-rank approximation, which mitigates the effect of noise on the estimated statics and stacks displayed in Figure 2.21 when compared to those in Figures 2.18(e), 2.18(b) and 2.19(b). The obtained stack section and auto-picked horizon (Figure 2.21) show higher degree of improvement compared to those after elevation statics and conventional residual statics corrections (Figures 2.18(a), 2.19(a) and 2.18(c), 2.19(c) respectively). The stack power after residual statics correction with our proposed method becomes 42% higher than that of data with elevation statics correction, while the improvement with stack power maximization is only 6%. Note that to achieve the same LR-ReS estimation and correction performance of data with and without NMO correction (Figures 2.21, 2.18(b) and 2.19(b)) on this challenging field data, improved pre-processing becomes essential. Even though the data's condition and transform domain without NMO correction are not optimal, the estimated LR-ReS provide reflections with improved continuity (Figures 2.16 and 2.17). Moreover, its stack section (Figure 2.18(b)) shows better performance compared to conventional residual statics cor-

rection (Figure 2.18(c)), which requires NMO velocity estimation and correction.

In contrast to the field data, the synthetic data set is ideal, except for the added short-wavelength statics. Therefore, removing the direct wavefield and part of the reflected wavefield imposed by NMO correction can decrease the redundancy of the data and influence the singular values decay. To observe that, we examine the singular values decay prior to LR-ReS correction on data with and without NMO correction (Figure 2.22(b)). We notice that NMO correction results in slower singular values decay compared to data without NMO correction due to the applied mute. In this case, the midpoint-offset domain without NMO correction is the better one for statics estimation.

We indicate that conventional surface- and non-surface-consistent residual statics estimation methods can be influenced by noise. In this case, surface-consistent statics can provide a more stable solution compared to the non-surface-consistent ones. For the shown field data set, which is contaminated with noise, it is not necessary to enforce surface-consistency. However, if a surface-consistent solution is desired, it can be computed by averaging the estimated statics along the rows and columns in the source-receiver domain.

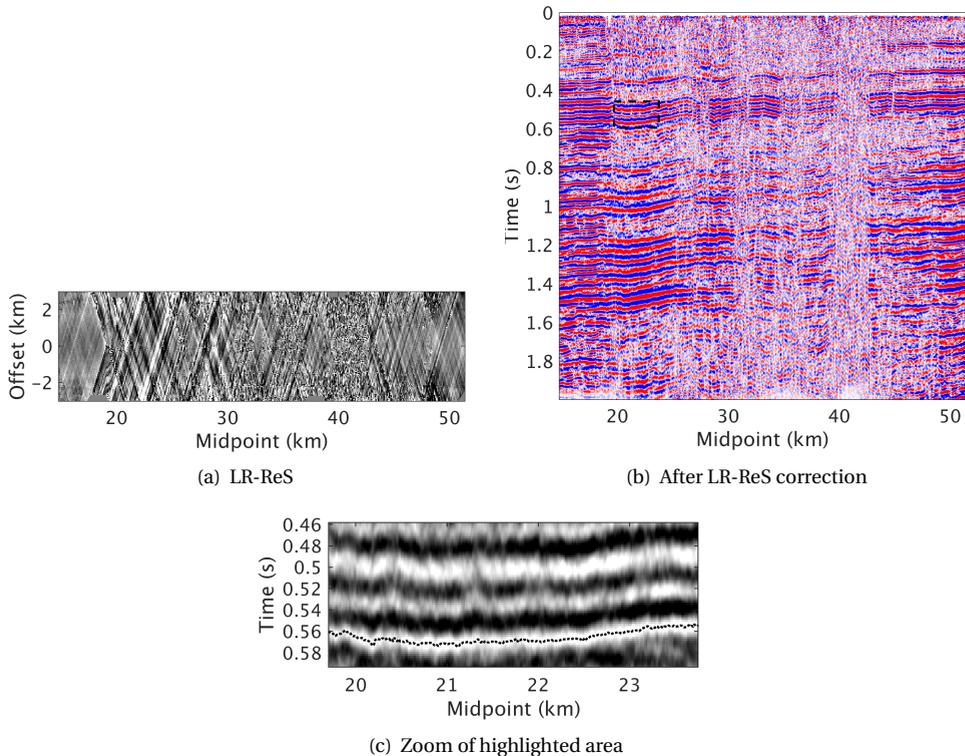


Figure 2.21: (a) The total estimated statics in the midpoint-offset domain clipped to ± 20 ms, (b) stack and (c) zoom at the highlighted area of (b) with the auto-picked horizon after LR-ReS correction of NMO corrected data.

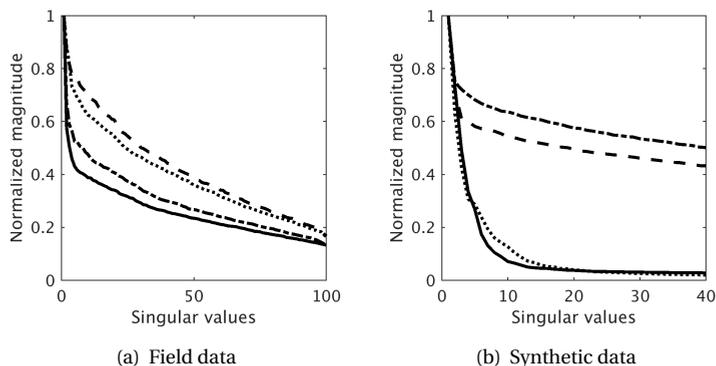


Figure 2.22: The largest singular values computed from frequency slices in the midpoint-offset domain of (a) field data and (b) synthetic data without NMO correction (dashed curves), with NMO correction (dash-dot curves) and after LR-ReS estimation and correction of data without (dotted curves) and with (solid curves) NMO correction.

2.5.3. FURTHER EXTENSIONS AND APPLICATIONS

Residual statics correction usually compensates for the short-wavelength component of the statics. Therefore, we assume that the data can be redatumed to a flat surface (roughly) using long-wavelength statics correction methods, after which residual statics correction can be applied. To apply LR-ReS estimation and correction to field data with varying topography prior to long-wavelength statics correction, further tests need to be carried out. We envision that if the near-surface renders coherent energy incoherent and the largest singular values preserve the coherent energy, LR-ReS estimation and correction can still be applicable.

As evident from the field data's stack (Figure 2.18(a)), the data set is affected by various near-surface conditions. Between 15 and 18.5 km as well as between 43 and 52 km, there is minor distortion due to the near-surface. As a result, only small LR-ReS are estimated (Figures 2.18(e) and 2.21(a)). The majority of the improvement is obtained between 18.5 and 40 km (Figures 2.18(b) and 2.21(b)). On the other hand, between 40 and 43 km, the data set is heavily distorted and exhibits no coherency due to the near-surface conditions and 3D effects not captured by the 2D data. In this case, a 3D wave-equation-based solution is necessary to overcome the limitations of elevation statics correction, which can be followed by LR-ReS estimation and correction.

The proposed LR-ReS estimation and correction framework is currently implemented on 2D data. However, it can be easily extended to 3D, where the 5D data-volume becomes parameterized by midpoints and offsets along the x- and y-dimensions (m_x, m_y, h_x, h_y). Other tensor- or matrix-based transform-revealing parameterizations can also be used. The inclusion of multi-dimensional data can further benefit the LR-ReS estimation by increasing the redundancy of the data, which can lead to better performance compared to processing separate 2D lines. However, poor sampling with coarse sources and receivers

may prevent that if the few largest singular values do not capture the coherent energy. We note that there can be a difference between the field data grid and the computational grid. In the shown field data example, we place the common midpoints at the nearest grid points to the actual field positions. After statics estimation and correction, the traces can be sorted back to the original domain. Therefore, a further point of research can be related to binning of off-the-grid data. Currently, the estimated LR- ReS are offset-variant, which can be the case when raypaths in the near-surface diverge from the normal-incidence. The statics may also become time-variant, which requires additional analysis to implement LR-ReS estimation and correction in this situation.

A potential application of LR-ReS estimation and correction is S-wave short-wavelength statics correction, as the proposed method can estimate large statics with the frequency-band-dependent approach. In addition to using residual statics correction in conventional processing workflows, it can also be used as a pre-processing step prior to full wavefield imaging and inversion such as full waveform inversion and joint migration inversion [39–41] as residual statics can be beyond the resolution of these methods. Or even better, residual statics correction should be combined together with these dynamic processes to explain the whole data and obtain more complete near- and sub-surface models. [42] have shown initial steps towards this direction. Not requiring an NMO velocity model and being able to correct for the surface- and non-surface-consistent short-wavelength statics makes the LR-Res estimation and correction more favorable compared to existing methods.

2.6. CONCLUSIONS

The surface-consistency assumption commonly used in short-wavelength statics correction can be successful in many occasions. However, when raypaths in the near-surface diverge from the normal incidence, residual statics correction with stack power maximization can be shown to fail to fully correct for short-wavelength statics. In this case, additional non-surface-consistent statics correction process is required, which adds to the computational costs and can be prone to errors in the presence of noise or when picking noncontinuous horizons. To overcome these limitations, we propose a low-rank-based residual statics estimation and correction framework that can simultaneously correct for the surface- and non-surface-consistent statics. The method utilizes the redundant nature of midpoint-offset frequency slices to iteratively estimate multi-rank-scale and frequency-band-dependent statics. The proposed LR-ReS estimation and correction can estimate short-wavelength statics without the need for NMO correction, which can reduce the efforts- and time-consuming multiple passes of NMO velocity estimation and residual statics correction. Consequently, it does not also require windowing over a noise-free area containing primaries or windowing to avoid the NMO stretch effect. Results on synthetic and field data show significant improvements that conventional residual statics correction could not achieve. Since the method can be directly applied to the total wavefield without extensive pre-processing, e.g. multiple removal and NMO velocity estimation, residual statics correction becomes feasible prior to other processing steps to improve their performance. Further potential applications include short-wavelength statics correction prior to first breaks picking, traveltimes inversion or full wavefield imaging and inversion techniques. These findings make the proposed method preferable to correct for

short-wavelength statics compared to conventional methods.

2.7. ACKNOWLEDGMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his Ph.D. studies. We also would like to thank Saudi Aramco for providing the field data.

REFERENCES

- [1] M. T. Taner, F. Koehler, and K. A. Alhilali, *Estimation and correction of near-surface time anomalies*, *Geophysics* **39** (1974), pp. 441–463, <https://doi.org/10.1190/1.1440441>.
- [2] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <https://library.seg.org/doi/pdf/10.1190/1.9781560801580>.
- [3] J. Ronen and J. F. Claerbout, *Surface-consistent residual statics estimation by stack-power maximization*, *Geophysics* **50** (1985), pp. 2759–2767, <https://doi.org/10.1190/1.1441896>.
- [4] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999).
- [5] D. Marsden, *Static corrections—a review, part II*, *The Leading Edge* **12** (1993), pp. 115–120, <https://doi.org/10.1190/1.1436936>.
- [6] R. E. Sheriff, *Encyclopedic Dictionary of Applied Geophysics, Fourth edition* (Society of Exploration Geophysicists, 2002) <https://library.seg.org/doi/pdf/10.1190/1.9781560802969>.
- [7] J. Xu, H. Zhang, and J. Zhang, *A robust surface-consistent residual phase correction method based on migrated gathers*, *Exploration Geophysics* **49** (2018), pp. 336–344, <https://doi.org/10.1071/EG17017>.
- [8] D. C. Henley, *Interferometric application of static corrections*, *Geophysics* **77** (2012), pp. Q1–Q13, <https://doi.org/10.1190/geo2011-0082.1>.
- [9] C. P. Ursenbach and J. C. Bancroft, *Playing with fire: Noise alignment in trim and residual statics*, in *SEG Technical Program Expanded Abstracts 2001* (Society of Exploration Geophysicists, 2001) pp. 1973–1976, <https://library.seg.org/doi/pdf/10.1190/1.1816525>.
- [10] A. Breuer, N. Etrich, and P. Habelitz, *Deep learning in seismic processing: Trim statics and demultiple*, in *SEG Technical Program Expanded Abstracts 2020* (Society of Exploration Geophysicists, 2020) pp. 3199–3203, <https://library.seg.org/doi/pdf/10.1190/segam2020-3427887.1>.
- [11] T. J. Ulrych, S. Freire, and P. Siston, *Eigenimage processing of seismic sections*, in *SEG Technical Program Expanded Abstracts 1988* (Society of Exploration Geophysicists, 1988) pp. 1261–1265, <https://doi.org/10.1190/1.1892508>.

- [12] M. Bekara and M. Van der Baan, *Local singular value decomposition for signal enhancement of seismic data*, *Geophysics* **72** (2007), pp. V59–V65, <https://doi.org/10.1190/1.2435967>.
- [13] S. Trickett and L. Burroughs, *Prestack rank-reduction-based noise suppression*, *CSEG Recorder* **34** (2009), pp. 24–31.
- [14] S. Trickett, L. Burroughs, A. Milton, L. Walton, and R. Dack, *Rank-reduction-based trace interpolation*, in *SEG Technical Program Expanded Abstracts 2010* (Society of Exploration Geophysicists, 2010) pp. 3829–3833, <https://doi.org/10.1190/1.3513645>.
- [15] N. Moldoveanu, *A singular value decomposition algorithm for the attenuation of high energy towed-streamer noise*, *First Break* **29** (2011), pp. 101–108, <https://doi.org/10.3997/1365-2397.29.12.55902>.
- [16] Y. Chen, Y. Zhou, W. Chen, S. Zu, W. Huang, and D. Zhang, *Empirical low-rank approximation for seismic noise attenuation*, *IEEE Transactions on Geoscience and Remote Sensing* **55** (2017), pp. 4696–4711, <http://dx.doi.org/10.1109/TGRS.2017.2698342>.
- [17] M. Zhang, Y. Liu, H. Zhang, and Y. Chen, *Incoherent noise suppression of seismic data based on robust low-rank approximation*, *IEEE Transactions on Geoscience and Remote Sensing* **58** (2020), pp. 8874–8887, <https://doi.org/10.1109/TGRS.2020.2991438>.
- [18] V. Oropeza and M. Sacchi, *Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis*, *Geophysics* **76** (2011), pp. V25–V32, <https://doi.org/10.1190/1.3552706>.
- [19] N. Kreimer and M. D. Sacchi, *A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation*, *Geophysics* **77** (2012), pp. V113–V122, <https://doi.org/10.1190/geo2011-0399.1>.
- [20] A. Aravkin, R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, *Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation*, *SIAM Journal on Scientific Computing* **36** (2014), pp. S237–S266, <https://doi.org/10.1137/130919210>.
- [21] R. Kumar, C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, *Efficient matrix completion for seismic data reconstruction*, *Geophysics* **80** (2015), pp. V97–V114.
- [22] M. Maraschini, R. Dyer, K. Stevens, and D. Bird, *Source separation by iterative rank reduction - theory and applications*, in *74th EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2012) pp. 1–5, <https://doi.org/10.3997/2214-4609.20148370>.
- [23] H. Wason, R. Kumar, F. J. Herrmann, and A. Y. Aravkin, *Source separation via SVD-free rank minimization in the hierarchical semi-separable representation*, in *SEG Technical Program Expanded Abstracts 2014* (Society of Exploration Geophysicists, 2014) pp. 120–126, [10.1190/segam2014-1583.1](https://doi.org/10.1190/segam2014-1583.1).

- [24] J. Cheng and M. D. Sacchi, *Separation and reconstruction of simultaneous source data via iterative rank reduction*, *Geophysics* **80** (2015), pp. V57–V66.
- [25] A. M. Alfaraj, R. Kumar, and F. J. Herrmann, *Automatic statics and residual statics correction with low-rank approximation*, in *80th EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2018) pp. 1–5, <https://doi.org/10.3997/2214-4609.201801107>.
- [26] A. Alfaraj, M. Almubarak, and F. Herrmann, *Correcting for short-wavelength statics with low rank approximation*, in *81st EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2019) pp. 1–5.
- [27] A. Gholami, *Residual statics estimation by sparsity maximization*, *Geophysics* **78** (2013), pp. V11–V19.
- [28] E. Candes, L. Demanet, D. Donoho, and L. Ying, *Fast discrete curvelet transforms*, *Multiscale Modeling & Simulation* **5** (2006), pp. 861–899, <https://doi.org/10.1137/05064182X>.
- [29] A. Stanton, N. Kazemi, and M. D. Sacchi, *Processing seismic data in the presence of residual statics*, in *SEG Technical Program Expanded Abstracts 2013* (Society of Exploration Geophysicists, 2013) pp. 1838–1842, <https://doi.org/10.1190/segam2013-1453.1>.
- [30] A. Gholami, *Phase retrieval through regularization for seismic problems*, *Geophysics* **79** (2014), pp. V153–V164, <https://doi.org/10.1190/geo2013-0318.1>.
- [31] J. W. Thorbecke and D. Draganov, *Finite-difference modeling experiments for seismic interferometry*, *Geophysics* **76** (2011), pp. H1–H18, <https://doi.org/10.1190/geo2010-0039.1>.
- [32] G. H. Golub and C. Reinsch, *Singular value decomposition and least squares solutions*, in *Linear algebra* (Springer, 1971) pp. 134–151.
- [33] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, *Psychometrika* **1** (1936), pp. 211–218.
- [34] S. R. Trickett, *F-xy eigenimage noise suppression*, *Geophysics* **68** (2003), pp. 751–759, <https://doi.org/10.1190/1.1567245>.
- [35] M. Dukalski, D. Rovetta, S. van der Linde, M. Möller, N. Neumann, and F. Phillipson, *Quantum computer-assisted global optimization in geophysics illustrated with stack power maximization for refraction residual statics estimation*, *Geophysics* **88** (2022), pp. 1–74.
- [36] S. Chopra and J. P. Castagna, *AVO* (Society of Exploration Geophysicists, 2014) <https://library.seg.org/doi/pdf/10.1190/1.9781560803201>.
- [37] M. Al-Ali and D. Verschuur, *An integrated method for resolving the seismic complex near-surface problem*, *Geophysical Prospecting* **54** (2006), pp. 739–750.

- [38] D. J. Verschuur, *Seismic multiple removal techniques: past, present and future (Revised edition)* (EAGE publications Houten, Netherlands, 2013).
- [39] A. Tarantola, *Inversion of seismic reflection data in the acoustic approximation*, *Geophysics* **49** (1984), pp. 1259–1266, <https://doi.org/10.1190/1.1441754> .
- [40] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, *Geophysics* **74** (2009), pp. WCC1–WCC26, <https://doi.org/10.1190/1.3238367> .
- [41] A. Berkhout, *An outlook on the future of seismic imaging, Part III: Joint migration inversion*, *Geophysical Prospecting* **62** (2014), pp. 950–971, <https://doi.org/10.1111/1365-2478.12158> .
- [42] A. Alfaraj and E. Verschuur, *Mitigating the near-surface effect in velocity and reflectivity estimation with multi-scale low-rank approximated image updates*, in *82nd EAGE Conference and Exhibition* (European Association of Geoscientists & Engineers, 2020) pp. 1–5, <https://doi.org/10.3997/2214-4609.202011619> .

3

NEAR-SURFACE PRE-PROCESSING FOR RTM AND LS-RTM

In the presence of near-surface weathering layers, wave propagation may become complex and accurate velocity estimation can be challenging. As a result, reverse-time migration (RTM) and least-squares (LS)-RTM may provide inaccurate images of low-resolution contaminated with artifacts. Therefore, prior data conditioning with near-surface correction is routinely applied on land data. In the previous chapter, we propose and apply a model-independent rank-reduction-based near-surface correction that overcomes the challenges of existing near-surface correction methods. However, with rank-based methods, there can be a concern that subtle subsurface structures may not get preserved. In this chapter, we demonstrate using the SEAM Phase II Arid model that the proposed near-surface correction can preserve subsurface structures and restore the high-resolution of reverse time migration images without the need to access of the subsurface velocity model.

This chapter is a modified version of the proceeding "A. M. Alfaraj, and D. J. Verschuur, Land near-surface pre-processing for RTM and LS-RTM, in 83rd EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2022)

3.1. INTRODUCTION

Wave-equation depth imaging of seismic data with reverse-time migration (RTM) can be used to create subsurface models of steeply dipping structures [1, 2]. However, RTM can be challenged by numerous factors such as limitations in data sampling and wavelet estimation. Using an inversion-based least-squares (LS)-RTM [3, 4], where the difference between observed and modelled data is minimized, can mitigate the acquisition effects and improve the resolution of the migrated image. However, when the data quality is low, e.g. due to low signal-to-noise ratio, prior data conditioning becomes essential.

On land, the addition of complex near-surface weathering layers may form a barrier that prevents imaging targets of interest [5]. This can be the case when the forward modelling operator does not account for the complexity of the weathering layers. As a result, migration with RTM and LS-RTM may result in blurring and inaccuracy as we demonstrate. Therefore, near-surface analysis is routinely carried out when processing land data to perform statics correction and bypass the near-surface weathering layers [6, 7]. However, this process may not be able to fully account for the near-surface effects leading to degraded images, e.g. when violating assumptions of near-surface correction methods. Furthermore, there is an associated risk that near-surface correction with static time shifts may remove details from the subsurface model or may even create non-existing structures, e.g. see [8].

Recently, [9] utilized a data-driven rank-based approach to correct for the short-wavelength component of the near-surface effects. The method accounts for rapid variations of surface-elevation and properties of the near-surface layers, e.g. velocity and thickness, without the need of having a velocity model, while considering the offset dimension [10]. Consequently, the low-rank-base residual statics (LR-ReS) correction is capable of estimating more accurate statics as it is not limited by the surface-consistency assumption. Contrary to other techniques that require normal moveout (NMO) corrected gathers and consequently windowing to avoid the NMO stretch, LR-ReS correction can perform statics correction of the far offsets. Therefore, we propose to utilize it as a pre-processing step for wave-equation-based imaging techniques to correct for the short-wavelength part of the near-surface effects. We also evaluate whether LR-ReS correction is capable of enhancing and preserving the amplitude and structure of the subsurface model in the presence of faults.

3.2. RTM AND LS-RTM

RTM [1, 2] aims at estimating an image of the subsurface in depth, which is based on linearization of the two-way wave equation. It assumes that the subsurface model \mathbf{m} is the sum of a low-wavenumber background model \mathbf{m}_0 that provides the incident wavefield and a high-wavenumber one $\delta\mathbf{m}$ that results in the scattered wavefield:

$$\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{m}. \quad (3.1)$$

The first order approximation, which neglects multiple scattering, of the non-linear forward modelling operator $\mathbf{F}(\mathbf{m}, \mathbf{q})$ can be found by:

$$\mathbf{F}(\mathbf{m}, \mathbf{q}) \approx \mathbf{F}(\mathbf{m}_0, \mathbf{q}) + \frac{\partial \mathbf{F}(\mathbf{m}_0, \mathbf{q})}{\partial \mathbf{m}} \delta\mathbf{m}, \quad (3.2)$$

where \mathbf{q} is the source term. $\frac{\partial \mathbf{F}(\mathbf{m}_0, \mathbf{q})}{\partial \mathbf{m}}$ is the linear Born modelling operator $\mathbf{L}(\mathbf{m}_0, \mathbf{q})$ that can be used to model linearized data $\delta \mathbf{d}$:

$$\delta \mathbf{d} = \mathbf{L}(\mathbf{m}_0, \mathbf{q}) \delta \mathbf{m}. \quad (3.3)$$

Therefore, the RTM image $\delta \mathbf{m}$ can be estimated from a migration velocity model \mathbf{m}_0 and observed data with single scattering $\delta \mathbf{d}_{\text{obs}}$ as follows:

$$\delta \mathbf{m} = \mathbf{L}(\mathbf{m}_0, \mathbf{q})^T \delta \mathbf{d}_{\text{obs}}, \quad (3.4)$$

where $\mathbf{L}(\mathbf{m}_0, \mathbf{q})^T$ is the migration operator.

Since RTM uses the adjoint of the Born modelling operator and not its inverse, migration artifacts can arise. To mitigate these artifacts and enhance the resolution, inversion-based least-squares RTM has been proposed [3, 4], which solves the following linear minimization problem over all shots:

$$\min_{\delta \mathbf{m}} \frac{1}{2} \|\delta \mathbf{d}_{\text{obs}} - \mathbf{L}(\mathbf{m}_0, \mathbf{q}) \delta \mathbf{m}\|_2. \quad (3.5)$$

In the presence of near-surface weathering layers with rapid variations in velocity and thickness, wave propagation becomes complex and accurate velocity estimation can be challenging. As a result, both RTM and LS-RTM (equations 3.4 and 3.5) can result in inaccurate migrated images with low-resolution. Therefore, prior data conditioning that mitigates rapid variations in surface-elevation and properties of the weathering layers, but preserves the subsurface structures becomes essential. To achieve that, we propose to utilize LR-ReS correction prior to RTM and LS-RTM.

3.3. LR-RES ESTIMATION AND CORRECTION

LR-ReS correction [9, 10] exploits the redundancy of seismic data in the midpoint-offset domain. When the data is influenced by near-surface weathering layers, the redundancy of the data decreases, which maps into slowly decaying singular values of monochromatic frequency slices. On the contrary, frequency slices of data unaffected by the weathering layers exhibit higher redundancy and consequently rapidly decaying singular values. To promote the low-rank structure, LR-ReS correction performs not only low-rank approximation, but also statics-estimation by crosscorrelation of low-rank approximated data and data with statics. As a result, amplitude losses due to thresholding the singular values is prevented. The framework operates in an iterative multi-rank-scale fashion over multiple frequency bands to extract multi-scale statics and improve low-rank approximation of high frequency slices. Algorithm 3.1 summarizes the necessary steps for near-surface correction with LR-ReS.

3.4. RESULTS

We demonstrate the performance of pre-processing with LR-ReS correction prior to RTM and LS-RTM using part of the SEAM Phase II Arid model [11]. The selected part consists of channels in the near-surface and flat subsurface layers dissected by a fault to examine whether LR-ReS correction preserves it or not (Figure 3.1(a)). We model linearized data

Algorithm 3.1: Summary of low-rank-based residual statics correction

Input: Data in the source-receiver (s-r) domain, frequency bands and ranks for each frequency and rank-scale

Output: Data after near-surface correction and statics in the midpoint-offset (m-h) and s-r domain

- 1 $\mathbf{D}_{mh} \leftarrow$ Coordinate transformation from the s-r to m-h domain: $h = s - r, m = \frac{s + r}{2}$
- 2 $\tilde{\mathbf{D}}_{mh} \leftarrow$ Fourier transformation from the time to the frequency domain
- 3 **for** loop over rank-scales **do**
- 4 **for** loop over frequency slices **do**
- 5 Singular values decomposition
- 6 $\tilde{\mathbf{D}}_{lr} \leftarrow$ Low-rank approximation
- 7 **if** the desired frequency band for statics-estimation is reached **then**
- 8 $\underline{\mathbf{D}}_{mh} \leftarrow$ Inverse Fourier transformation of $\tilde{\mathbf{D}}_{mh}$
- 9 $\mathbf{D}_{lr} \leftarrow$ Inverse Fourier transformation of $\tilde{\mathbf{D}}_{lr}$
- 10 Statics estimation by crosscorrelation of $\underline{\mathbf{D}}_{mh}$ and \mathbf{D}_{lr}
- 11 $\mathbf{D}_{mh} \leftarrow$ statics correction of \mathbf{D}_{mh}
- 12 $\tilde{\mathbf{D}}_{mh} \leftarrow$ Fourier transformation of \mathbf{D}_{mh}
- 13 Coordinate transformation of the estimated data and statics from the midpoint-offset to the source-receiver domain

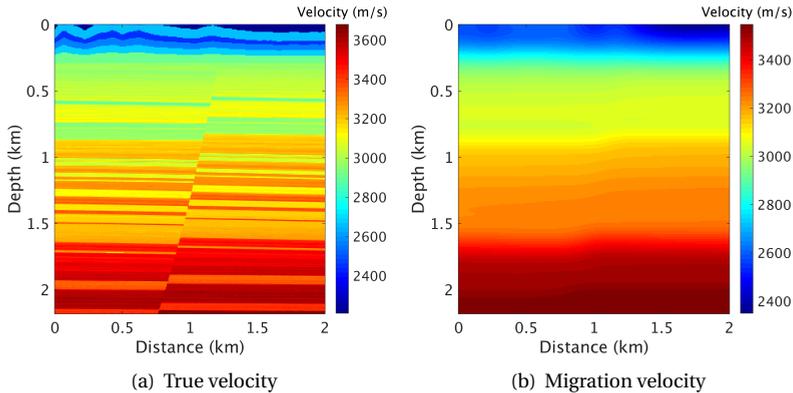


Figure 3.1: (a) True Velocity model and (b) its smooth version (migration velocity model).

with Born modelling (Figure 3.3(a)). The spacing of sources and receivers is 12.5 m and the central frequency of the Ricker wavelet is 35 Hz. RTM of this data (equation 3.4) with a smooth migration velocity (Figure 3.1(b)) provides the image displayed in Figure 3.2(b). Note that the operators we use for modelling and migration [12] are the same to isolate the near-surface effects from other factors such as acquisition footprint.

To imitate rapid variations in the near-surface properties, we add time-shifts that range between -28 and $+28$ ms to each trace of the modelled data (Figure 3.3(b)). The

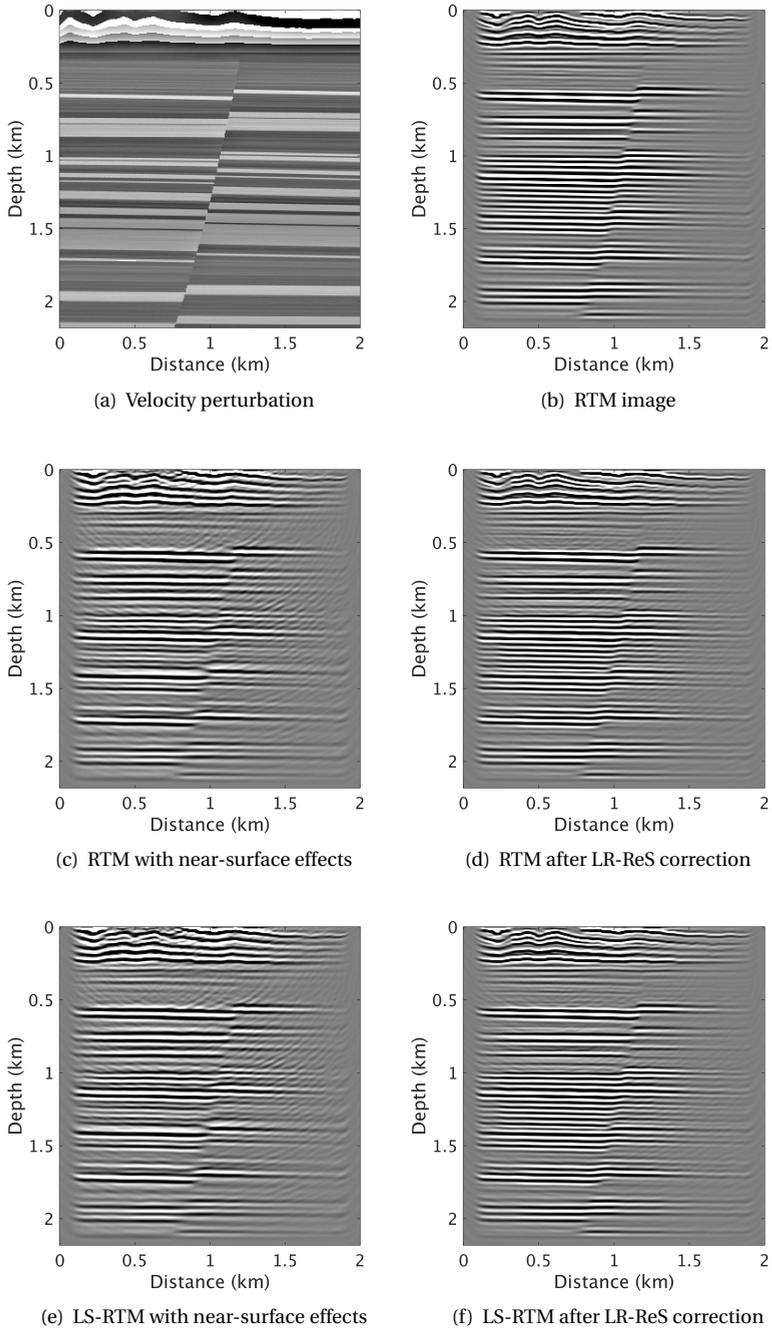


Figure 3.2: (a) Velocity perturbation. RTM image of (b) linearized data, (c) data with near-surface effects and (d) after LR-ReS correction. LS-RTM image of (e) data with near-surface effects and (f) after LR-ReS correction.

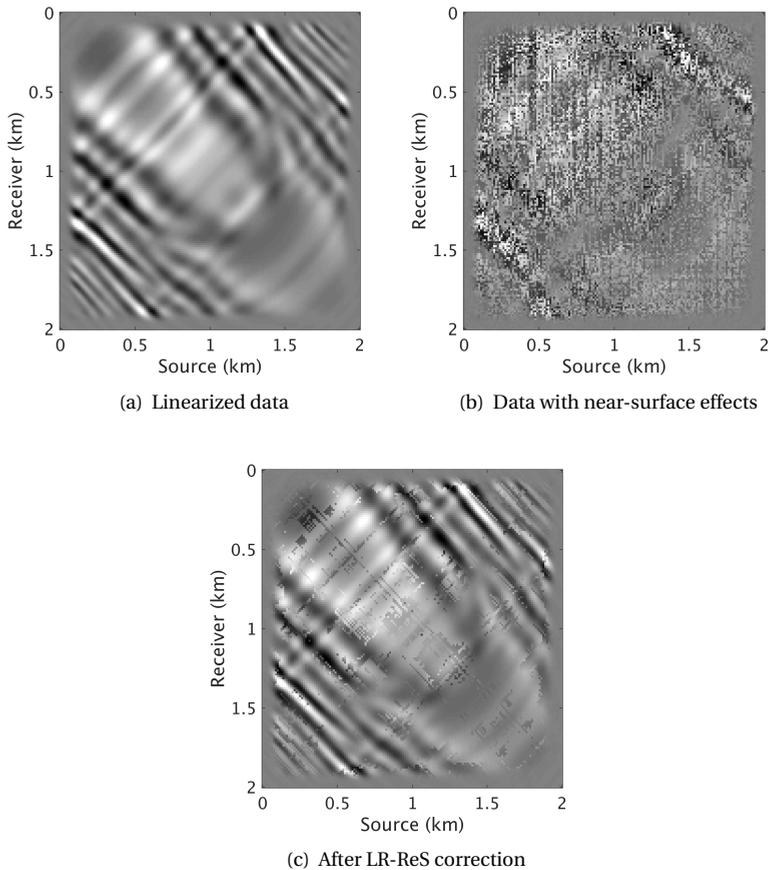


Figure 3.3: Time slices of (a) linearized data, (b) data with near-surface effects and (c) after LR-ReS correction.

time-shifts consist of surface-consistent and non-surface-consistent statics. Consequently, RTM results in a low-resolution image with artifacts (Figure 3.2(c)). Therefore, prior data conditioning is necessary. To do so, we use LR-ReS correction (Algorithm 3.1), which does not require a velocity model, to obtain the data shown in Figure 3.3(c). Performing RTM of this data reduces the artifacts and restores the high-resolution information as visible from the recovered distinct layers (Figure 3.2(d)). In addition, it also preserves the fault and near-surface structures. The average amplitude spectra of the migrated images and traces extracted at location 925 m confirms that data conditioning with LR-ReS correction can overcome the rapid near-surface variations to provide highly similar results to those of statics-free data (Figure 3.4). Repeating the same exercise, but for LS-RTM leads to the same conclusions, but with more meaningful amplitudes (Figures 3.2(e) and 3.2(f)). Adding larger statics (-48 to 40 ms) exacerbates the distortion of the

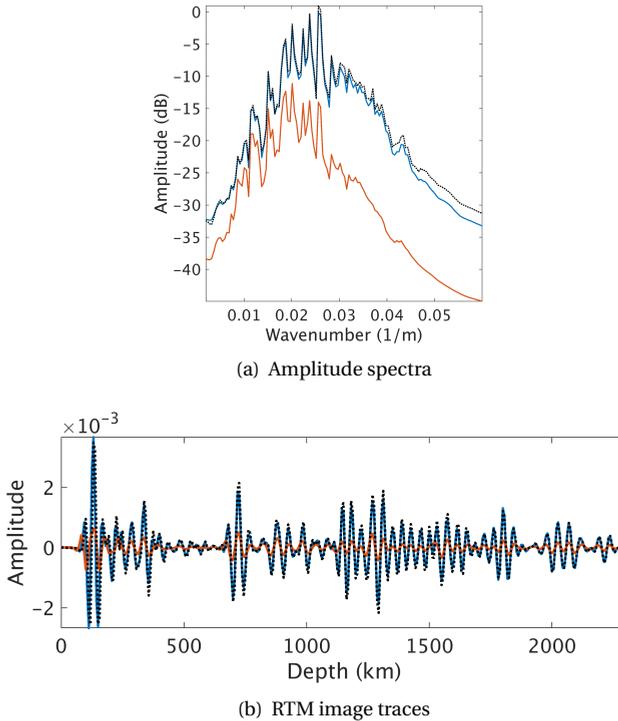


Figure 3.4: (a) Amplitude spectra and (b) RTM image traces at 925 m of linearized data (Figure 3.2(b)) in black, data with near-surface effects (Figure 3.2(c)) in red and after LR-ReS correction (Figure 3.2(d)) in blue.

RTM image and introduces erroneous structures (Figure 3.5(a)), while pre-processing with LR-ReS correction still improves the migrated image (Figure 3.5(b)).

3.5. CONCLUSIONS

In the presence of rapid variations in surface-elevation and properties of the weathering layers, RTM and LS-RTM may provide distorted low-resolution subsurface images. We demonstrate that pre-processing with LR-ReS correction, which is independent of the velocity model can increase the resolution and reduce the artifacts of the migrated images. Moreover, it is capable of preserving the accuracy of the subsurface structures, e.g. faults, as they are captured in the largest singular values of the data.

3.6. ACKNOWLEDGEMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his Ph.D. studies.

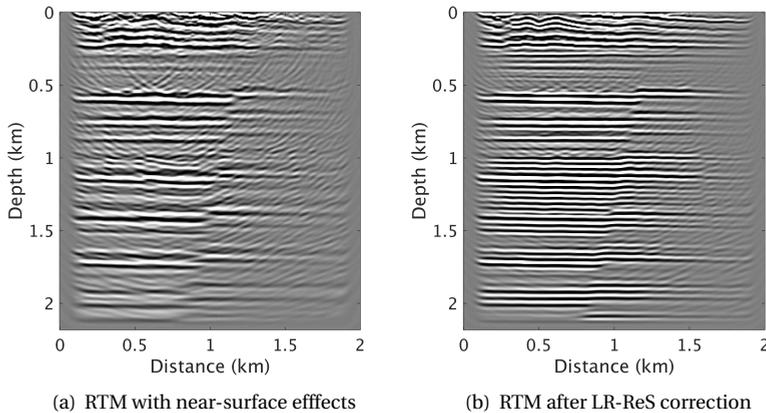


Figure 3.5: RTM image of (a) data with larger near-surface effects and (b) after LR-ReS correction.

REFERENCES

- [1] E. Baysal, D. D. Kosloff, and J. W. Sherwood, *Reverse time migration*, *Geophysics* **48** (1983), pp. 1514–1524, <https://doi.org/10.1190/1.1441434>.
- [2] N. D. Whitmore, *Iterative depth migration by backward time propagation*, in *SEG Technical Program Expanded Abstracts 1983* (Society of Exploration Geophysicists, 1983) pp. 382–385, <https://doi.org/10.1190/1.1893867>.
- [3] S. Dong, J. Cai, M. Guo, S. Suh, Z. Zhang, B. Wang, and Z. Li, *Least-squares reverse time migration: towards true amplitude imaging and improving the resolution*, in *SEG Technical Program Expanded Abstracts 2012* (2012) pp. 1–5, <https://library.seg.org/doi/pdf/10.1190/segam2012-1488.1>.
- [4] C. Zeng, S. Dong, and B. Wang, *Least-squares reverse time migration: Inversion-based imaging toward true reflectivity*, *The Leading Edge* **33** (2014), pp. 962–968, <https://doi.org/10.1190/le33090962.1>.
- [5] T. H. Keho and P. G. Kelamis, *Focus on land seismic technology: The near-surface challenge*, *The Leading Edge* **31** (2012), pp. 62–68, <https://doi.org/10.1190/1.3679329>.
- [6] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999) <https://doi.org/10.1190/1.9781560801818>.
- [7] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <http://dx.doi.org/10.1190/1.9781560801580>.
- [8] C. P. Ursenbach and J. C. Bancroft, *Playing with fire: Noise alignment in trim and residual statics*, in *SEG Technical Program Expanded Abstracts 2001* (2001) pp. 1973–1976, <https://library.seg.org/doi/pdf/10.1190/1.1816525>.

- [9] A. M. Alfaraj, E. Verschuur, and F. J. Herrmann, *Non-surface-consistent short-wavelength statics correction for dense and subsampled data: A rank-based approach*, in *SEG Technical Program Expanded Abstracts 2020* (2020) pp. 2784–2788, <https://doi.org/10.1190/segam2020-3426918.1>.
- [10] A. M. Alfaraj, E. Verschuur, and F. J. Herrmann, *Residual statics correction without NMO – a rank-based approach*, in *First International Meeting for Applied Geoscience & Energy Expanded Abstracts* (2021) pp. 2565–2569, <https://library.seg.org/doi/pdf/10.1190/segam2021-3583455.1>.
- [11] M. Oristaglio, *SEAM update: The arid model–seismic exploration in desert terrains*, *The Leading Edge* **34** (2015), pp. 466–468, <https://doi.org/10.1190/tle34040466.1>.
- [12] P. A. Witte, M. Yang, and F. J. Herrmann, *Sparsity-promoting least-squares migration with the linearized inverse scattering imaging condition*, in *79th EAGE Conference and Exhibition* (European Association of Geoscientists & Engineers, 2017) pp. 1–5, <https://doi.org/10.3997/2214-4609.201701125>.

4

RECONSTRUCTION OF SUBSAMPLED DATA

Imaging and inversion of land seismic data affected by complex weathering layers near the surface are challenging. In chapter 2, we propose a rank-based method to overcome the rapid variations of surface elevation and properties of the near-surface weathering layers. We also demonstrate in chapter 3 that the method can preserve subtle subsurface structures when applied prior to reverse time migration, independently of the subsurface velocity model.

When the data are additionally subsampled for economical reasons such as monitoring of sequestered carbon dioxide and hydrogen, the problem is further exacerbated due to the combined influence of subsampling and weathering layers. On the one hand, interpolation performs poorly since the weathering layers reduce the data's coherency. On the other hand, near-surface correction followed by interpolation requires knowledge of the subsurface model such as separation between primaries and multiples as well as subsurface velocity estimation, which are difficult to perform from subsampled data. Using a model-independent low-rank-based approach, near-surface correction followed by interpolation can be feasible. Nevertheless, both the weathering layers and randomized subsampling render coherent energy incoherent. Therefore, they both contribute to destruction of the low-rank structure commonly associated with statics-free densely-sampled data. Frugal data acquisition in complex near-surface regimes makes separation of the distinct sampling and weathering effects on the rank structure difficult, which as a result lead to poor reconstruction.

To overcome aforementioned impediments, we propose to reconstruct the data with joint rank-reduction-based near-surface correction and interpolation. The method simultaneously accounts for the weathering and subsampling effects to provide accurate reconstruction. Since low-rank approximation is used for near-surface correction, we

This chapter is a modified version of the manuscript undergoing review "A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, Seismic data reconstruction in complex near-surface regimes, Geophysics."

also utilize it in rank-minimization interpolation as a cost-free initial solution to the optimization problem. As both near-surface correction and interpolation operate in the midpoint-offset domain, we avoid the cost of transformations back and forth from the source-receiver to midpoint-offset transform domain. Consequently, the proposed reconstruction, which shows its potential on synthetic and field data, additionally increases the computational efficiency.

4.1. INTRODUCTION

Due to their low velocity and rapidly varying nature, the weathering layers near the surface dictate how land seismic data get acquired [1]. An additional layer of complexity arises when collecting data in the presence of acquisition gaps. This can be due to acquisition limitations such as obstacles. Or to acquire economical data, which are of high demand for engineering purposes or monitoring of sequestered captured carbon dioxide and reservoirs, the field of compressive sensing [2–4] suggests avoiding the strict Nyquist sampling criterion and acquire randomly subsampled data. This is another scenario, where the data contain gaps that require reconstruction to fill-in the missing traces.

In the literature, there exist many reconstruction techniques. Transform-domain-based methods are popular as they are capable of distinguishing and using properties of dense and subsampled data. Such domains include the Fourier [5], curvelet [6–8] and Radon domains [9, 10]. An alternative approach is to use rank-based methods, which can process multidimensional large scale seismic data in a computationally efficient manner [11–14]. However, rank-based methods that require repetitive computations of the singular values of large scale matrices can be computationally demanding. This issue can be mitigated with truncated singular value decomposition or matrix factorization techniques [15, 16]. Another challenging issue is rank-based interpolation of high frequencies, which [17] improve by using common information amongst neighbouring frequencies. While the aforementioned reconstruction methods demonstrate their success in reconstruction of statics-free and therefore low-rank seismic data, reconstruction of data affected by weathering layers, which lead to low data coherency, remains challenging.

[18] studies 5D sparse Fourier interpolation and finds that interpolation becomes difficult when there are imperfections in the applied statics before interpolation, which we also illustrate. [19] use sparsity promotion in the Fourier domain with projection onto convex sets to reconstruct subsampled data contaminated with residual statics. Even though it shows improved reconstruction of data synthetically shifted with ± 10 ms of residual statics, it requires data windowing to ensure sparsity, which is not a trivial task and may influence the results. When the windowed data are merged back together, there is a potential of introducing errors along the spatial and temporal dimensions as each window may get assigned different statics. [20] applies phase retrieval using only the amplitude spectrum with sparsity promoting regularization to interpolate data affected by residual statics. Sparsity promotion in the Fourier and curvelet domains has been also used for statics correction [21], which also require data windowing. [22] acquire and reconstruct compressively sampled land data, where statics correction due to ice lakes is one of the challenging issues that need to be resolved. In complex near-surface regimes, where statics correction on its own is a major challenge, data reconstruction is also expected to suffer.

The current approach to tackle the combined effects of weathering and acquisition gaps is to first perform near-surface correction to make the data more coherent, followed by data interpolation of normal moveout (NMO) corrected data [18]. However, interpolation and statics estimation depend on knowledge of the subsurface model such as NMO velocity, which requires distinction between primaries and multiples [23–25]. Since the velocity model is usually influenced by the statics, and vice versa [26], velocity estimation becomes more challenging when the data are additionally subsampled as the num-

ber of traces in each common midpoint (CMP) decreases. Despite the abundant experience with NMO velocity estimation, it remains efforts- and time-consuming, and prone to errors in the above situations or in the presence of multiples that can be confused with primaries. To avoid dependency on the subsurface model for near-surface correction of periodically- and densely-sampled data, [27, 28] use a model-independent low-rank (LR)-based approach. However, acquisition of such data is prohibitively expensive. On the other hand, reduction of the acquisition costs with randomized subsampling influences the rank structure, which consequently affects the LR-based near-surface correction.

4.1.1. CONTRIBUTIONS

Our contributions are towards reconstruction of randomly subsampled land data influenced by complex weathering layers. Such problem requires solving for the near-surface and subsampling effects as both contribute to the low-rank structure destruction (both effects render coherent energy incoherent that breaks the low-rank structure typically associated with statics-free densely sampled data). The low-rank structure destruction means that the singular values no longer decay fast, while in the appropriate domain, seismic data can be well-approximated by low-rank matrices. When collecting economical data in complex near-surface regimes, separation between the subsampling and weathering effects becomes difficult. Consequently, low-rank-based model-independent near-surface correction followed by interpolation, or vice-versa performs poorly. This is because interpolation and near-surface correction require that data incoherency comes only from subsampling or statics, respectively, but not both. To obtain accurate densely-sampled data, we propose joint reconstruction with rank-reduction-based near-surface correction and interpolation. With this approach, near-surface correction improves interpolation, and vice versa. Since we need to compute low-rank approximation for near-surface correction, we utilize it as a cost-free initial solution for the rank-minimization optimization problem. We also lower the number of required transformations between the acquisition and transform domain as both interpolation and near-surface correction operate in the midpoint-offset domain. Therefore, the proposed reconstruction additionally improves the computational efficiency. To demonstrate its potential, we apply it to synthetic data generated with acoustic finite difference modelling and field data affected by complex weathering layers and noise.

4.1.2. OUTLINE

The paper is organized as follows. We first study the singular values decay behaviour in the combined presence of weathering and subsampling effects. According to the findings, we propose joint rank-reduction-based near-surface correction and interpolation. We provide the details of the proposed algorithm and its components, which we subsequently apply to synthetic and field data. We then discuss the parameters selection, computational efficiency, limitations and extensions of the method.

4.2. THE SINGULAR VALUES DECAY

Similar to other rank-reduction-based methods, our proposed reconstruction exploits the redundant nature of seismic data that leads to a low-rank structure, where the singular

values decay rapidly. When the data are acquired with randomized subsampling [2–4] or when influenced by the weathering layers, where both processes render coherent energy incoherent, the low-rank structure gets destroyed leading to slowly decaying singular values [28, 29]. To observe their combined influence, we analyze the singular values decay behaviour.

For the analysis, we simulate densely sampled data with 10 m source- and receiver-intervals using the velocity model displayed in Figure 4.1 with finite difference modelling [30]. To exploit the redundant nature of the seismic data, we use the midpoint-offset domain:

$$m = \frac{s + r}{2} \quad (4.1a)$$

and

$$h = s - r, \quad (4.1b)$$

where s , r , m and h are the source, receiver, midpoint and offset coordinates. Since the transformation rotates a matrix (complex-valued frequency slice) in the source-receiver domain by 45° , the matrix columns become more linearly dependent in the midpoint-offset transform domain. This can be visually seen when comparing frequency slices in the source-receiver and midpoint-offset domain (Figures 4.2, where the columns in the latter exhibit lower variability compared with the former. Consequently, the singular values in the midpoint-offset domain decay rapidly compared with those in the source-receiver domain. However, when the data become incoherent, the low-rank structure gets destroyed because linear dependency of the matrix columns become lower.

To replicate complex near-surface conditions, we shift the simulated data with up to ± 52 ms of statics composed of surface- and non-surface-consistent elements. We then subsample the data by randomly removing 75% of the shots to mimic a compressive sensing acquisition design (Figures 4.3(a), 4.3(b)). We note that removing 75% of the shots may sound harsh, but that is analogous to an average of 40 m periodic shot spacing, which is common for data acquisition in large scale areas. While it is optimal to collect densely sampled data, it is also important to reduce the acquisition costs. For example, economical data is necessary for engineering purposes or subsurface monitoring of captured carbon dioxide and reservoirs. However, in complex near-surface regimes, the singular values (Figures 4.3(e) and 4.3(f)) become slowly decaying due to the effects of statics and subsampling compared to the statics-free densely sampled data. Therefore, we can use rank-reduction methods to estimate the latter from subsampled data affected by the weathering layers.

4.3. METHODOLOGY

Promotion of the low-rank structure requires solving for the near-surface and subsampling effects as both contribute to the low-rank structure destruction. We propose to do so with a *novel* joint rank-reduction-based near-surface correction and interpolation.

4.3.1. LR-BASED NEAR SURFACE CORRECTION

The LR-based near-surface estimation and correction iteratively estimates multi-scale short-wavelength statics in the midpoint-offset domain independently of knowledge of

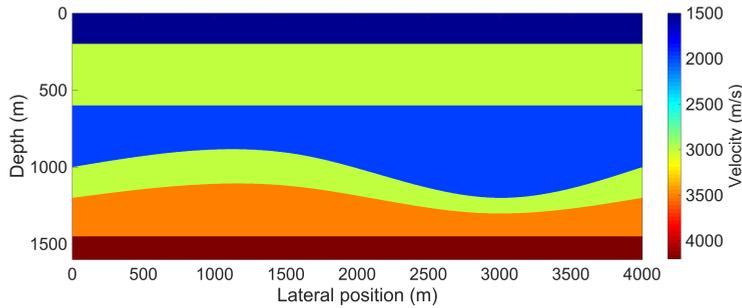


Figure 4.1: Velocity model used to simulate the synthetic data.

4

the subsurface model [27, 28]. It promotes the low-rank structure, which corresponds to matrices that can be approximated by a low-rank matrix in the appropriate domain. It is based on the singular value decomposition (SVD)[31]:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \quad (4.2)$$

where $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$ is a midpoint-offset frequency slice of dimensions n_m (number of midpoints) and n_h (number of offsets). $\mathbf{S} \in \mathbb{R}^{k \times k}$ is the block diagonal matrix containing the non-negative real-valued singular values $\mathbf{S} = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k)$, where $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_k \geq 0$ and $k = \min\{n_m, n_h\}$. $\mathbf{U} \in \mathbb{C}^{n_m \times k}$ and $\mathbf{V} \in \mathbb{C}^{n_h \times k}$ are the orthogonal matrices that hold the left \mathbf{u} and right \mathbf{v} singular vectors, respectively, while H denotes the Hermitian transpose. A low-rank approximation of \mathbf{X} can be obtained by selecting few singular vectors that correspond to the few largest singular values. The statics can be estimated by calculating the lag of the maximum cross-correlation between subsampled data with statics and low-rank approximated data. The process is performed on band-pass-filtered data in the time-domain after multi-scale low-rank approximation of frequency slices. This enables exploiting the relationship amongst multiple frequency bands and avoiding spurious statics from corrupted frequencies. The method starts with using the low frequencies for statics estimation as they are less influenced by short-wavelength statics compared to high frequencies. Then it applies these statics to the full-band data. As a result, the coherency of the high frequencies gets improved, which allows for updating the statics when including them in the next iteration.

The method shows its potential for statics estimation and correction of periodically and densely sampled data [28]. Figure 4.4 displays the results after three multi-scale iterations when additionally including randomized subsampling. The data in the time-domain becomes more coherent (Figures 4.4(b) and 4.4(e)) compared to the incoherent data we start with (Figures 4.4(a) and 4.4(d)). However, there are imperfections in the estimated statics, which are noticeable when comparing the statics-corrected data (Figures 4.4(b) and 4.4(e)) with the statics-free data (Figures 4.4(c) and 4.4(f)). They are more obvious from the statics error displayed in Figures 4.4(g) and 4.4(h). Next, we investigate whether this error influences data interpolation or not.

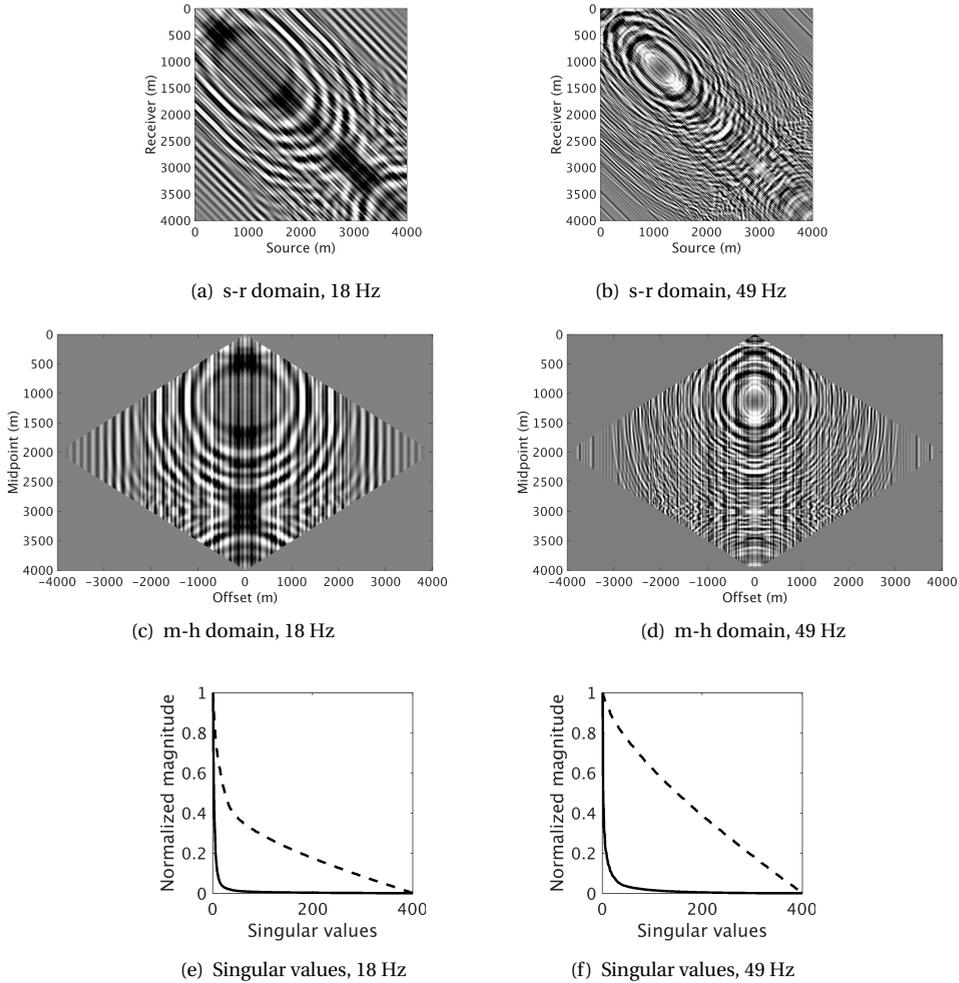


Figure 4.2: Densely-sampled, statics-free frequency slices (real component) of synthetic data in the (a, b) source-receiver domain and (c, d) midpoint-offset domain along with (e, f) their singular values decay curves computed from the complex-valued frequency slices at (a, c, e) 18 and (b, d, f) 49 Hz. Dashed curves correspond to frequency slices in the source-receiver domain and solid curves correspond to those in the midpoint-offset domain.

4.3.2. RANK-MINIMIZATION-BASED INTERPOLATION

To interpolate the data, we use rank-minimization-based interpolation with matrix factorization due its computational efficiency [15, 29]. For each midpoint-offset frequency slice $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$, we solve the following basis pursuit denoising problem (BPDN):

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \epsilon, \quad (4.3)$$

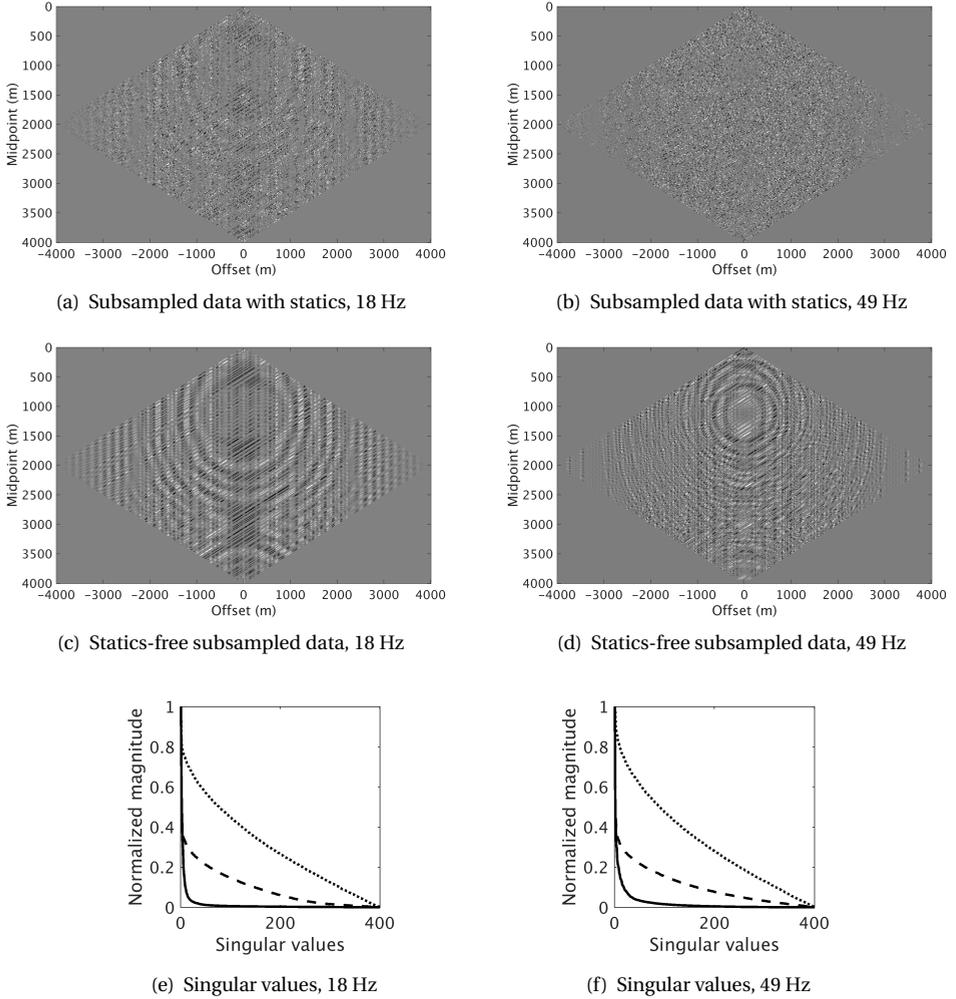


Figure 4.3: Frequency slices in the midpoint-offset domain after 75% randomized subsampling of shots (a, b) affected by ± 52 ms of non-surface-consistent statics and (c, d) statics-free data along with (e, f) their singular values decay at (a, c, e) 18 and (b, d, f) 49 Hz. Dotted curves correspond to (a, b), dashed curves correspond to (c, d) and solid curves correspond to statics-free frequency slices shown in Figures 4.2(c) and 4.2(d).

where $\mathbf{b} \in \mathbf{C}^{n_p}$ is the observed frequency slice on the subsampled source-receiver grid $n_{\bar{s}} \times n_{\bar{r}}$ organized as a vector with dimension $n_p = n_{\bar{s}} \times n_{\bar{r}} < n_s \times n_r$. ϵ is the noise-level, $\|\mathbf{X}\|_* = \sum_{i=1}^k \sigma_i$ is the nuclear norm of \mathbf{X} , while σ_i contains the singular values defined in equation 5.1. The operator \mathcal{A} is the sampling-transform operator composed of a midpoint-offset

transform domain operator \mathcal{S} and a measurement operator \mathbf{M} :

$$\mathcal{A} = \mathbf{M}\mathcal{S}^\top, \quad (4.4)$$

where \mathcal{S}^\top (equation 4.1) transforms the data from the midpoint-offset to the source-receiver domain. Therefore, \mathcal{A} maps the data from $\mathbf{C}^{n_m \times n_h} \rightarrow \mathbf{C}^{n_p}$, where $n_p \ll n_m \times n_h$. The transform domain operator rotates a matrix in the source-receiver domain by 45° , which makes its columns close to linearly-dependent and therefore, of low-rank nature. The measurement operator decimates the data by removing the unmeasured entries, i.e. it is composed of ones and zeros, which correspond to the measured and unmeasured samples, respectively. For a compressive sensing scenario, the sensing matrix \mathbf{M} aims to destroy the coherency, e.g. with uniform randomized sampling, which results in slowly decaying singular values as shown in Figures 4.3(e) and 4.3(f). To solve equation 4.3, we use an extension of the SPG ℓ_1 solver [32] that solves a sequence of LASSO subproblems [33]:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \quad \text{subject to} \quad \|\mathbf{X}\|_* \leq \delta, \quad (4.5)$$

where $\|\mathbf{X}\|_* \leq \delta$ is the projection on the nuclear norm ball. It requires computation of the singular values with SVD (equation 5.1) followed by thresholding. Calculation of the singular values at each iteration per frequency slice is computationally demanding for large scale matrices, which can be avoided by using a more efficient matrix factorization technique (see [15, 16, 29, 32–34] for more details).

Rank-minimization interpolation demonstrates its capability in reconstruction of marine data, which are usually not influenced by the weathering layers. Therefore, they exhibit higher coherency compared to land data. That can be seen from Figure 4.3, where the low-rank structure destruction is low for the statics-free subsampled data compared to that influenced by statics. Consequently, interpolation of the incoherent randomly subsampled data with statics leads to a noisy reconstruction (Figures 4.5(a) and 4.5(b)) since the low-rank structure destruction due to the weathering layers is not considered. In this case, statics correction after interpolation will still be noisy. Interpolation after improving the coherency with the LR-based near-surface correction provides a better reconstruction (Figures 4.5(c) and 4.5(d)). However, the reconstruction is erroneous and contains noise as can be seen from the frequency slices and the time-domain gathers (Figures 4.5(c), 4.5(d), 4.6(a) and 4.6(c)). Consequently, the reconstruction error is large (Figures 4.7(a) and 4.7(c)). That is a consequence of the imperfections in the estimated statics (Figure 4.4), which is due to the mixed effects of weathering and subsampling on the rank structure. Therefore, without fully accounting for the weathering layers, interpolation becomes challenging as also indicated by [18] and [22].

4.3.3. JOINT NEAR-SURFACE CORRECTION AND INTERPOLATION

To reconstruct economical data in complex near-surface regimes, we propose to simultaneously correct for the weathering layers and interpolate the data. For interpolation with equation 4.3, \mathbf{b} refers to the observed subsampled frequency slice arranged as a vector. Since the LR-based near-surface estimation and correction operate on matrices in the midpoint-offset domain on the desired grid of densely sampled data, we modify equation

4.3 to:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathcal{M}_{\text{mh}}(\mathbf{X}) - \widehat{\mathbf{b}}_{\text{mh}}\|_2 \leq \epsilon, \quad (4.6)$$

where

$$\mathbf{b}_{\text{mh}} = \mathcal{M}_{\text{mh}} \mathcal{S}(\mathbf{M}^\top \mathbf{b}) \quad (4.7)$$

is the subsampled frequency slice on the desired midpoint-offset grid, $\widehat{\mathbf{b}}_{\text{mh}}$ is the same slice, but after near-surface estimation and correction and $\mathcal{M}_{\text{mh}} = \mathcal{S} \circ \mathbf{M}$ is the measurement operator in the midpoint-offset domain. Compared to equation 4.3, equation 4.6 removes the need to apply the adjoint of the transform domain operator \mathcal{S}^\top at each iteration. Therefore, since the subsampled measured data after statics correction $\widehat{\mathbf{b}}_{\text{mh}}$ is already in the midpoint-offset domain and since the optimization problem is solved over many iterations per frequency slice in the same domain, we reduce the computational costs of going back and forth to the transform domain while solving the optimization problem.

The initial solution \mathbf{X}_0 of equation 4.6 is critical as it determines the number of required iterations. By using the low-rank approximated data $\widehat{\mathbf{D}}_{\text{lr}}^{(f)}$ as the initial solution, we get closer to the desired one, which decreases the number of iterations compared to starting with an empty or random matrix. The reason is that low-rank approximation can be an interpolator (see Figure 5.6), but an inaccurate one if the neglected singular vectors below the selected singular value are of importance. Therefore, the joint reconstruction additionally increases the computational efficiency because low-rank approximation is used for both near-surface correction and interpolation as an initial solution. Algorithm 4.1 shows the details of our proposed method.

Throughout the algorithm, (\cdot) , $(\hat{\cdot})$ and $(\check{\cdot})$ indicate subsampled data, estimated data and data in the frequency domain, respectively. The subscripts $_{\text{mh}}$ and $_{\text{sr}}$ correspond to data in the midpoint-offset and source-receiver domains, respectively, while $_{\text{mh},i}$ and $_{\text{sr},i}$ specify interpolated data. The input to the algorithm is subsampled data in the source-receiver domain $\check{\mathbf{D}}_{\text{sr}}$, frequency bands \mathbf{f}_{b} at which we estimate the statics and ranks used for interpolation \mathbf{k} and for low-rank approximation \mathbf{K} of each frequency slice. The output is the interpolated and statics-corrected data in the source-receiver $\check{\mathbf{D}}_{\text{sr},i}$ and midpoint-offset $\widehat{\mathbf{D}}_{\text{mh},i}$ domains.

At the first step of the algorithm, we apply the transpose of the sampling matrix \mathbf{M}^\top , which inserts to the acquired data empty rows or columns that will be filled by interpolation. Therefore, the subsampled source-receiver grid $n_{\check{r}} \times n_{\check{s}}$ becomes the desired grid of densely sampled data $n_r \times n_s$. Since rank selection is more difficult for data in the time-domain, which contains more variability as it encompasses all frequencies, we carry out the low-rank approximation required for statics correction as well as interpolation in the frequency domain as indicated by the second step. The LR-based near-surface correction exploits the relationships amongst multiple frequency bands. By applying the estimated statics at low frequencies to the full-band data, the coherency of the higher frequencies improves since neighbouring frequency bands share common statics. Consequently, the accuracy of their low-rank approximation increases. After partial statics correction, which renders incoherent energy coherent, the singular values decay becomes faster. This allows for using a smaller rank for more accurate low-rank approximation to estimate a signal with less statics influence. To do so, the algorithm contains loops over the number

Algorithm 4.1: Joint rank-reduction-based near-surface correction and interpolation

Input: $\tilde{\mathbf{D}}_{\text{sr}} \in \mathbb{R}^{n_t \times n_f \times n_s}$, $\mathbf{k} \in \mathbb{R}^{n_f}$, $\mathbf{K} \in \mathbb{R}^{n_f \times n_l}$, $\mathbf{f}_b \in \mathbb{R}^{n_{f_b}}$

Output: $\tilde{\mathbf{D}}_{\text{mh},i} \in \mathbb{R}^{n_t \times n_h \times n_m}$, $\tilde{\mathbf{D}}_{\text{sr},i} \in \mathbb{R}^{n_t \times n_r \times n_s}$

- 1 Apply \mathbf{M}^T to each time-slice of $\tilde{\mathbf{D}}_{\text{sr}}$ to obtain $\mathbf{D}_{\text{sr}} \in \mathbb{R}^{n_t \times n_r \times n_s}$
 - 2 Transform \mathbf{D}_{sr} to $\mathbf{D}_{\text{mh}} \in \mathbb{R}^{n_t \times n_h \times n_m}$ with equation 4.1
 - 3 fft \mathbf{D}_{mh} to $\tilde{\mathbf{D}}_{\text{mh}} \in \mathbb{C}^{n_m \times n_h \times n_f}$
 - 4 **for** $i \leftarrow 1$ to $n_{f_b} - 1$ **do**
 - 5 **for** $l \leftarrow 1$ to n_l **do**
 - 6 **for** $f \leftarrow \mathbf{f}_b^{(i)} + 1$ to $\mathbf{f}_b^{(i+1)}$ **do**
 - 7 Calculate SVD for $\tilde{\mathbf{D}}_{\text{mh}}^{(f)}$ with equation 4.2
 - 8 $\tilde{\mathbf{D}}_{\text{lr}}^{(f)} \leftarrow \sum_{j=1}^{\mathbf{K}^{(f,l)}} s^{(j)} \mathbf{u}^{(j)} \mathbf{v}^{(j)H}$
 - 9 **if** $f = \mathbf{f}_b^{(i+1)}$ **then**
 - 10 ifft $\tilde{\mathbf{D}}_{\text{lr}}$ to \mathbf{D}_{lr} , ifft $\tilde{\mathbf{D}}_{\text{mh}}$ to \mathbf{D}_{mh}
 - 11 $\mathbf{T}_{\text{mh}} \leftarrow \mathcal{C}(\mathbf{D}_{\text{mh}}, \mathbf{D}_{\text{lr}})$
 - 12 $\hat{\mathbf{D}}_{\text{mh}} \leftarrow \tau(\mathbf{D}_{\text{mh}}, \mathbf{T}_{\text{mh}})$
 - 13 $\mathbf{D}_{\text{mh}} \leftarrow \hat{\mathbf{D}}_{\text{mh}}$, fft \mathbf{D}_{mh} to $\tilde{\mathbf{D}}_{\text{mh}}$
 - 14 **if** $l = n_l$ **then**
 - 15 **for** $f \leftarrow \mathbf{f}_b^{(i)} + 1$ to $\mathbf{f}_b^{(i+1)}$ **do**
 - 16 $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}^{(f)} \leftarrow$ solution of equation 4.6:
 - 17 $\hat{\mathbf{b}}_{\text{mh}} \leftarrow \mathcal{M}_{\text{mh}}(\tilde{\mathbf{D}}_{\text{mh}}^{(f)}), \mathbf{X}_0 \leftarrow \tilde{\mathbf{D}}_{\text{lr}}^{(f)}$
 - 18 Repeat steps 7 and 8 with $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}^{(f)}$ instead of $\tilde{\mathbf{D}}_{\text{mh}}$ to get $\tilde{\mathbf{D}}_{\text{LR},i}$
 - 19 Repeat steps 10 – 12 with $\tilde{\mathbf{D}}_{\text{LR},i}$ and $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}^{(f)}$ instead of $\tilde{\mathbf{D}}_{\text{lr}}$ and $\tilde{\mathbf{D}}_{\text{mh}}$, respectively, to estimate $\mathbf{T}_{\text{mh},i}$
 - 20 Apply $\mathbf{T}_{\text{mh},i}$ to $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}^{(f)}$ and \mathbf{D}_{mh}
 - 21 Repeat step 13
 - 22 Transform $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}$ to $\hat{\tilde{\mathbf{D}}}_{\text{sr},i}$ with equation 4.1
 - 23 ifft $\hat{\tilde{\mathbf{D}}}_{\text{mh},i}$ to $\hat{\mathbf{D}}_{\text{mh},i}$ and $\hat{\tilde{\mathbf{D}}}_{\text{sr},i}$ to $\hat{\mathbf{D}}_{\text{sr},i}$
-

of frequency bands n_{f_b} (step 4) and rank-scales n_l (step 5). For each frequency slice in a certain frequency band (step 6), we calculate the singular values (step 7), followed by low-rank approximation to obtain low-rank approximated data $\tilde{\mathbf{D}}_{\text{lr}}$ (step 8), where s , \mathbf{u} and \mathbf{v} correspond to the singular values, the left and right singular vectors, respectively. After low-rank approximation of all the frequencies in the current frequency band (step 9), we inverse Fourier transform the low-rank approximated data \mathbf{D}_{lr} and data with statics \mathbf{D}_{mh} to the time-domain (step 10), where we estimate the statics \mathbf{T}_{mh} (step 11) by calculation of the lag that corresponds to the maximum cross-correlation coefficient:

$$t_{\text{mh}} = \arg \max_t (\mathbf{d}_{\text{mh}}(t) \star \mathbf{d}_{\text{lr}}(t)), \quad (4.8)$$

where \star indicates cross-correlation and the notation (t) is used to show time-dependence. t_{mh} is the time-shift of one trace estimated from data with statics \mathbf{d}_{mh} and low-rank approximated data \mathbf{d}_{r} . After estimating the statics \mathbf{T}_{mh} from all the traces \mathbf{D}_{mh} and \mathbf{D}_{r} , they are used for statics correction (step 12) to estimate statics-corrected data $\hat{\mathbf{D}}_{\text{mh}}$:

$$\hat{\mathbf{d}}_{\text{mh}} = \mathbf{d}_{\text{mh}}(t + t_{\text{mh}}). \quad (4.9)$$

The subsampled data with statics is then updated to be the statics-corrected data (step 13). After partial statics correction of the subsampled data, it becomes ready for interpolation as it exhibits higher coherency. After the multi-rank-scale iterations (step 14), we solve equation 4.6 to interpolate the statics-corrected frequency slices that belong to the current frequency band as indicated by steps 15 and 16.

Since the low-rank structure destruction is not only a result of the statics, but also randomized subsampling (Figure 4.3), low-rank approximation of subsampled data, which we need for statics estimation, can not preserve data without the statics influence. We need more accurate low-rank approximation that can preserve the statics-free data, which requires data with higher coherency. Since randomized subsampling decreases the coherency, we use the interpolated frequencies, which now exhibit faster singular values decay to allow for more accurate low-rank approximation. We repeat the statics estimation and correction, but using the interpolated data (steps 17 – 20). The estimated statics $\mathbf{T}_{\text{mh},i}$ are then used for statics correction of the interpolated and subsampled (to-be used for statics estimation at the next iteration and to-be interpolated) data $\hat{\mathbf{D}}_{\text{mh},i}$ and \mathbf{D}_{mh} , respectively. Therefore, not only statics correction improves interpolation, but also vice versa. Consequently, reconstruction with the proposed joint scheme provides better results with less noise compared to the stepwise one (Figure 4.5). Finally, the data can be transformed to the source-receiver and time-domain (steps 21 and 22). We further analyse the results in the next section.

4.4. RESULTS

Using synthetic and field data, we further demonstrate the potential of our proposed reconstruction method.

4.4.1. SYNTHETIC DATA

As illustrated by Figures 4.6(a) and 4.6(c), imperfect statics correction results in poor reconstruction and large error (Figures 4.7(c) and 4.7(a)). On the other hand, the proposed reconstruction, which simultaneously corrects for the near-surface effects and interpolates the data (Algorithm 5.1), provides improved results with less noise and better quality as shown in Figures 4.6(b) and 4.6(d). Therefore, by using the interpolated frequencies, we can improve the estimated statics as iterations progress. This can be observed from the error of the estimated statics, which is minimal for the proposed reconstruction (Figures 4.6(e) and 4.6(f)) compared to the stepwise one (Figures 4.4(g) and 4.4(h)). Consequently, the reconstruction amplitude error is lower for the former in comparison with the latter (see Figures 4.7(d) and 4.7(c)). The reconstruction error of each frequency slice relative to the statics-free densely sampled one shown in Figure 4.9 also confirms that the joint approach improves the reconstruction as its error is lower compared to the stepwise one.

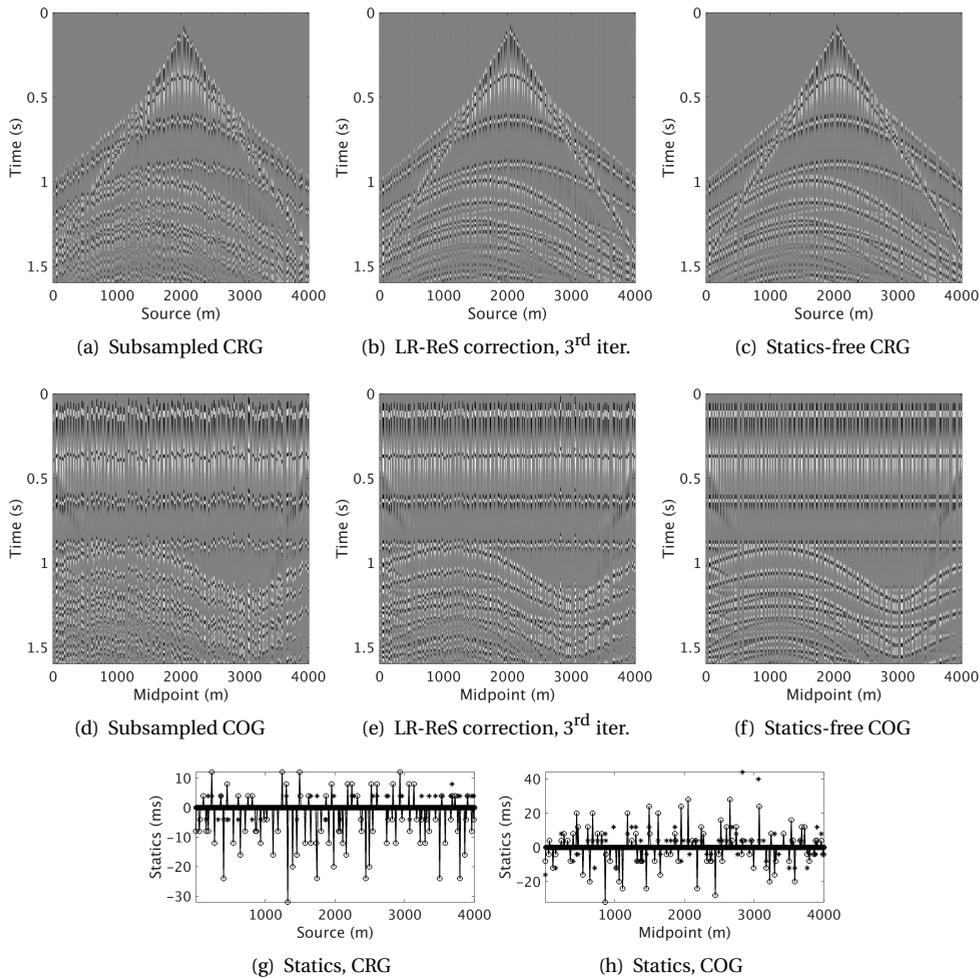


Figure 4.4: Common (a, b, c) receiver and (d, e, f) offset gathers of (a, d) 75% randomly subsampled data with statics, after (b, e) near-surface correction at the third iteration and (c, f) statics-free data. (g, h) The estimated total statics of the displayed (g) common receiver gather (CRG) and (h) common offset gather (COG) plotted in solid lines with circle markers and their error plotted with asterisks.

The same figure also shows that interpolation without statics correction leads to a large error, which makes statics estimation and correction after interpolation challenging.

4.4.2. FIELD DATA

To further examine the potential of the proposed joint reconstruction, we test its performance on field data. The field data set is affected by complex weathering layers, which

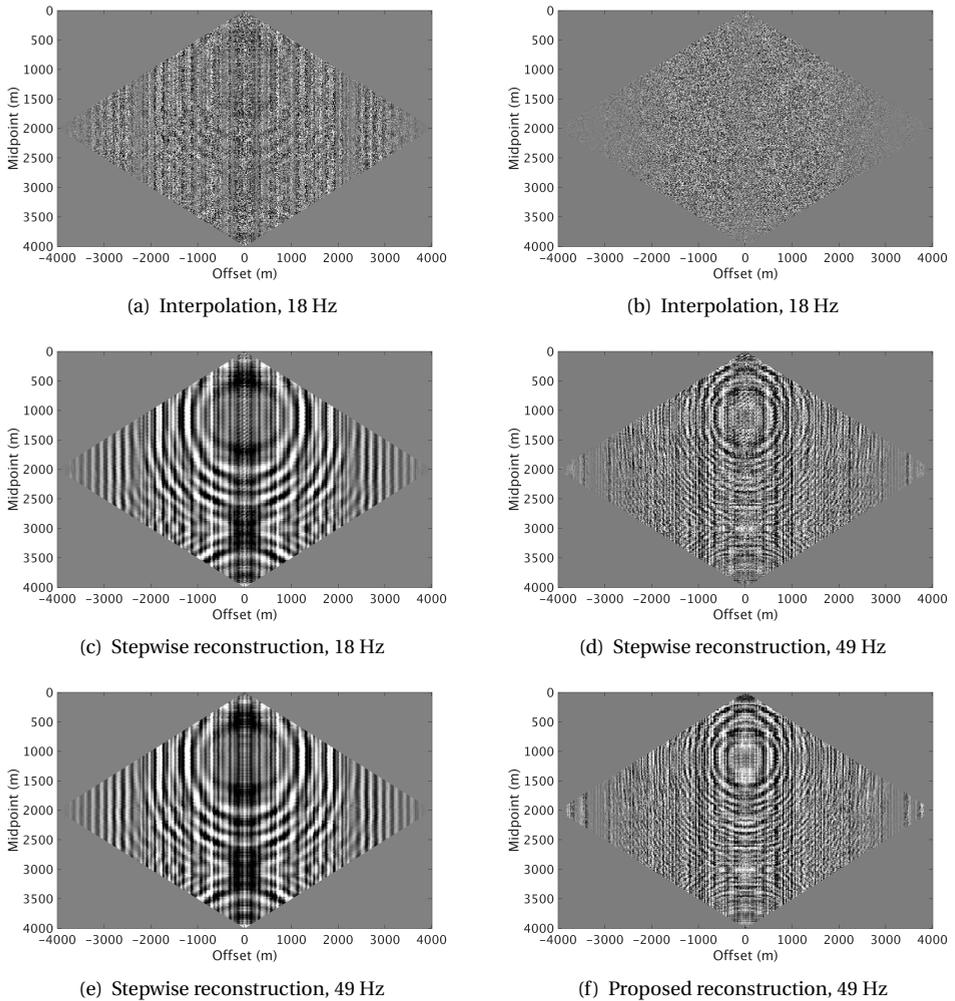


Figure 4.5: Midpoint-offset frequency slices after (a, b) interpolation without statics correction, (c, d) stepwise reconstruction and (e, f) proposed reconstruction at (a, c, e) 18 and (b, d, f) 49 Hz.

include loose sand, fast carbonates, karsts and gravel with surface elevation changes of about 100 m across the line [35]. The data also contain residual ground roll due to a simple ground roll attenuation and random noise with different characteristics given the different near-surface conditions. In this case, the near-surface on its own is challenging to data processing as can be seen from the gathers and stack, which contain poorly continuous reflectors affected by statics (Figures 4.10(a), 4.11(a) and 4.12(a)). Note that NMO correction is only used for display, but not during the reconstruction process. Additionally, the data set is acquired with 30 m source and receiver intervals, which we further decimate

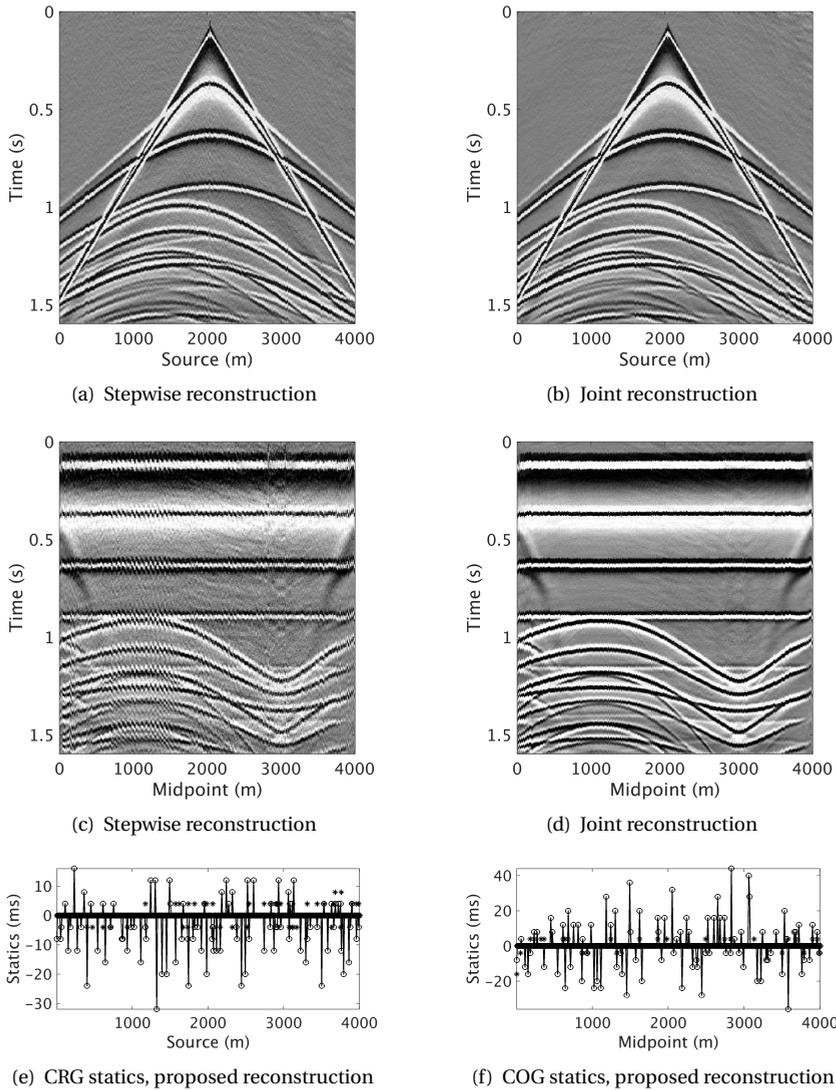


Figure 4.6: Common (a, b) receiver and (c, d) offset gathers after (a, c) stepwise and (b, d) joint reconstruction of 75% randomly subsampled data with statics. (e, f) The estimated statics with the proposed reconstruction for the displayed CRG (b) COG (d) plotted in solid lines with circle markers and their error plotted with asterisks.

by removing 50% of the sources with uniform randomized subsampling (Figures 4.10(b), 4.11(b) and 4.12(b)). We note that 50% randomized subsampling equates to 60 m periodic shot spacing on average, which is a frugal acquisition design. Using few upholes, which are inadequate to compensate for the near-surface variations, we apply elevation statics

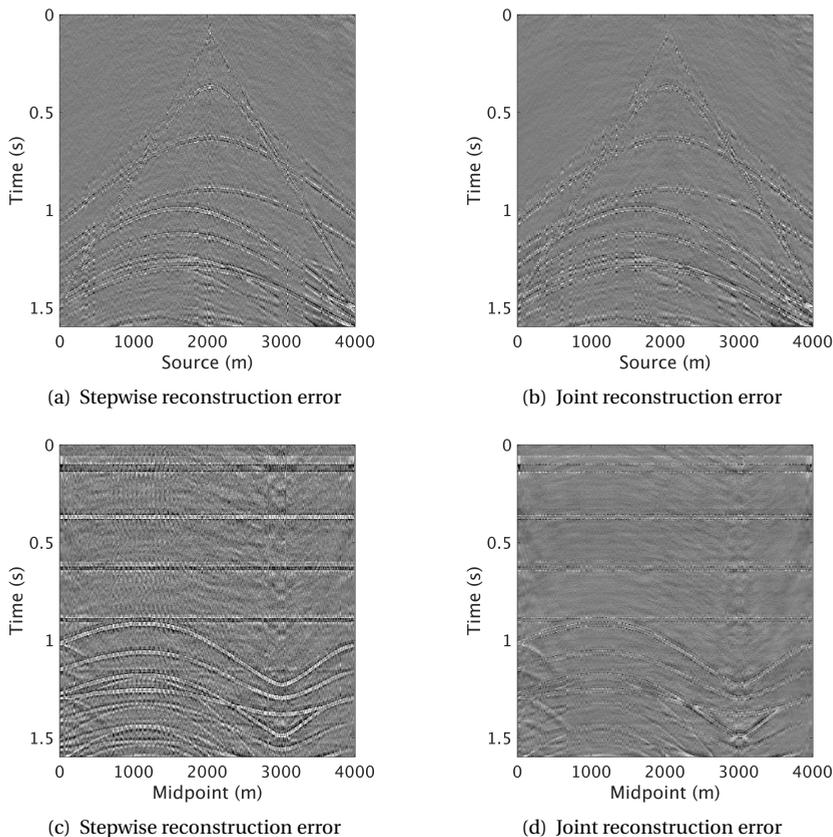


Figure 4.7: Residual amplitudes of common (a, b) receiver and (c, d) offset gathers after (a, c) stepwise and (b, d) joint reconstruction of 75% randomly subsampled data with statics. The reconstructed data are shown in Figure 4.6.

correction and bring the sources and receivers to a flat datum. From the stack of the subsampled data (Figure 4.12(a)), it is obvious that interpretation is challenging, which calls for data reconstruction.

To reconstruct the data, we apply rank-minimization-based interpolation without prior statics correction, which results in the data and stacks shown in Figures 4.10(c), 4.11(c) and 4.12(c). Interpolation in this situation provides an improvement compared to the subsampled data (Figures 4.10(b), 4.11(b) and 4.12(b)). However, due to not accounting for rapid variations in the weathering layers, the interpolated data exhibit more noise compared to the original data (Figures 4.10(a), 4.11(a) and 4.12(a)). Also, the events between 19 – 24 km and 0.4 – 0.6 s are distorted. To improve the results, we use the proposed reconstruction (Algorithm 5.1). The estimated statics for the displayed gathers are shown in Figure 4.10(e), while the estimated statics for the whole data are shown in Figure 4.12(e). The latter shows that larger statics are estimated from data between 19

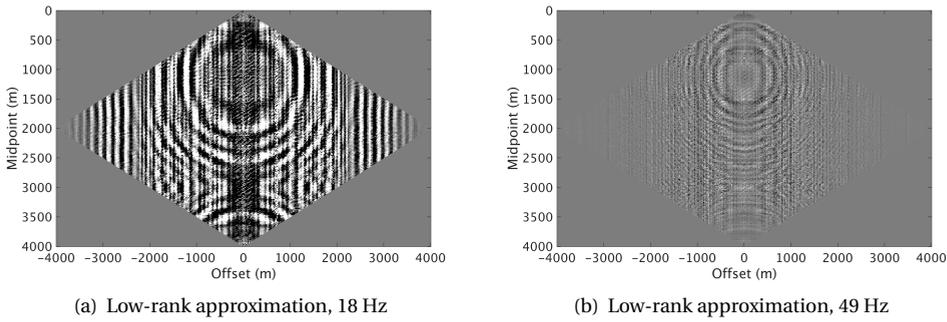


Figure 4.8: Low-rank approximation after partial statics correction of Figures 4.3(a) and 4.3(b), which we use as a starting solution \mathbf{X}_0 of equation 4.6 for interpolation at (a) 18 and (b) 49 Hz.

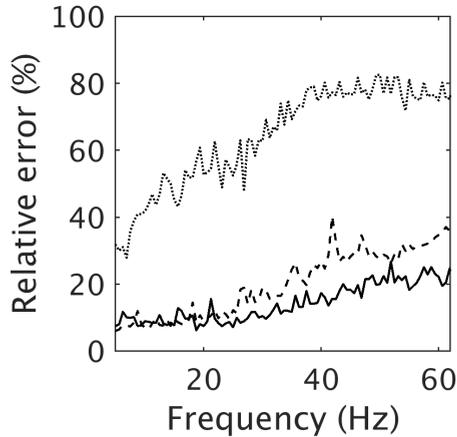


Figure 4.9: Error of reconstruction of the 75% randomly subsampled data with statics per frequency slice using interpolation (dotted curve), stepwise reconstruction (dashed curve) and joint reconstruction (solid curve).

and 24 km, which contain more complexity in the weathering layers compared to the data between 15 and 18 km (see Figure 4.12(a)). Reconstruction with our proposed method (Figures 4.10(d), 4.11(d) and 4.12(d)) reduces the noise compared to interpolation without statics correction. Moreover, it leads to higher power and more continuity of the events compared to the original data as demonstrated by the stack section in Figure 4.12(d). The improvement noticed on the stack between 19 and 24 km also coincides with the observation that the majority of the estimated statics (Figure 4.12(e)) belong to the same section.

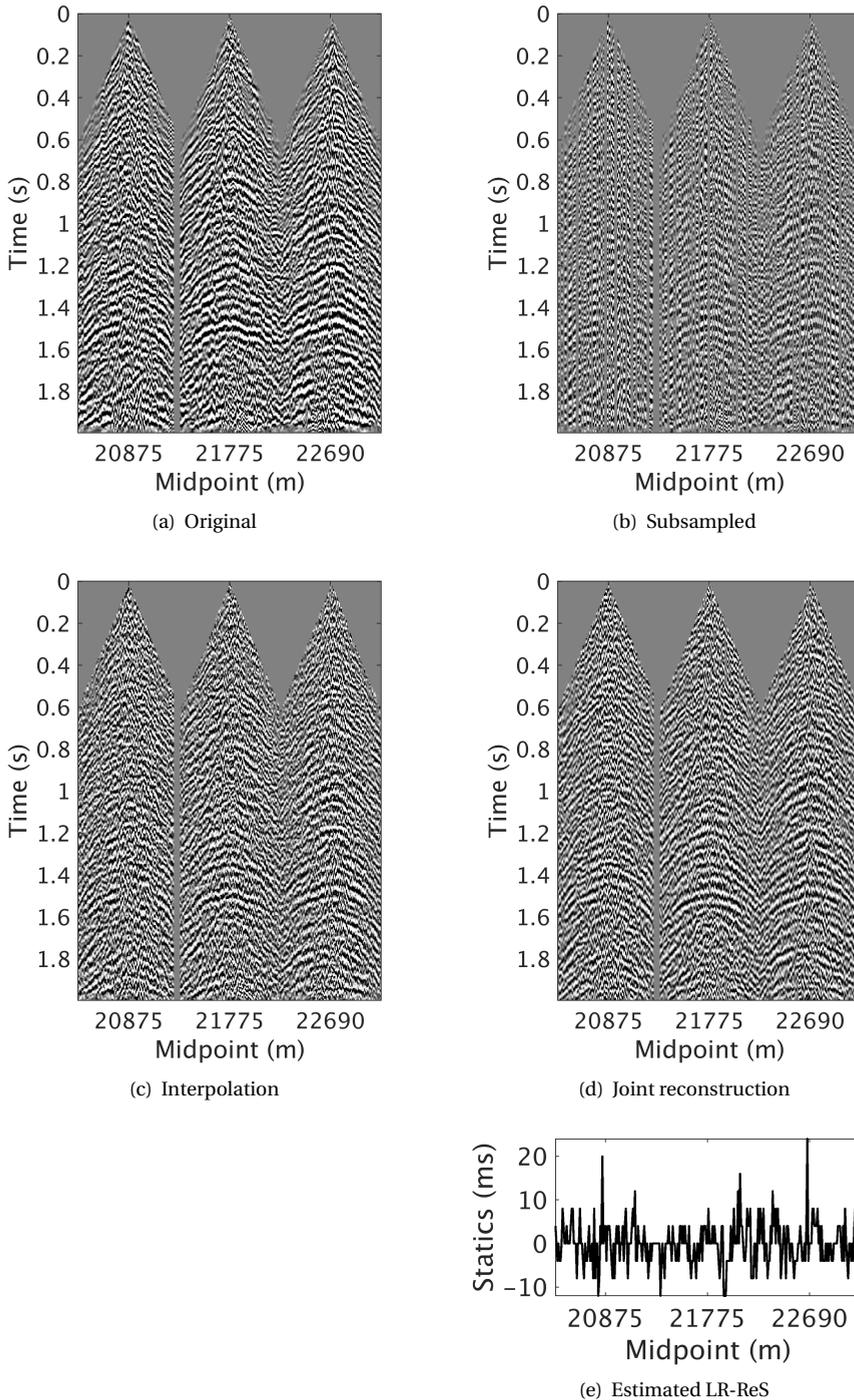


Figure 4.10: Part of the field data's common midpoints: (a) originally acquired data and data after (b) 50% randomized shots subsampling, (c) interpolation without statics correction and (d) joint reconstruction along with (e) the estimated statics of the displayed gathers.

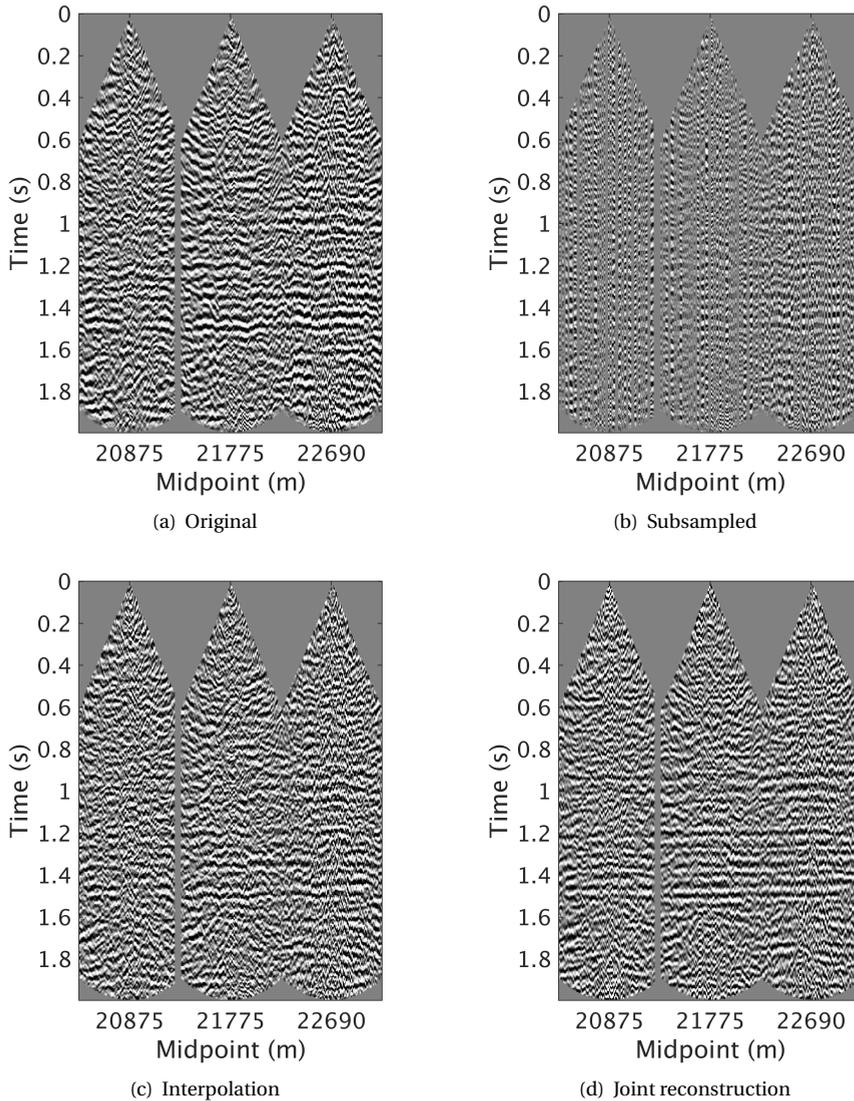


Figure 4.11: Part of the field data's common midpoints after NMO correction: (a) originally acquired data, and data after (b) 50% randomized shots subsampling, (c) rank-minimization interpolation and (d) our proposed joint reconstruction.

4.5. DISCUSSION

In this section, we discuss the use of data windowing, rank selection, computational efficiency, limitations and extensions.

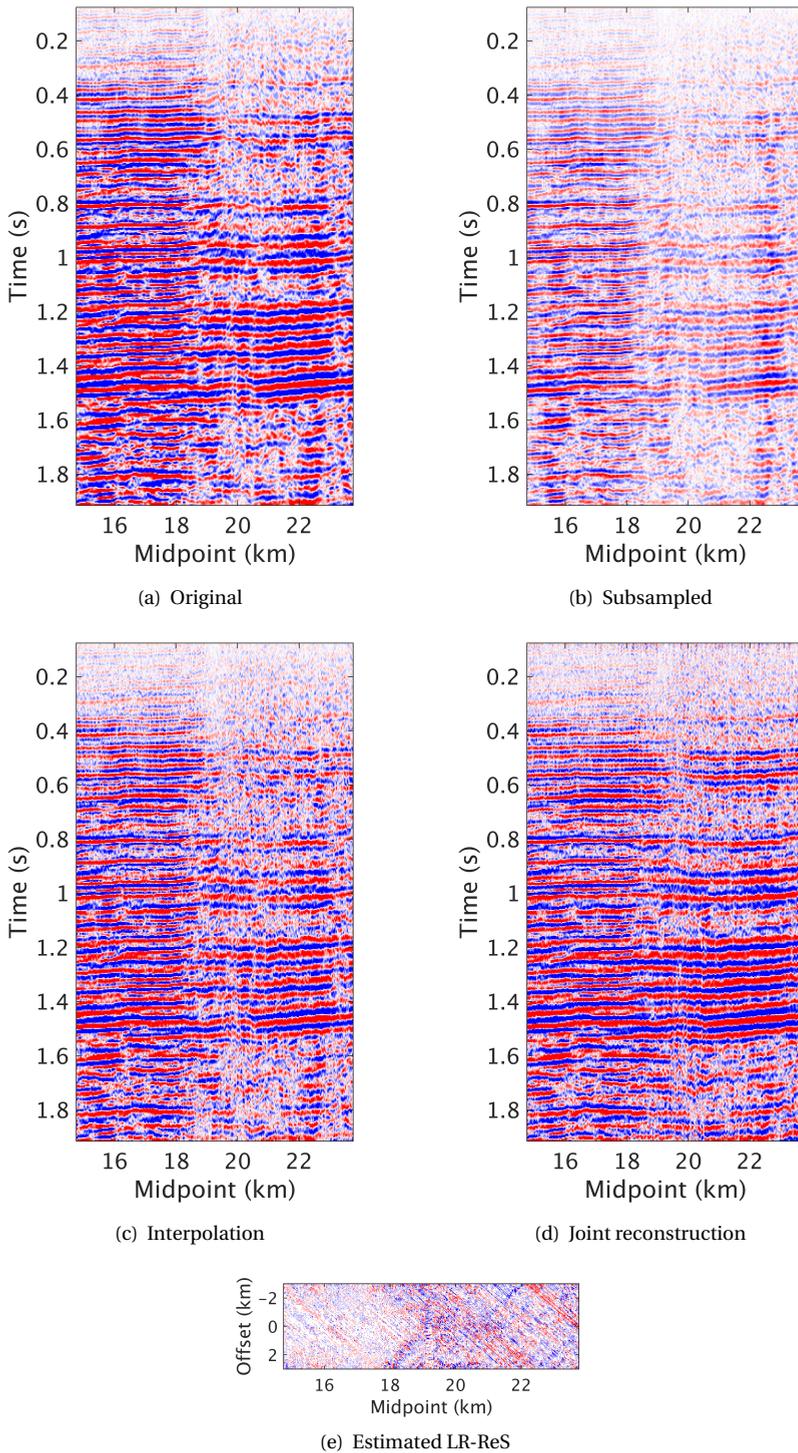


Figure 4.12: Field data stacks of the (a) original data and after (b) 50% randomized subsampling, (c) interpolation without statics correction, and (d) our proposed joint reconstruction along with (e) the estimated statics clipped to ± 12 ms.

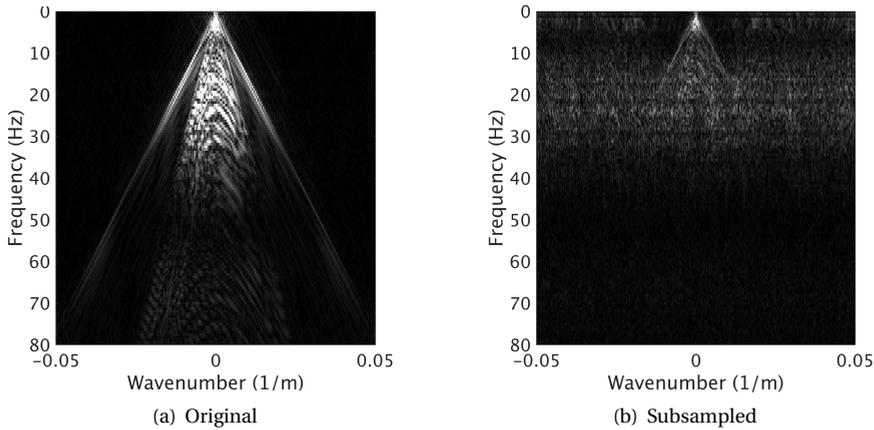


Figure 4.13: Frequency-wavenumber spectrum of a selected common receiver gather from the synthetic data: (a) statics-free and (b) after adding statics and removing 50% of the shots at random.

4.5.1. DATA WINDOWING

It is common for rank-based reconstruction techniques to rely on data windowing to ensure that the data are sufficiently of low rank nature. Moreover, the number of linearly dipping events in each window can be used as a guide for rank selection, see [12]. However, selection of the windowing parameters such as the window dimensions, number of windows, which can be overlapping or not, and how to merge back the separate windows are not trivial. In the presence of weathering layers, merging the windows after statics correction can be a hurdle as the separate windows will get assigned different statics to result in mis-ties along the horizontal and vertical directions. This could be the case if windowing is used to reconstruct the field data example, where the section between 15 and 18 km exhibits low influence of the weathering layers compared to the one between 19 and 24 km (Figure 4.12). To avoid the non-trivial task of data windowing for statics correction and interpolation, we use the midpoint-offset-frequency domain, where the data can be well-approximated by low-rank matrices. Therefore, we can successfully apply the proposed reconstruction techniques to the whole line without windowing, which is demonstrated by the synthetic and field data examples.

The other processes that requires data windowing is statics estimation, e.g. with stack power maximization [36]. Conventional methods require multiple-free data, as only primaries are aligned after NMO correction or migration. To lower that requirement, a window with primaries is usually selected. The proposed method does not need access to NMO corrected or migrated data. Therefore, data windowing is also not necessary for statics estimation.

To avoid cycle skipping, a limit on the maximum cross-correlation time lag is necessary. This parameter also depends on the desired maximum allowable time-shift, which relies on the complexity of the weathering layers. Since the statics in the synthetic data are large, we allow a maximum shift of 32 ms. The statics in the field data are lower than the

synthetic data. Therefore, we choose 12 ms as the maximum allowable time shift.

4.5.2. RANK SELECTION

The main parameters that need to be selected for Algorithm 5.1 are the ranks used for interpolation \mathbf{k} of all the frequencies and those used for multi-scale low-rank approximation \mathbf{K} . After partial statics correction, the singular values decay faster because the data exhibit higher coherency. Since we do not account for all the statics at the first iteration, there will remain incoherent energy, which is represented by the smaller singular values. Therefore, we can neglect singular values in the tail of the decay curve to capture the coherent energy, which we use for statics estimation and correction. The further the iterations progress, the faster the singular values will decay. Accordingly, we begin the iterations with a high-rank approximation and further decrease the rank as iterations progress. For interpolation, we need to preserve as much as possible of the signal, while for statics estimation, we only need a signal with less statics imprint that we use for cross-correlation. The latter tends to be captured by the few largest singular values. Since amplitude preservation is essential after interpolation, the ranks used for interpolation are higher than those needed for statics estimation. Since the high frequencies contain more variability compared to the low ones, they usually require higher ranks. Therefore, we suggest to start with low rank for low frequencies and linearly increase it with increasing frequency content for both interpolation [29] and near-surface correction [28]. User analysis is necessary for determination of the optimal ranks since complexity of the frequency slices also play a role.

We choose the lowest and highest ranks, which correspond to the lowest and highest frequencies of the synthetic data according to the above strategy. The rank values we use for interpolation and low-rank approximation at the first, second and third iterations are (15 - 75), (15 - 30), (5 - 15) and (3 - 5), respectively. We determine the rank values for the in-between frequencies by linearly increasing the rank between the lowest and highest ranks in each iteration. To quantify the influence of the rank selection, we use the stack power as a measure. The stack power increases by 202% and decreases by 13% compared to the densely-sampled data with statics and statics-free data, respectively. Increasing the rank values by 50% leads to almost the same results ($\pm 0.5\%$). On the other hand, decreasing and increasing the rank values by 50% result in 210% and -10% change in the stack power compared to the stacks of densely-sampled data with statics and statics-free data, respectively. Therefore, the sensitivity of the joint reconstruction to the rank selection of this synthetic data is low.

For the field data, we follow the same strategy to select the rank values. The lowest and highest rank values we use for interpolation and statics correction at the first, second and third iterations are (30 - 90), (5 - 15), (3 - 15) and (3 - 5). These values lead to 19% increase in the stack power after applying our proposed joint reconstruction compared to the densely-sampled data (compare Figures 4.12(a) and 4.12(d)). To examine the method's sensitivity for the rank selection, we test a number of scenarios outlined in Table 4.1. We notice that varying the rank values for interpolation has a higher impact on the stack power than for statics estimation. This conclusion is expected as the rank values used for interpolation explicitly estimate new data (phase and amplitude). On the other hand, the low-rank-based statics correction uses the rank values to estimate the statics. Moreover, the multi-scale approach implemented for statics correction alleviates the need

Interpolation ranks	Statics-correction ranks	Stack power
+25%	0%	+9%
0%	+25%	+17%
+25%	+25%	+9%
-25%	0%	+22%
0%	-25%	+21%
-25%	-25%	+19%

Table 4.1: Sensitivity of the joint reconstruction to rank selection of the field data. The change in the interpolation and statics correction rank values is with respect to the chosen values stated in section 4.5.2. The stack power column indicates the power of the estimated stack relative to the stack of densely-sampled data.

of using accurate rank values.

4.5.3. COMPUTATIONAL EFFICIENCY

Other than obtaining improved reconstruction, the proposed method provides better efficiency compared to stepwise reconstruction. The computational efficiency of the proposed reconstruction is related to the computational efficiency of the separate processes themselves. For LR-based near-surface estimation and correction, it is determined by the numbers of low-rank approximations $O(\min\{n_m^2 \times n_h, n_m \times n_h^2\})$, and cross-correlations $O(n_t \times n_g)$, where n_g is the number of cross-correlation lags, which turned out to be more efficient compared to residual statics estimation by stack power maximization [36]. Truncated SVD, where only a subset of the largest singular vectors and singular values are calculated, can speedup the low-rank approximation, which can be necessary for 3D large scale seismic data. The efficiency of rank-minimization-based interpolation, which operates on the whole line can be better than those that operate on separate gathers. To further enhance the computational efficiency, low-rank approximation and interpolation of frequency slices as well as the cross-correlation of traces can be performed in parallel over multiple cores.

The joint reconstruction approach further improves the computational efficiency. As we show in Figure 5.6, low-rank approximation can interpolate the randomly subsampled data displayed in Figures 4.3(a) and 4.3(b). However, this interpolation is poor as the approximation neglects data of importance. Nonetheless, it provides a good initial solution that we can use to reduce the number of required iterations to solve equation 4.3. Since low-rank approximation is a by-product of LR-based near-surface estimation and correction, we obtain the initial guess cost-free with our proposed joint reconstruction. Moreover, since near-surface estimation and correction is performed in the midpoint-offset domain, we also modify the interpolation such that the input measured data is used in the same domain (equation 4.6). Therefore, we reduce the requirement to go back and forth between the acquisition and midpoint-offset domains. Additionally, our model-independent approach requires less manual interaction compared to existing techniques, which usually need multiple iterations of velocity and statics estimation as they influence each others, followed by data interpolation of NMO-corrected data [18].

4.5.4. LIMITATIONS AND EXTENSIONS

Different near-surface conditions lead to different effects. Even for the shown field data example, the data between 15 and 18 km is less influenced by the near-surface weathering layers compared to the data between 18 and 24 km (Figure 4.12(a)). Therefore, interpolation of the former section can be achieved without the need to any statics correction (Figure 4.12(c)), which is not the case for the latter one that gets improved with our proposed method (Figure 4.12(d)). Even though the field data contain noise that we do not remove as part of the pre-conditioning, the proposed reconstruction is still robust. Further improvement could be obtained with prior noise attenuation or by adding a noise reduction step as part of the joint reconstruction scheme, since noise can influence the coherency of the data. In the used field data example, ground roll is aliased due to the coarse 30 m source and receiver intervals. Therefore, we remove it with a simple frequency-wavenumber filter before reconstruction as it is challenging to interpolate.

In terms of sampling schemes, we only consider randomized sampling. There are multiple ways to optimize the locations of sources and receivers, e.g. see [7] and recent studies by [37, 38]. However, there are other limiting factors for land data such as the challenging near-surface that needs to be considered during the acquisition design. When the data are periodically subsampled, rank-based interpolation becomes inapplicable because such sampling does not lead to low-rank structure destruction. This results in less complexity for the discussed reconstruction schemes compared to the case of randomly subsampled data as the low-rank structure destruction becomes influenced only by the near-surface. In this case, the proposed technique can be modified by replacing the rank-based interpolator with another one that is applicable for interpolation of periodically subsampled data. Similarly, other interpolation methods for randomly subsampled data, e.g. curvelet-, Fourier- or Radon-based methods, can replace the used rank-minimization interpolation, if desired.

Interpolation and statics correction can be performed on separate gathers, e.g. by applying sparsity promotion on receiver gathers in the frequency-wavenumber domain since statics and randomized subsampling decrease the sparsity (Figure 4.13(b)) that densely sampled statics-free data exhibit (Figure 4.13(a)). As noticeable, the low frequencies are less influenced by the subsampling and near-surface effects compared to the higher frequencies (Figure 4.13(b)). Therefore, a similar scheme to Algorithm 5.1 can be implemented, where the statics estimated from subsampled and interpolated low frequencies and applied to the full-band data allow for statics estimation and reconstruction of the higher frequencies as they improve their coherency. However, operating on the full line in the midpoint-offset domain allows us to exploit structural relationships between the different common midpoints, which cannot be achieved using separate gathers. We note that interpolation of separate CMP gathers still allows to exploit the structural relationships up to a certain extent compared to separate shot or receiver gathers. However, the separate CMP gathers are not linked in this case. Therefore, there might be variability across the line, which is minimized with our proposed approach.

To further extend the proposed reconstruction to 3D data, Algorithm 5.1 can be implemented, but finding a suitable transform-revealing domain is essential. Luckily, there are different options that can be investigated to matricize the 3D data into 2D matrices, which include the midpoint-offset domain along the x- and y-directions, the source and receiver

coordinates along the x- and y-directions or the x- and y-coordinates of the sources and receivers. It is essential that the chosen domain reveals the necessary properties that allow for low-rank approximation and rank-minimization-based interpolation.

4.6. CONCLUSIONS

Complex near-surface regimes with rapid variations in properties of the weathering layers degrade the quality of subsurface models of interest. The addition of acquisition gaps, e.g. to reduce the acquisition costs, exacerbates the challenge. To obtain accurate subsurface models, data reconstruction becomes essential.

We examine four reconstruction strategies. Interpolation without accounting for the near-surface effects fails to provide satisfactory results. Alternatively, the conventional reconstruction approach with surface-consistent short-wavelength statics correction followed by interpolation requires distinction between primaries and multiples and NMO velocity estimation, which are challenging to perform with subsampled data influenced by the weathering layers.

Using a model-independent low-rank-based near-surface correction followed by rank-minimization interpolation improves the reconstruction. However, it results in noisy and erroneous reconstruction due to imperfections in the estimated statics. These imperfections are due to the combined influence of weathering and subsampling, which reduce the coherency and destroy the low-rank structure.

To overcome the aforementioned limitations, we propose to reconstruct the data with simultaneous rank-reduction based near-surface correction and interpolation. With this approach, near-surface correction and interpolation improve each others performance, as data with improved coherency from one is fed into the other. We additionally increase the computational efficiency by avoiding repetitive transformations between the source-receiver and midpoint-offset domains and by using low-rank approximated data (cost-free product from the LR-based near-surface correction) as initial solution to the rank-minimization optimization problem. The application of our proposed method to randomly subsampled synthetic data and a challenging field data set affected by complex weathering layers demonstrates its potential.

4.7. ACKNOWLEDGMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his Ph.D. studies. We also would like to thank Saudi Aramco for providing the field data.

REFERENCES

- [1] T. H. Kebo and P. G. Kelamis, *Focus on land seismic technology: The near-surface challenge*, *The Leading Edge* **31** (2012), pp. 62–68, <https://doi.org/10.1190/1.3679329>.
- [2] D. L. Donoho, *Compressed sensing*, *IEEE Transactions on information theory* **52** (2006), pp. 1289–1306, <https://doi.org/10.1109/TIT.2006.871582>.
- [3] E. J. Candès and M. B. Wakin, *An Introduction To Compressive*

- Sampling*, IEEE signal processing magazine **25** (2008), pp. 21–30, <http://dx.doi.org/10.1109/MSP2007.914731> .
- [4] F. J. Herrmann, M. P. Friedlander, and O. Yilmaz, *Fighting the curse of dimensionality: compressive sensing in exploration seismology*, IEEE Signal Processing Magazine **29** (2012), pp. 88–100, <http://dx.doi.org/10.1109/MSP2012.2185859> .
- [5] P. M. Zwartjes and M. D. Sacchi, *Fourier reconstruction of nonuniformly sampled, aliased seismic data*, Geophysics **72** (2007), pp. V21–V32, <http://geophysics.geoscienceworld.org/content/72/1/V21.full.pdf> .
- [6] G. Hennenfent and F. J. Herrmann, *Application of stable signal recovery to seismic data interpolation*, in *SEG Technical Program Expanded Abstracts 2006* (2006) pp. 2797–2801, <https://library.seg.org/doi/pdf/10.1190/1.2370105> .
- [7] G. Hennenfent and F. J. Herrmann, *Simply denoise: Wavefield reconstruction via jittered undersampling*, Geophysics **73** (2008), pp. V19–V28, <http://geophysics.geoscienceworld.org/content/73/3/V19.full.pdf> .
- [8] A. M. Alfaraj, F. Herrmann, and R. Kumar, *Reconstruction of s-waves from low-cost randomized and simultaneous acquisition by joint sparse inversion*, in *SEG Technical Program Expanded Abstracts 2017* (2017) pp. 2533–2538, <https://library.seg.org/doi/pdf/10.1190/segam2017-17681376.1> .
- [9] M. N. Kabir and D. Verschuur, *Restoration of missing offsets by parabolic radon transform1*, Geophysical Prospecting **43** (1995), pp. 347–368, <http://dx.doi.org/10.1111/j.1365-2478.1995.tb00257.x> .
- [10] D. O. Trad, T. J. Ulrych, and M. D. Sacchi, "accurate interpolation with high-resolution time-variant radon transforms", Geophysics **67** (2002), pp. 644–656, <http://dx.doi.org/10.1190/1.1468626> .
- [11] S. Trickett, L. Burroughs, A. Milton, L. Walton, and R. Dack, *Rank-reduction-based trace interpolation*, in *80th Annual International Meeting, SEG, Expanded Abstracts* (2010) pp. 3829–3833, <http://dx.doi.org/10.1190/1.3513645> .
- [12] V. Oropeza and M. Sacchi, *Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis*, Geophysics **76** (2011), pp. V25–V32, <https://doi.org/10.1190/1.3552706> .
- [13] N. Kreimer and M. D. Sacchi, *A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation*, Geophysics **77** (2012), pp. V113–V122, <https://doi.org/10.1190/geo2011-0399.1> .
- [14] A. Aravkin, R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, *Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation*, SIAM Journal on Scientific Computing **36** (2014), pp. S237–S266, <https://doi.org/10.1137/130919210> .

- [15] B. Recht, M. Fazel, and P. Parrilo, *Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization*, *SIAM Review* **52** (2010), pp. 471–501, <https://doi.org/10.1137/070697835>.
- [16] B. Recht and C. Ré, *Parallel stochastic gradient algorithms for large-scale matrix completion*, *Mathematical Programming Computation* **5** (2013), pp. 201–226, <https://doi.org/10.1007/s12532-013-0053-8>.
- [17] Y. Zhang, S. Sharan, and F. J. Herrmann, *High-frequency wavefield recovery with weighted matrix factorizations*, in *89th Annual International Meeting, SEG, Expanded Abstracts* (Society of Exploration Geophysicists, 2019) pp. 3959–3963, <https://doi.org/10.1190/segam2019-3215103.1>.
- [18] D. Trad, *Five-dimensional interpolation: Recovering from acquisition constraints*, *Geophysics* **74** (2009), pp. V123–V132, <https://doi.org/10.1190/1.3245216>.
- [19] A. Stanton, N. Kazemi, and M. D. Sacchi, *Processing seismic data in the presence of residual statics*, in *SEG Technical Program Expanded Abstracts 2013* (Society of Exploration Geophysicists, 2013) pp. 1838–1842, <https://doi.org/10.1190/segam2013-1453.1>.
- [20] A. Gholami, *Phase retrieval through regularization for seismic problems*, *Geophysics* **79** (2014), pp. V153–V164, <https://doi.org/10.1190/geo2013-0318.1>.
- [21] A. Gholami, *Residual statics estimation by sparsity maximization*, *Geophysics* **78** (2013), pp. V11–V19, <https://doi.org/10.1190/geo2012-0035.1>.
- [22] C. Mosher, C. Li, Y. Ji, F. D. Janiszewski, B. Bankhead, L. Williams, J. Hand, and J. Anderson, *Compressive seismic imaging: Moving from research to production*, in *SEG Technical Program Expanded Abstracts 2017* (2017) pp. 74–78, <https://library.seg.org/doi/pdf/10.1190/segam2017-17679803.1>.
- [23] M. T. Taner, F. Koehler, and K. A. Alhilali, *Estimation and correction of near-surface time anomalies*, *Geophysics* **39** (1974), pp. 441–463, <https://doi.org/10.1190/1.1440441>.
- [24] R. A. Wiggins, K. L. Larner, and R. D. Wisecup, *Residual statics analysis as a general linear inverse problem*, *Geophysics* **41** (1976), pp. 922–938, <https://doi.org/10.1190/1.1440672>.
- [25] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999) <https://doi.org/10.1190/1.9781560801818>.
- [26] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <https://library.seg.org/doi/pdf/10.1190/1.9781560801580>.
- [27] A. M. Alfaraj, E. Verschuur, and F. J. Herrmann, *Residual statics correction without NMO — A rank-based approach*, in *First International Meeting for Applied Geoscience & Energy Expanded Abstracts* (2021) pp. 2565–2569, <https://doi.org/10.1190/segam2021-3583455.1>.

- [28] A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, *Low-rank-based residual statics estimation and correction*, *Geophysics* **88** (2023), pp. V215–V231, <https://doi.org/10.1190/geo2022-0415.1>.
- [29] R. Kumar, C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, *Efficient matrix completion for seismic data reconstruction*, *Geophysics* **80** (2015), pp. V97–V114, <https://doi.org/10.1190/geo2014-0369.1>.
- [30] J. W. Thorbecke and D. Draganov, *Finite-difference modeling experiments for seismic interferometry*, *Geophysics* **76** (2011), pp. H1–H18, <https://doi.org/10.1190/geo2010-0039.1>.
- [31] G. H. Golub and C. Reinsch, *Singular value decomposition and least squares solutions*, in *Linear algebra* (Springer, 1971) pp. 134–151.
- [32] E. v. Berg and M. P. Friedlander, *Probing the Pareto frontier for basis pursuit solutions*, *SIAM Journal on Scientific Computing* **31** (2008), pp. 890–912, <https://doi.org/10.1137/080714488>.
- [33] A. Y. Aravkin, J. V. Burke, and M. P. Friedlander, *Variational properties of value functions*, *SIAM Journal on Optimization* **23** (2013), pp. 1689–1717, <https://doi.org/10.1137/120899157>.
- [34] J. D. M. Rennie and N. Srebro, *Fast maximum margin matrix factorization for collaborative prediction*, in *Proceedings of the 22nd international conference on Machine learning*, ICML '05 (ACM, New York, NY, USA, 2005) pp. 713–719, <https://doi.org/10.1145/1102351.1102441>.
- [35] M. Al-Ali and D. Verschuur, *An integrated method for resolving the seismic complex near-surface problem*, *Geophysical Prospecting* **54** (2006), pp. 739–750, <https://doi.org/10.1111/j.1365-2478.2006.00575.x>.
- [36] J. Ronen and J. F. Claerbout, *Surface-consistent residual statics estimation by stack-power maximization*, *Geophysics* **50** (1985), pp. 2759–2767, <https://doi.org/10.1190/1.1441896>.
- [37] Y. Guo and M. D. Sacchi, *Data-driven time-lapse acquisition design via optimal receiver-source placement and reconstruction*, in *90th Annual International Meeting, SEG, Expanded Abstracts* (2020) pp. 66–70, <https://doi.org/10.1190/segam2020-3426928.1>.
- [38] Y. Zhang, M. Louboutin, A. Siahkoohi, Z. Yin, R. Kumar, and F. J. Herrmann, *A simulation-free seismic survey design by maximizing the spectral gap*, in *Second International Meeting for Applied Geoscience & Energy* (2022) pp. 15–20, <https://doi.org/10.1190/image2022-3751690.1>.

5

3D MODEL-INDEPENDENT NEAR-SURFACE ESTIMATION AND CORRECTION

Due to limitations of acquisition and model-building algorithms, estimation of accurate near-surface models that capture their rapidly varying nature is challenging. Therefore, near-surface correction with short-wavelength statics is routinely used to minimize the near-surface influence on subsurface reflectors prior to model building. However, existing short-wavelength statics estimation techniques rely on a subsurface velocity model, e.g. NMO velocity, which is also affected by the weathering layers. As a result, iterative velocity and statics estimation are necessary to improve their performance, which are efforts- and time-consuming, particularly for 5D data. Moreover, many pre-processing steps can be necessary to prepare the data for the velocity and statics estimation stage.

To the aforementioned challenges, we extend the model-independent rank-based near-surface estimation and correction to 5D data. In chapters 2, 3 and 4, we use the midpoint-offset transform domain to reveal desirable rank structures. In the proper matricization domain, where we organize 5D data volumes into matrices, we show that the data exhibit low-rank structures, which get destroyed when influenced by incoherency due to the rapid variations in the weathering layers. Accordingly, we promote the low-rank structure to estimate statics that improve the data coherency in an iterative and multi-scale approach, similar to the 2D situation.

Since we do not rely on a velocity model and consequently bypass the associated pre-processing steps, we avoid the need of iterative velocity and statics estimation. At the same time, the proposed method requires minimal pre-processing and can estimate accurate surface- and non-surface-consistent statics at once, which typically require two statics estimation steps.

This chapter is a modified version of the manuscript undergoing review "A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, 3D model-independent rank-based near-surface correction, Geophysics".

These benefits, in addition to being computationally efficient, make the proposed method favorable compared to existing methods. We demonstrate its performance on 5D synthetic and field data. A potential application of our proposed method is short-wavelength near-surface correction prior to full waveform inversion of land data.

5.1. INTRODUCTION

5D seismic data collected onshore with sources and receivers at the surface can be influenced by the loosely compacted weathering layers near the surface. The rapid variations in topography and properties of the weathering layers can lead to undesired time-shifts of subsurface reflectors commonly referred to as 'short-wavelength statics'. One of the main reasons to treat the near-surface variations with 'statics' is the limited capability of delineating accurate and high-resolution near-surface models. Despite the advancements in model-building technology, short-wavelength statics estimation and correction remains one of the most essential steps for land data [1] that is still necessary even prior to full waveform inversion of land data, e.g. see also [2].

Similar to 2D data, short-wavelength statics estimation and correction of 3D data can be performed with data-driven methods in a surface-consistent manner [3–6], where the time-shifts depend only on the positions of sources and receivers without considering their offsets, which is a simplification of the actual situation [7, 8]. Moreover, these surface-consistent as well as existing non-surface-consistent methods normally rely on the availability of a subsurface model, e.g. NMO or migration velocity and NMO-corrected or migrated gathers [9], horizon picking and pilot trace construction [10, 11]. In addition to requiring multiple pre-processing steps and being efforts and time consuming, these processes when containing errors can lead to inaccurate statics estimation. In such cases, iterative subsurface model and statics estimation become essential, which exacerbate the already laborious situation, especially when considering 3D data.

To overcome the aforementioned impediments for 2D data, [12, 13] use a data-driven rank-reduction-based approach to estimate short-wavelength statics without the need of NMO velocity and NMO-corrected gathers. The method transforms the data to the midpoint-offset domain in order to reveal the low-rank structure without data windowing and use low-rank approximation for statics estimation. The method shares the same principles with other rank-based methods that are proposed for example for interpolation [14–18], deblending [19–21], denoising [22–26], and residual statics correction [27, 28], where ideal seismic data are of low-rank nature, but non-ideal data result in slow decay of the singular values. However, these methods may differ in the choice of the transform domain, the use of 5D data and data windowing. For statics correction of 3D data, the 2D midpoint-offset transform can be applied on separate 2D lines. Nevertheless, obtaining 2D lines from sparse and irregularly sampled 3D data may require interpolation to fill-in the missing traces, which may perform poorly when the data is affected by the weathering layers. Therefore, the midpoint-offset domain may not be optimally suitable for statics estimation and correction of 3D data.

5.1.1. CONTRIBUTIONS

To avoid the dependency of statics estimation on the subsurface model and the associated processing steps commonly need prior to model estimation, which may include multiple reflections analysis or removal, denoising and data reconstruction, we propose a novel 3D model-independent rank-reduction-based near-surface estimation and correction. To compute the singular value decomposition (SVD) of 5D data volumes with 1 temporal and 4 spatial dimensions, which is necessary for low-rank approximation, we need to perform matricization of the 5D data, i.e. organization of the 5D data into matrices. At the same

time, it is essential that the chosen organization domain reveals the underlying low-rank structure. Therefore, we first analyze different matricization domains that can be used to organize the 5D data. We show that — in the proper domain — the near-surface weathering layers render coherent energy incoherent, which results in slowly decaying singular values compared to the statics-free data that are of low-rank nature. According to the findings, we describe the details of our proposed algorithm. In short, it performs low-rank approximation to obtain coherent 3D data, which we cross-correlate with the original data to estimate the statics. We then perform statics correction to update the original data. Since accurate rank selection for low-rank approximation can influence the results, we use an iterative and multi-scale approach, which improves the statics estimation and alleviates the need of accurate rank selection. Since migrated or NMO corrected gathers are no longer required, we bypass the need of the multiple iterations of velocity and statics estimation that are routinely applied with the conventional approach. The proposed method can also estimate both surface- and non-surface-consistent statics at once, which circumvents the necessity of separate surface- and non-surface consistent short wavelength statics estimation steps. We demonstrate the performance of the method on 3D synthetic and field data.

5

5.2. MATRICIZATION DOMAIN

Due to rapid variations of topography and properties of the near-surface weathering layers, seismic data get influenced by short-wavelength statics that result in distorted subsurface models [6]. To enhance the subsurface models using a rank-based approach, we first need to find a suitable domain that captures the weathering layers effect on the rank-structure of 5D seismic data. While using small windows — even though not trivial [18] — may do the purpose, merging the windows back together can lead to mis-ties as the different windows will have different statics correction. Therefore, we search for a domain that allows for short-wavelength statics estimation and correction independently of the subsurface model without the need of data windowing.

Organization of 5D seismic data volumes into 2D matrices is necessary to compute the SVD, an essential component of rank-based methods. Let $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$ be a frequency slice, its SVD can be computed with [29]:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \quad (5.1)$$

where $\mathbf{S} \in \mathbb{R}^{k \times k}$ is a block diagonal matrix of non-negative real-valued singular values $\mathbf{S} = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k)$, where $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_k \geq 0$ and $k \leq \min\{n_1, n_2\}$ is the rank of the matrix. $\mathbf{U} \in \mathbb{C}^{n_1 \times k}$ and $\mathbf{V} \in \mathbb{C}^{n_2 \times k}$ are orthogonal matrices that hold the left \mathbf{u} and right \mathbf{v} singular vectors, respectively. H denotes the Hermitian transpose.

We explore two possible domains that will determine the dimensions n_1 and n_2 of \mathbf{X} . We will denote (i) and (ii) the two domains: (i) source-receiver domain and (ii) x - y domain. In the source-receiver domain, sources ($x_{\text{src}}, y_{\text{src}}$) and receivers ($x_{\text{rec}}, y_{\text{rec}}$) along the x - and y -directions are placed along two different dimensions of the matrix, where n_1 and n_2 become the number of shots and receivers. In the x - y domain, coordinates of sources and receivers along the x -direction ($x_{\text{src}}, x_{\text{rec}}$) are placed on one dimension of the matrix, e.g. n_1 and the sources and receivers along the y -direction ($y_{\text{src}}, y_{\text{rec}}$) are placed on the other dimension of the matrix. The potential domain should provide a ma-

trix composed of close to linearly-dependent columns when it has coherent energy. In this case, the singular values decay rapidly, which allows for accurate low-rank approximation, where $k \ll \min\{n_1, n_2\}$. On the other hand, the columns of matrices influenced by the weathering layers, which render coherent energy incoherent, should become less linearly-dependent as the low-rank structure gets destroyed to result in slow singular values decay.

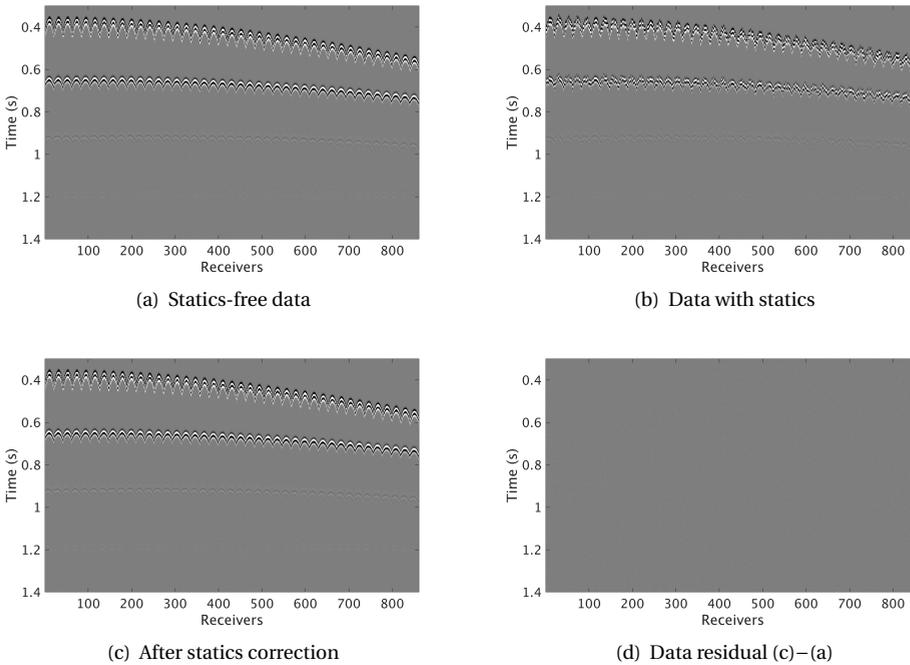


Figure 5.1: 3D shot gathers: (a) statics-free, (b) statics-contaminated, (c) after near-surface correction with our proposed method and (d) its error.

5.2.1. SIMULATED DATA

To observe the matricization domain effect on the rank structure, we use 5D synthetic data with primaries and internal multiples for illustration (Figure 5.1(a)). The model is composed of two reflectors, a flat one and dipping one with conflicting dips along the inline and crossline directions (Figures 5.2(a) and 5.3(a)). The data are collected with sources and receivers spaced at 25m in the inline direction and 50m in the crossline direction. To mimic rapid variations in surface elevation and properties of the near-surface weathering layers, we timeshift the data with timeshifts in the range of ± 24 ms (Figure 5.1(b)). These statics contain surface- and non-surface-consistent components. As a result, the stack power and resolution of the stacks decrease to result in high discrepancy compared to the statics-free stacks (Figures 5.2(b), 5.2(c), 5.3(b), 5.3(c)). From frequency slices corresponding to these two data sets, we analyze their singular values decay in the different

matricization domains.

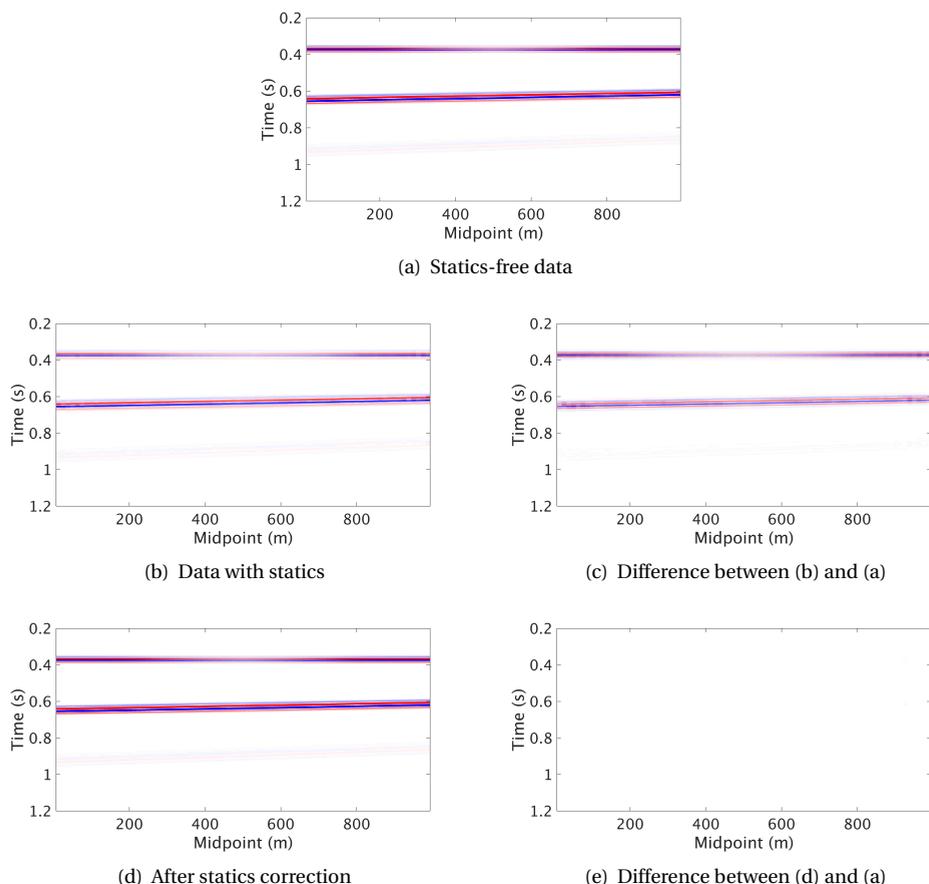


Figure 5.2: Inline stack sections of (a) statics-free data, (b) data with statics and (d) after statics-correction with our proposed method. (c) and (e) contain the error of (b) and (d), respectively. All plots are made at the same amplitude level.

Data matricization (*i*) (Figures 5.4(a) and 5.4(b)) shows that the data are highly varying. Therefore, the columns and rows of the matrix are far from linearly dependent, i.e. the data are not of low-rank nature as the singular values are not rapidly decaying (Figure 5.4(e)). Consequently, low-rank approximation will yield inaccurate data due to the significance of a large portion of the singular values. At the same time, the singular values decay curve of data affected by the weathering layers (Figures 5.4(c) and 5.4(d) and 5.4(e)) shows a different trend compared with that of statics-free data, which makes it hard to estimate the latter from the former. Therefore, there is no desirable rank-structure that we can make use of in the source-receiver domain.

On the other hand, data matricization (*ii*) in the x - y domain shows that the statics-free data are well-structured with lower variability amongst the rows and columns of the

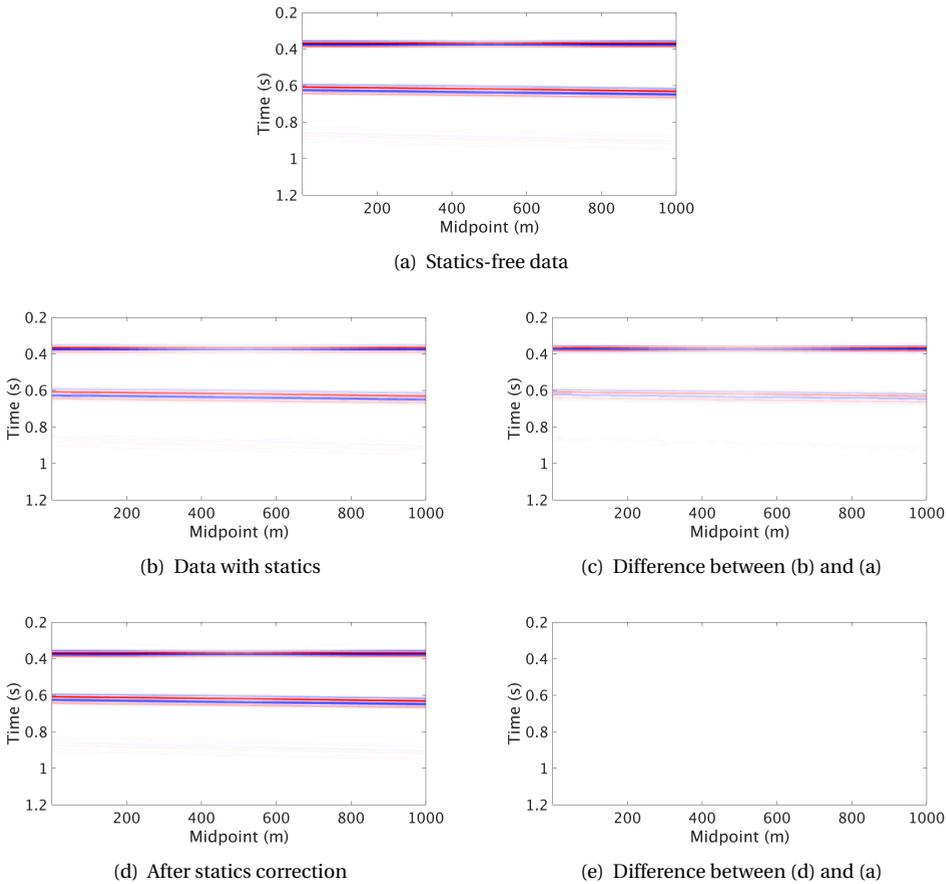


Figure 5.3: Crossline stack sections of (a) statics-free data, (b) data with statics and (d) after statics-correction with our proposed method. (c) and (e) contain the error of (b) and (d), respectively.

matrix (Figures 5.5(a) and 5.5(b)) compared with matricization in the source-receiver domain (Figures 5.4(a) and 5.4(b)). As a result, the singular values decay rapidly in the x - y domain (Figure 5.5(e)) to allow for accurate low-rank approximation. In the presence of weathering layers, the columns and rows of the matrix become less linearly-dependent (Figures 5.5(c) and 5.5(d)) compared with the statics-free situation (Figures 5.5(a) and 5.5(b)). Consequently, the singular values decay becomes slower (Figure 5.5(e)). Therefore, matricization in the x - y domain satisfies the requirements to process the data with a rank-based approach. We note that this domain is also used for rank-minimization-based interpolation [18].

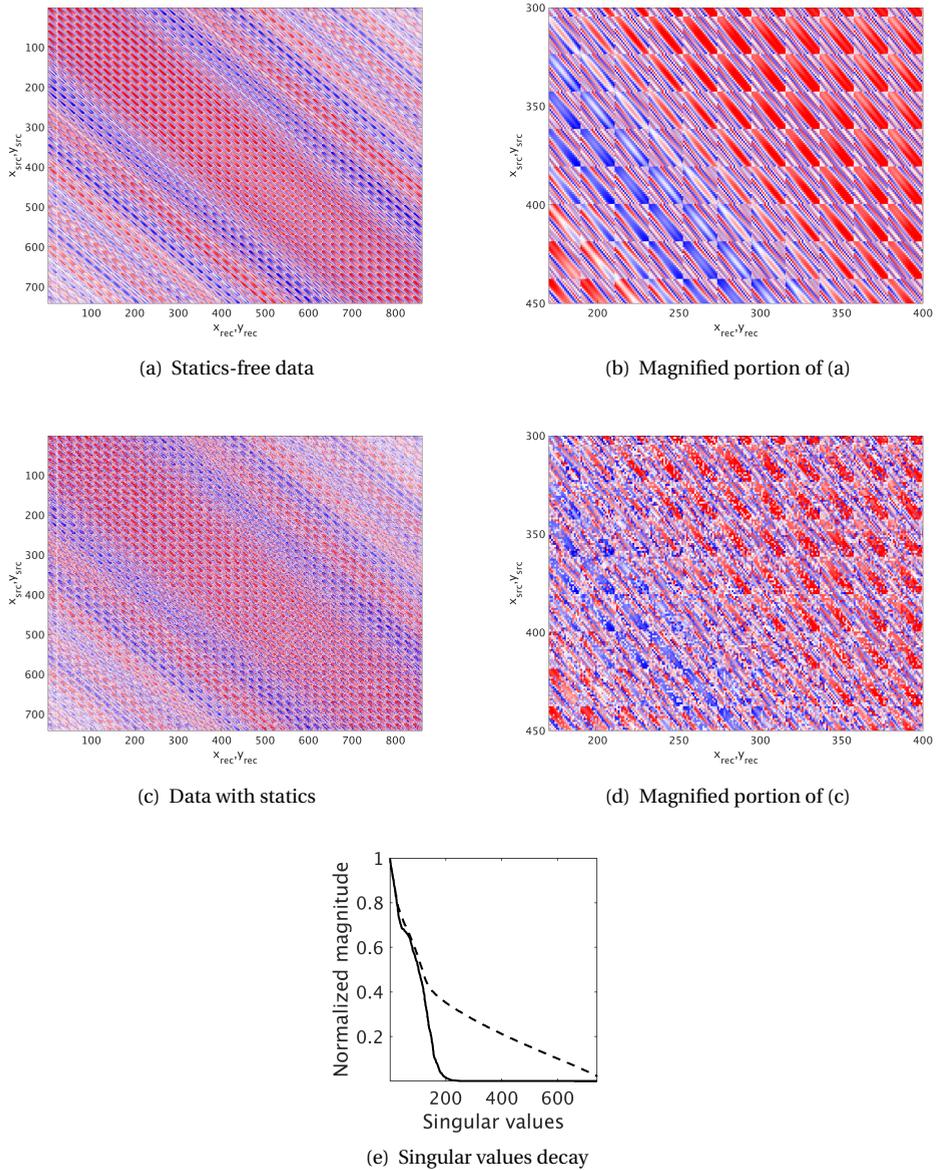


Figure 5.4: Data organization (i) in the source-receiver domain. 30 Hz frequency slices of (a) statics-free data and (c) data with statics. (b) and (d) are magnified parts of (a) and (c), respectively. (e) The singular values of (a) and (c) plotted with solid and dashed lines, respectively.

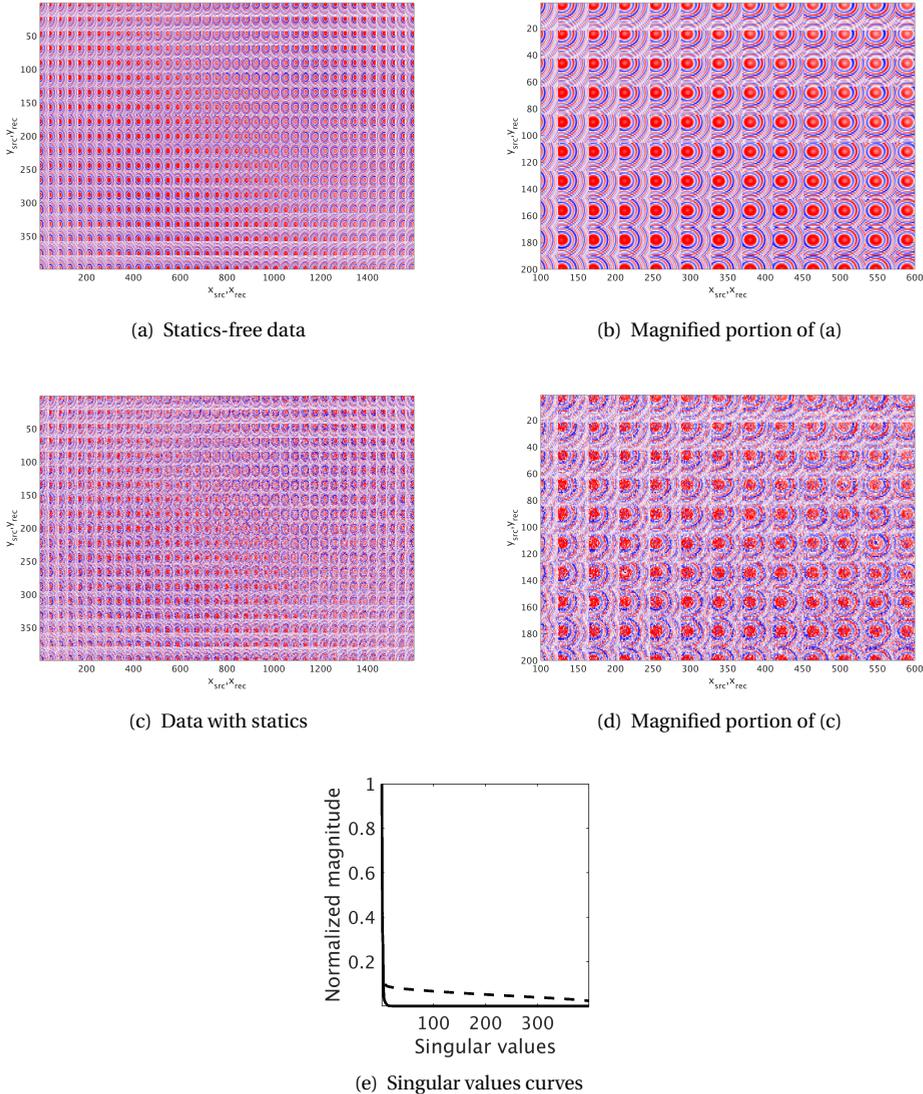


Figure 5.5: Data organization (ii) in the x - y domain. 30 Hz frequency slices of (a) statics-free data, (c) data with statics. (b) and (d) are magnified parts of (a) and (c), respectively. (e) The singular values of (a) and (b) plotted with solid and dashed lines, respectively.

5.3. METHODOLOGY

Since ideal 3D data in the x - y domain are of low-rank nature, whereas the weathering layers result in slow singular values decay (Figure 5.5), we can promote the low-rank structure to estimate data without the near-surface influence. In principle, the goal is to estimate

coherent energy captured by the largest singular values and their corresponding singular vectors. To do so, we need to solve the following rank minimization problem:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X} - \mathbf{Y}\|_F \quad \text{subject to} \quad \text{rank}(\mathbf{X}) \leq k, \quad (5.2)$$

where \mathbf{Y} and $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$ are the observed and estimated frequency slices. The dimensions of the matrices in the x - y domain become $n_1 = n_{y_{\text{src}}} \times n_{y_{\text{rec}}}$ and $n_2 = n_{x_{\text{src}}} \times n_{x_{\text{rec}}}$, which we will refer to as n_y and n_x , respectively, to avoid notation clutter. $\|\cdot\|_F$ is the Frobenius norm defined as:

$$\|\mathbf{X}\|_F = \sqrt{\sum_{i=1}^{n_y} \sum_{j=1}^{n_x} \mathbf{X}^{(ij)^2}}, \quad (5.3)$$

which is equivalent to the ℓ_2 norm of a vector and $\mathbf{X}^{(ij)}$ are the ij -elements of \mathbf{X} . We can solve equation 5.2 with SVD followed by thresholding the singular values. Using the condition $k \ll \min\{n_y, n_x\}$, we restrict the solution to be of low-rank nature [30]. However, thresholding the singular values and the associated singular vectors may result in removing important data. At this stage, we do not require that amplitudes are preserved. We use low-rank approximation as an intermediate solution that can provide coherent data with less near-surface influence, although possibly with erroneous amplitudes (Figure 5.6). To estimate the statics, we cross-correlate the low-rank approximated data \mathbf{D}_{lr} with the original data $\mathbf{D}_{xy} \in \mathbb{R}^{n_t \times n_x \times n_y}$ in the t - x - y domain to estimate the static (time-shift) corresponding to the maximum cross-correlation coefficient:

$$t_{xy} = \underset{t}{\text{argmax}} (\mathbf{d}_{xy}(t) \star \mathbf{d}_{\text{lr}}(t)), \quad (5.4)$$

where \mathbf{d}_{xy} and $\mathbf{d}_{\text{lr}} \in \mathbb{R}^t$ are traces extracted from \mathbf{D}_{xy} and \mathbf{D}_{lr} , respectively. \star indicates cross-correlation, which measures the similarity between two signals with a relative delay g :

$$\phi = \sum_t \mathbf{d}_{xy}(t) \mathbf{d}_{\text{lr}}(t + g). \quad (5.5)$$

In our case, the delay that provides the highest cross-correlation coefficient corresponds to the static, which makes the two signals of high similarity. To avoid cycle skipping, a limit on the maximum delay is necessary, which also depends on the complexity of the weathering layers. For statics correction, we shift each trace with the estimated static to update the data:

$$\hat{\mathbf{d}}_{xy} = \mathbf{d}_{xy}(t + t_{xy}), \quad (5.6)$$

where $\hat{\mathbf{d}}_{xy}$ is the trace after statics correction. We apply equations 5.4 – 5.6 to all the traces to estimate the statics \mathbf{T}_{xy} and statics-corrected data $\hat{\mathbf{D}}_{xy}$ of all the traces, while preserving the data amplitudes.

However, the approximation of high frequencies with low rank is known to perform poorly due to their highly varying nature. To improve that, we use a multi-scale approach to process the data using frequency bands starting from low to high frequencies, while exploiting the relationships amongst the different frequency bands [13]. At low frequencies, the influence of short-wavelength statics is low, which allows for accurate low-rank approximation and statics estimation. We use these statics for statics correction of the full-band data as the different frequency bands share common statics. By doing so, we

improve the coherency of the higher frequencies, which improves their low-rank approximation. However, statics estimated from the low frequencies are not accurate enough for the high frequencies. Therefore, we update the statics when including the higher frequencies during statics estimation.

Rank selection is another hurdle for rank-based methods. When using data windowing, the number of linearly dipping events in each window can be used as a guide to select the rank [15]. However, as mentioned earlier, mis-ties when merging back adjacent windows can be a problem along the spatial and temporal dimensions when the windows are corrected with different statics. On the other hand, the x - y domain allows us to avoid the limitations of data windowing, which we demonstrate in section 5.2. Nevertheless, rank selection remains a challenge. Since we only use low-rank approximated data for cross-correlation during statics estimation, we limit the influence of the rank compared to other rank-based methods, e.g. rank-based denoising and interpolation, where the rank is used explicitly to compute the output. Another approach to lower the influence of the rank selection is to use an iterative multi-rank-scale approach [13]. After statics correction, frequency slices exhibit higher coherency, which leads to faster singular values decay. Consequently, we can use a lower rank for low-rank approximation at the next iteration. Therefore, we start the iterations with higher rank, and reduce it at the next iterations. As for rank selection for frequency slices within each rank-scale iteration, it depends on the complexity of each frequency slice. Low frequencies have lower variability, which requires lower rank for accurate low-rank approximation compared to higher frequencies that require higher rank as they are more variable. Using the aforementioned guidelines, rank selection becomes feasible. Since the outputs are the statics and statics-corrected data, the results can be examined as part of the quality control process. In section 5.5, We perform sensitivity analysis for selection of the rank values.

Algorithm 5.1 describes the details of our proposed method. The input are the 5D data \mathbf{D} volumes, the rank-scales \mathbf{K} , which contain the ranks for each frequency slice f and scale l , the frequency bands \mathbf{f}_b at which we estimate the statics, and the minimum f_{\min} and maximum f_{\max} usable frequencies. We start with matricization of the data in the x - y domain, followed by their Fourier transform to the frequency domain (steps 1 and 2 of Algorithm 5.1). We then perform low-rank approximation after computing the SVD (steps 5 and 6) within the loop over frequencies and rank scales. When we reach the desired frequency band for statics estimation (step 7), we inverse Fourier transform the low-rank approximated data \mathbf{D}_{lr} and original data \mathbf{D}_{xy} to the time domain (step 9), followed by their cross-correlation for statics estimation from the largest cross-correlation coefficients (step 10). Using these statics, we perform statics correction (step 11) and update the original data \mathbf{D}_{xy} with statics-corrected data $\hat{\mathbf{D}}_{xy}^{(l)}$ at the current iteration. In the next section, we demonstrate the proposed method's performance.

5.4. RESULTS

Using our proposed method, we show that we can estimate and correct for the near-surface weathering layers effect on synthetic and field data without the need of either knowledge of the subsurface model or data windowing.

Algorithm 5.1: 3D rank-based near-surface correction

Input: $\mathbf{D} \in \mathbb{R}^{n_t \times n_{x_{\text{rec}}} \times n_{y_{\text{rec}}} \times n_{x_{\text{src}}} \times n_{y_{\text{src}}}}$, $\mathbf{K} \in \mathbb{R}^{n_f \times n_l}$, $\mathbf{f}_b \in \mathbb{R}^{n_{f_b}}$, $f_{\min} \in \mathbb{R}$, $f_{\max} \in \mathbb{R}$
Output: $\widehat{\mathbf{D}}_{xy} \in \mathbb{R}^{n_t \times n_x \times n_y \times n_l}$, $\mathbf{T}_{xy} \in \mathbb{R}^{n_y \times n_x \times n_l \times n_{f_b}}$

- 1 Matricize \mathbf{D} to the x - y domain $\mathbf{D}_{xy} \in \mathbb{R}^{n_t \times n_x \times n_y}$
- 2 fft \mathbf{D}_{xy} to $\widetilde{\mathbf{D}}_{xy} \in \mathbb{C}^{n_y \times n_x \times n_f}$
- 3 **for** $l \leftarrow 1$ **to** n_l **do**
- 4 **for** $f \leftarrow f_{\min}$ **to** f_{\max} **do**
- 5 Calculate SVD for $\widetilde{\mathbf{D}}_{xy}^{(f)}$ with equation 5.1
- 6 $\widetilde{\mathbf{D}}_{lr}^{(f)} \leftarrow \sum_{j=1}^{\mathbf{K}^{(f,l)}} s^{(j)} \mathbf{u}^{(j)} \mathbf{v}^{(j)H}$, equation 5.2
- 7 **if** f is contained within \mathbf{f}_b **then**
- 8 $i \leftarrow$ frequency band index
- 9 ifft $\widetilde{\mathbf{D}}_{lr}$ to \mathbf{D}_{lr} , ifft $\widetilde{\mathbf{D}}_{xy}$ to \mathbf{D}_{xy}
- 10 $\mathbf{T}_{xy}^{(l,i)} \leftarrow \mathcal{C}(\mathbf{D}_{xy}, \mathbf{D}_{lr})$, equations 5.4 and 5.5
- 11 $\widehat{\mathbf{D}}_{xy}^{(l)} \leftarrow \tau(\mathbf{D}_{xy}, \mathbf{T}_{xy}^{(l,i)})$, equation 5.6
- 12 $\mathbf{D}_{xy} \leftarrow \widehat{\mathbf{D}}_{xy}^{(l)}$, fft \mathbf{D}_{xy} to $\widetilde{\mathbf{D}}_{xy}$

5.4.1. SYNTHETIC DATA EXAMPLE

To validate the performance of the proposed method, we apply Algorithm 5.1 to the synthetic data described in the Simulated data subsection (Figures 5.1(b), 5.4(c), 5.4(d), 5.5(c), 5.5(d)). We use three iterations and estimate the statics at three frequency bands that contain low, mid and high frequencies. The estimated statics after three iterations are shown in Figure 5.7. While most of the statics are estimated at the first iteration, they are fine-tuned at the second and third iterations. Note that these statics contain surface- and non-surface-consistent elements, which we estimate at the same time without the need of two different processes. To determine the performance of the proposed method, we subtract the total estimated statics (Figure 5.7(d)) from the true statics (Figure 5.7(e)), which shows that the error is minimal (Figure 5.7(f)). To observe the effect on the data, we plot the estimated frequency slices and their error with respect to the statics-free data (Figure 5.8). After the first iteration, the estimated frequency slices are improved (Figures 5.8(a) and 5.8(b)), but are still far from the statics-free data (Figures 5.8(c) and 5.8(d)). After the third iteration (Figures 5.8(e) and 5.8(f)), the frequency slices are improved to become similar to the statics-free data (Figures 5.5(c) and 5.5(d)) as we correct for the majority of the statics. The error at this stage is minimal (Figures 5.8(g) and 5.8(h)). The estimated data in time-domain (Figure 5.1(c)) demonstrate the improvement we obtain in contrast to the data we start with (Figure 5.1(b)). The difference between the statics-free (Figure 5.1(a)) and estimated gathers (Figure 5.1(c)) displayed in Figure 5.1(d) is zero. To further evaluate the results, we examine the stack sections.

When the subsurface velocity is accurate, the gathers should stack in phase. However, the near-surface weathering layers result in lower stack power (Figure 5.2(b)) and therefore higher error (Figure 5.2(c)) compared to the statics-free stack as the gathers stack out of

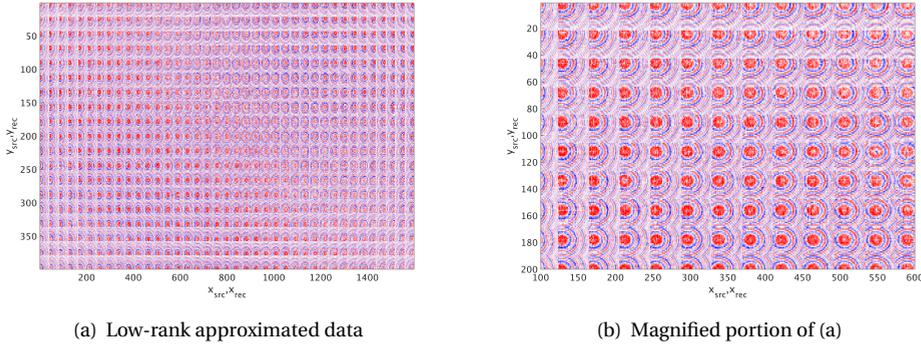


Figure 5.6: 30 Hz frequency slices of (a) Low-rank approximated data. (b) Magnified portion of (a).

phase. Therefore, one of the reasons to estimate and apply short-wavelength statics is to increase the stack's power in data processing. After applying our proposed method, the stack power increases (Figure 5.2(d)), which results in minimal error (Figure 5.2(e)). This confirms the potential of the proposed method.

5.4.2. FIELD DATA EXAMPLE

We test the method's performance on the Stratton 3D field data collected in South Texas. The data is acquired with source and receiver spacings of 220×880 and 110×1320 ft along the inline and crossline directions, respectively (Figure 5.9(a)). Arrays of 12 receivers and 4 vibrators are used at each receiver and shot point, respectively. The dataset is contaminated with spurious noisy traces and surface waves as can be seen from a raw shot gather in Figure 5.9(b). To prepare the data, we only apply simple frequency-wavenumber (FK) filter to attenuate surface waves and focus on the reflections (Figure 5.9(c)). Figure 5.10(a) shows the data in the x - y domain before statics correction. After we apply our proposed algorithm described in section 5.3, we obtain the gathers displayed in Figure 5.10(b). The short-wavelength statics due to the rapid variations of the weathering layers' properties noticeable in Figure 5.10(a) disappear in Figure 5.10(b). Note that we do not require any knowledge of the subsurface, e.g. velocity model or primary-multiple distinction, to reach to these results.

The improvement is also noticeable from the stack sections. After applying our proposed method, the stack exhibits 25% higher power compared to the stack before statics correction (Figures 5.11(a) and 5.11(b)). The increase in the stack power, which is due to the improved continuity of the reflections after statics correction, is more visible when comparing the magnified stack sections shown in Figures 5.12(a) and 5.12(b).

5.5. DISCUSSION

In this section, we discuss the advantages, sensitivity of the rank selection and limitations of the proposed method.

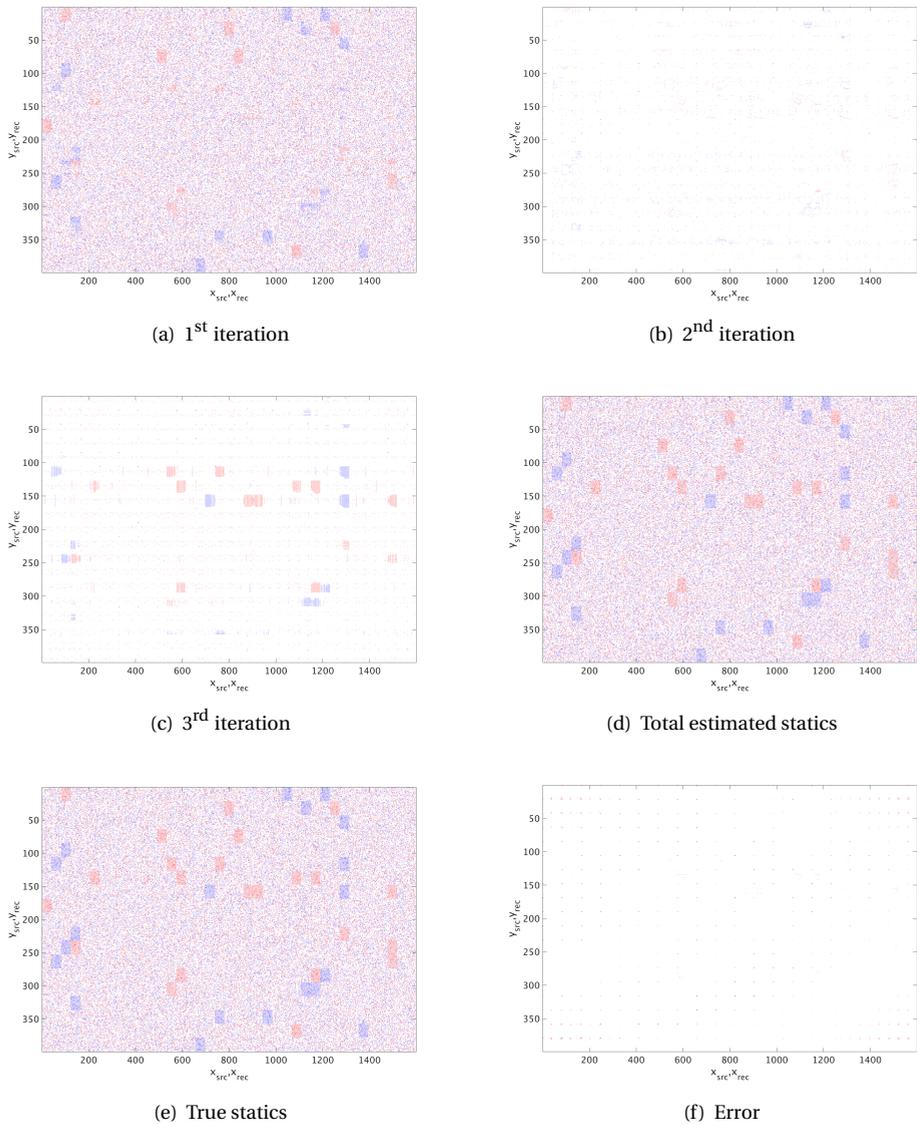
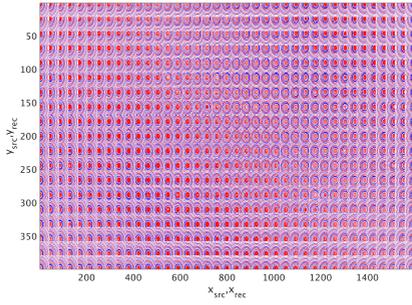
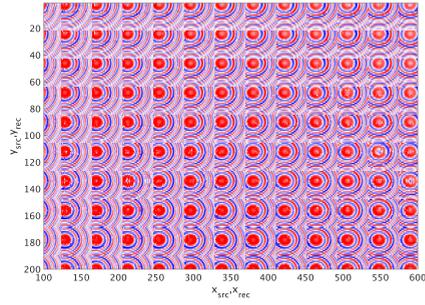
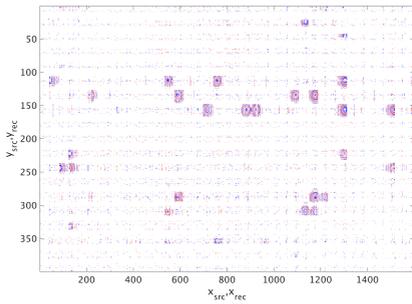
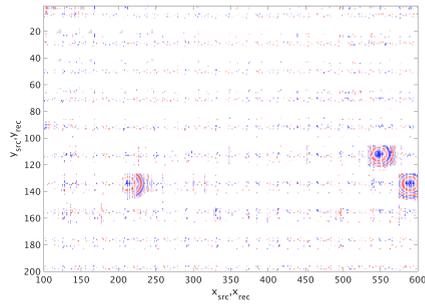


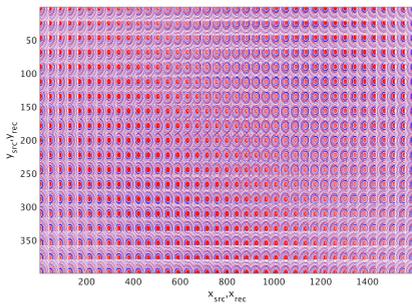
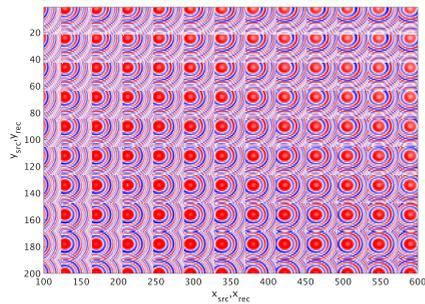
Figure 5.7: The estimated statics in the x - y domain after applying our proposed method at (a) first, (b) second and (c) third iterations. (e) The total estimated statics, (d) true statics and (f) the proposed method's error.

(a) Statics-corrected data, 1st iteration

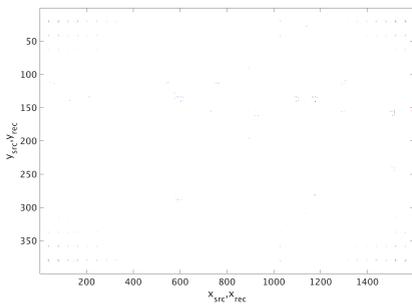
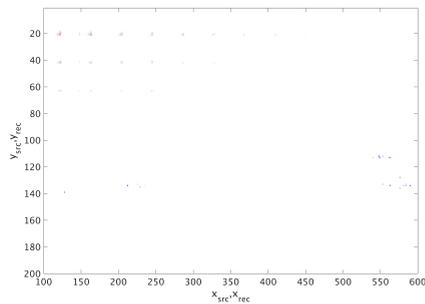
(b) Magnified portion of (a)

(c) Error after the 1st iteration

(d) Magnified portion of (c)

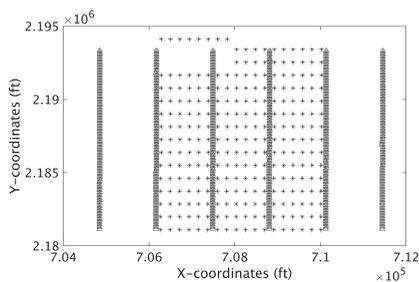
(e) Statics-corrected data, 3rd iteration

(f) Magnified portion of (e)

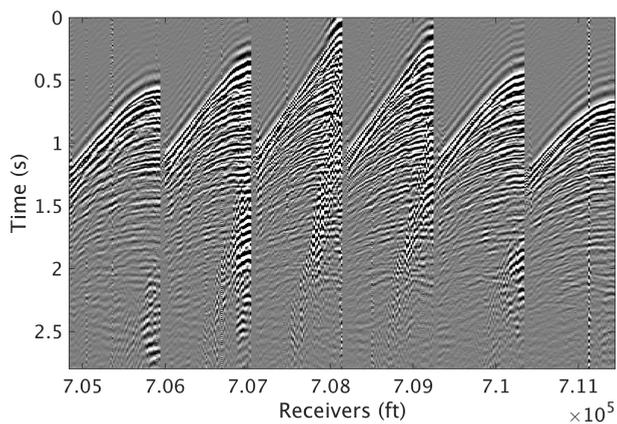
(g) Error after the 3rd iteration

(h) Magnified portion of (g)

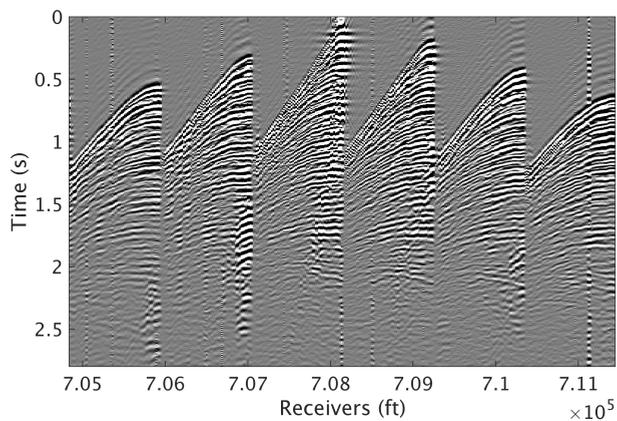
Figure 5.8: 30 Hz frequency slices of (a, b, e, f) statics-corrected data and (c, d, g, h) their error with respect to the statics-free data after the (a - d) first and (f - h) third iterations.



(a) Acquisition geometry

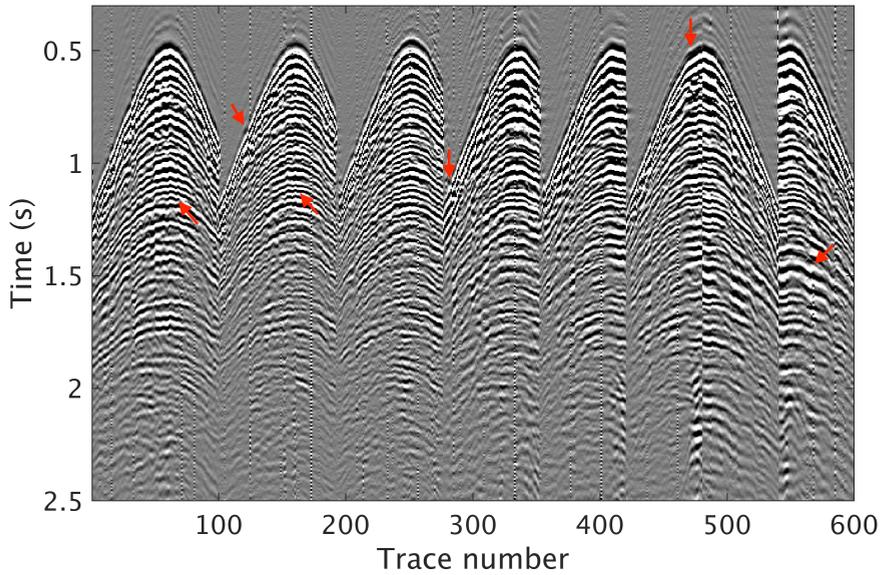


(b) Raw shot gather

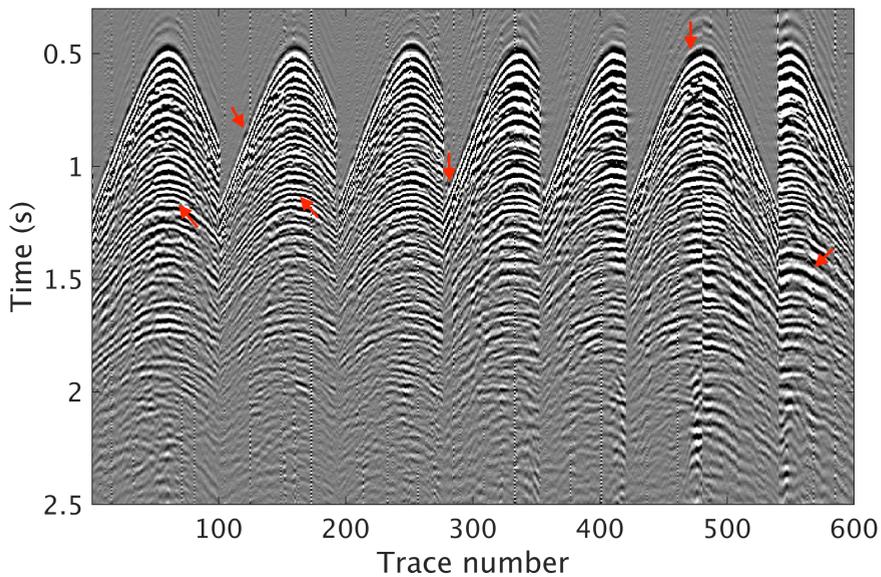


(c) Shot gather after FK filter

Figure 5.9: (a) Acquisition geometry of the field data, where triangles indicate receivers (6 lines) and asterisks indicate sources (a patch of $X \times Y$ sources). (b) Raw shot gather and (c) after FK filter for ground roll attenuation.



(a) Before statics correction



(b) After statics correction

Figure 5.10: Field data in the x - y domain (a) before and (b) after statics correction. The arrows indicate the improvements after statics correction.

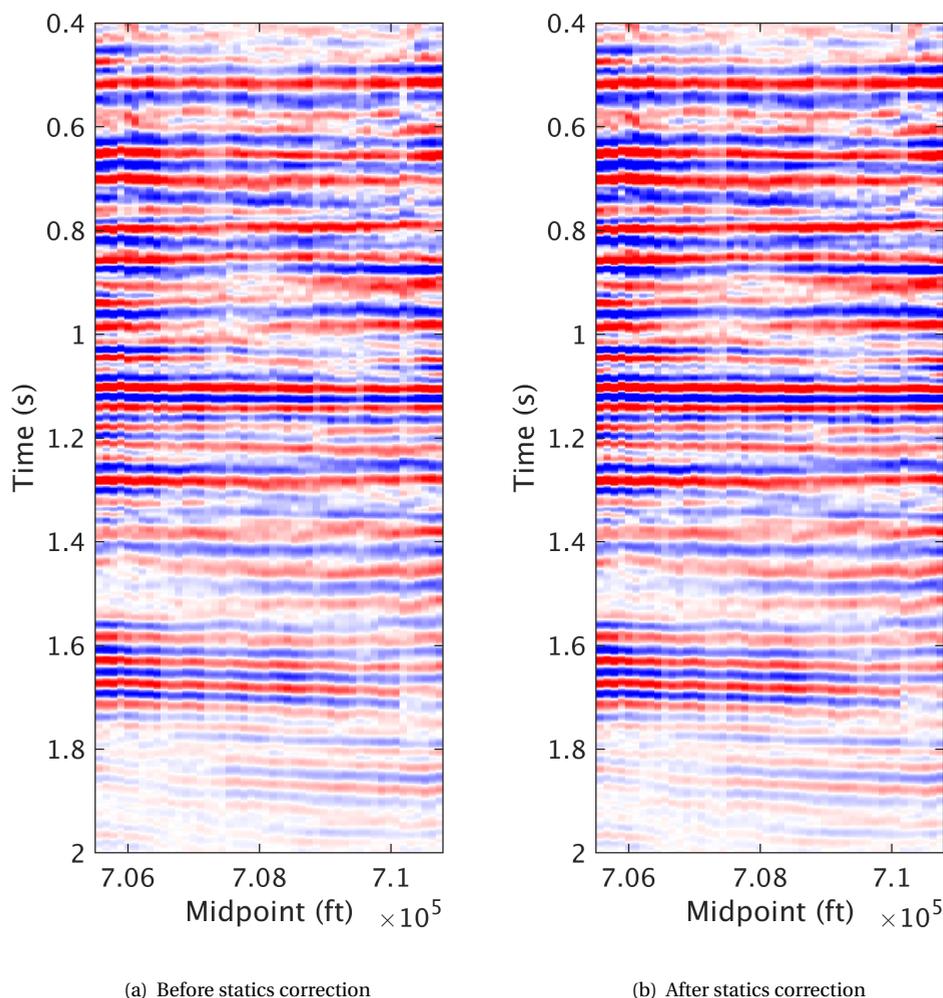


Figure 5.11: Stack sections of field data (a) before and (b) after near-surface correction.

5.5.1. ADVANTAGES

We demonstrate in the Results section that our proposed method corrects for rapid variations in properties of the weathering layers without the need of a subsurface velocity model using 3D synthetic and field data. The method estimates surface- and non-surface-consistent statics at the same time, which is usually achieved with at least two processes using conventional methods. With existing techniques, short-wavelength statics estimation and correction require many processing steps that include velocity estimation, multiples and primaries distinction and denoising. On the other hand, applying the proposed

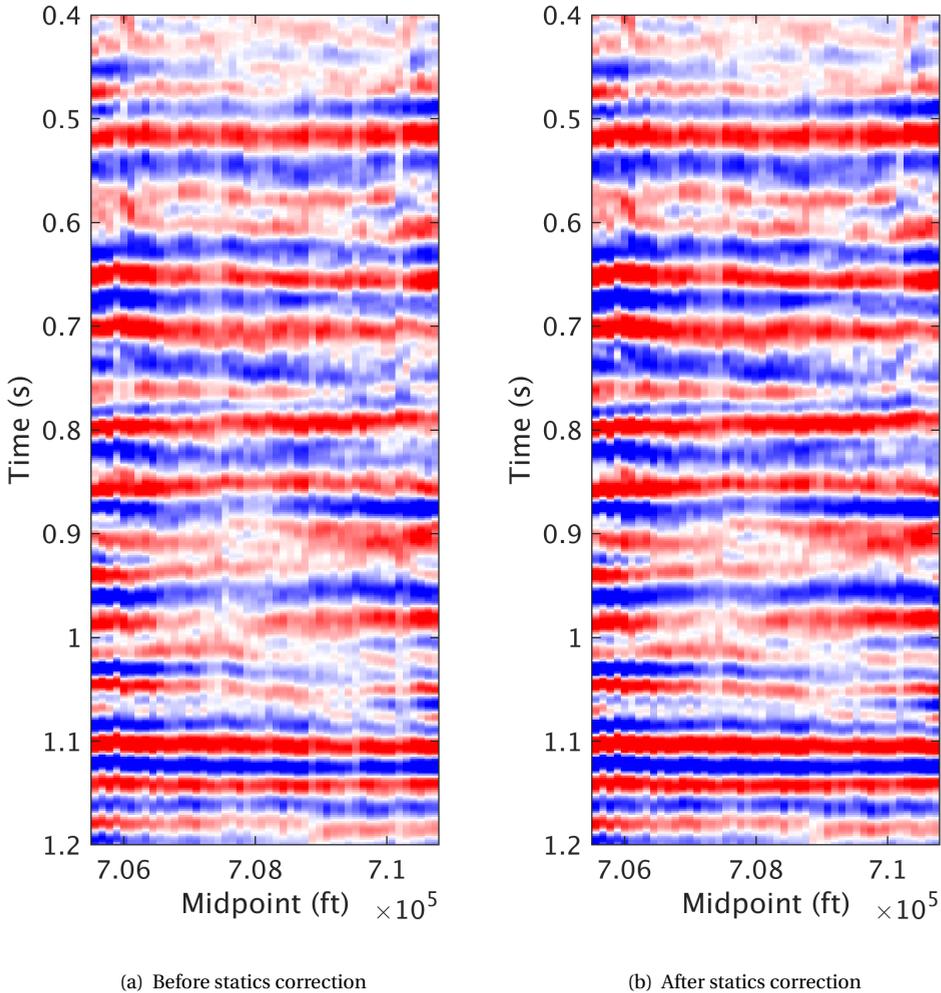


Figure 5.12: Magnified portions of the field data stack sections (a) before and (b) after near-surface correction.

method on this field dataset only requires simple FK filter to focus on the reflected wavefields. Moreover, the proposed method is computationally efficient. The aforementioned benefits make the proposed method favorable for short-wavelength statics estimation and correction. Potential applications include applying our proposed method prior to full waveform inversion of land data, where short-wavelength statics correction is an essential step [1].

For 3D data processing workflows, the proposed method can be applied at an early

stage, which has the potential to improve the performance of other processing methods such as denoising. Moreover, velocity estimation becomes no longer influenced by short-wavelength statics, which avoids the repetitive iterations of velocity and short-wavelength statics estimation. As a result, time- and efforts-consuming processes can be avoided.

5.5.2. SENSITIVITY TO RANK SELECTION

We perform sensitivity analysis of the proposed method with respect to the rank selection. Given that low-rank approximation is only used for statics estimation, the proposed method is less sensitive for the rank selection compared to other rank-based methods that use rank values for explicit estimation of new data, e.g. denoising and interpolation. The multi-rank-scale approach further lowers the requirements of selecting accurate rank values.

To quantify the influence of the rank selection, we use the stack power as it reflects how well the data are aligned. Higher stack power means better alignment, while lower stack power result in poor alignment. For the synthetic data, the rank values for the minimum and maximum frequency slices at the first, second and third iterations are (10–30), (6–20) and (2–5), respectively. The rank values for the in-between frequencies are obtained by linearly increasing the rank between the minimum and maximum value. The selected rank values result in 286% increase in the stack power compared to the stack of data with statics (Figure 6.4). Increasing the rank values by 50% result in 2% decrease in the power of the estimated stack, while decreasing the rank values by 50% lead to a 2% increase in the stack power.

For the Stratton field data, the selected rank values that correspond to the minimum and maximum frequency slices at the first and second iterations are (30–60) and (5–40), respectively. The rank values are higher for the field dataset due to its higher complexity compared to the synthetic dataset. Similar to the synthetic data, to obtain the rank for the rest of the frequencies, we linearly increase the rank between the minimum and maximum value. The stack power after applying our proposed method using the above ranks lead to 25% increase in the stack power. The limited increase in the stack power is due to the fact that the statics influence on the data is low (Figure 5.10). Increasing and decreasing the rank values by 50% result in 5% and 6% decrease in the stack power, respectively.

The above tests demonstrate that the proposed method is sensitive to the rank selection. The stack power changes between 2% and 6% when varying the rank values by $\pm 50\%$ for the synthetic and field data. Even though the change in the rank values is high, the stack sections exhibit low stack power change. Therefore, the sensitivity of the proposed method to the rank selection is low. The tests also show that the strategy of decreasing the rank as iterations progress is effective.

5.5.3. LIMITATIONS

In terms of limitations, the proposed method does not result in perfect statics correction, i.e. the statics and frequency slices residuals of the synthetic data displayed in the RESULTS section are not zero. However, they are close to zero, which makes their influence negligible. Since short-wavelength statics estimation is an approximation for the rapid variations in properties of the weather layers, it cannot provide perfect approximation. Nevertheless, short-wavelength statics correction is still effective for improving the accu-

racy and resolution of the subsurface models as we demonstrate. Therefore, the proposed method can be used to overcome the limitations of model building techniques that cannot capture the highly varying nature of the near-surface weathering layers.

5.6. CONCLUSIONS

In the presence of rapid variations of topography and properties of the near-surface weathering layers, estimation of accurate near-surface models can be challenging. Short-wavelength statics correction can be used to improve the quality and resolution of subsurface models. While existing techniques can be successful, they depend on an accurate velocity model of the subsurface. At the same time, the statics affect velocity estimation, and vice-versa, which consequently requires iterative velocity and statics estimation. Moreover, to reach to the velocity and short-wavelength statics estimation stage, many pre-processing steps are usually required that include multiples elimination or distinction and denoising.

We propose a 3D model-independent rank-based approach to overcome the above limitations. It's based on the x - y matricization of 5D data that reveals the data's low-rank structure, which gets destroyed due to the incoherency introduced by the near-surface weathering layers. We demonstrate that by means of low-rank approximation and cross-correlation, we can estimate and correct for short-wavelength statics of 5D synthetic and field data without the need of a subsurface velocity model. This way, the estimated statics are no longer influenced by the accuracy of the velocity model, which avoids the iterative velocity and statics estimation. Additionally, we can estimate accurate surface- and non-surface-consistent statics with one process. Moreover, the proposed method requires minimal pre-processing, which can considerably reduce the efforts and time-consuming processing workflow usually needed prior to velocity and short-wavelength statics estimation. The aforementioned benefits, in addition to its high computational efficiency and low sensitivity towards the rank selection, make our proposed method favorable compared to existing methods. Potential applications of the proposed method include short-wavelength statics estimation and correction prior to full waveform inversion of 5D land data.

5.7. ACKNOWLEDGMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his Ph.D. studies. The 3D Stratton field data were collected and made available for worldwide education and training by the Bureau of Economic Geology at the University of Texas at Austin.

REFERENCES

- [1] Ö. Yilmaz, K. Gao, M. Delic, J. Xia, L. Huang, H. Jodeiri, and A. Pugin, *A reality check on full-wave inversion applied to land seismic data for near-surface modeling*, The Leading Edge **41** (2022), pp. 40–46, <https://doi.org/10.1190/tle41010040.1>.
- [2] J. Sheng, C. Reta-Tang, F. Liu, A. Vázquez Cantú, and A. Cabrales Vargas, *Full-waveform inversion and FWI imaging for land data*, in *Second International Meeting for Applied Geoscience & Energy* (2022) pp. 827–831, [10.1190/image2022-3749887.1](https://doi.org/10.1190/image2022-3749887.1).

- [3] M. T. Taner, F. Koehler, and K. A. Alhilali, *Estimation and correction of near-surface time anomalies*, *Geophysics* **39** (1974), pp. 441–463, <https://doi.org/10.1190/1.1440441>.
- [4] J. Ronen and J. F. Claerbout, *Surface-consistent residual statics estimation by stack-power maximization*, *Geophysics* **50** (1985), pp. 2759–2767, <https://doi.org/10.1190/1.1441896>.
- [5] D. Marsden, *Static corrections—a review, part II*, *The Leading Edge* **12** (1993), pp. 115–120, <https://doi.org/10.1190/1.1436936>.
- [6] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <https://library.seg.org/doi/pdf/10.1190/1.9781560801580>.
- [7] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999) <https://doi.org/10.1190/1.9781560801818>.
- [8] R. E. Sheriff, *Encyclopedic Dictionary of Applied Geophysics, Fourth edition* (Society of Exploration Geophysicists, 2002) <https://doi.org/10.1190/1.9781560802969>.
- [9] J. Xu, H. Zhang, and J. Zhang, *A robust surface-consistent residual phase correction method based on migrated gathers*, *Exploration Geophysics* **49** (2018), pp. 336–344, <https://doi.org/10.1071/EG17017>.
- [10] C. P. Ursenbach and J. C. Bancroft, *Playing with fire: Noise alignment in trim and residual statics*, in *SEG Technical Program Expanded Abstracts 2001* (2001) pp. 1973–1976, <https://library.seg.org/doi/pdf/10.1190/1.1816525>.
- [11] D. C. Henley, *Interferometric application of static corrections*, *Geophysics* **77** (2012), pp. Q1–Q13, <https://doi.org/10.1190/geo2011-0082.1>.
- [12] A. M. Alfaraj, E. Verschuur, and F. J. Herrmann, *Residual statics correction without NMO — A rank-based approach*, in *First International Meeting for Applied Geoscience & Energy Expanded Abstracts* (2021) pp. 2565–2569, <https://doi.org/10.1190/segam2021-3583455.1>.
- [13] A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, *Low-rank-based residual statics estimation and correction*, *Geophysics* **88** (2023), pp. V215–V231, <https://doi.org/10.1190/geo2022-0415.1>.
- [14] S. Trickett, L. Burroughs, A. Milton, L. Walton, and R. Dack, *Rank-reduction-based trace interpolation*, in *SEG Technical Program Expanded Abstracts 2010* (Society of Exploration Geophysicists, 2010) pp. 3829–3833, <https://doi.org/10.1190/1.3513645>.
- [15] V. Oropeza and M. Sacchi, *Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis*, *Geophysics* **76** (2011), pp. V25–V32, <https://doi.org/10.1190/1.3552706>.
- [16] N. Kreimer and M. D. Sacchi, *A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation*, *Geophysics* **77** (2012), pp. V113–V122, <https://doi.org/10.1190/geo2011-0399.1>.

- [17] A. Aravkin, R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, *Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation*, *SIAM Journal on Scientific Computing* **36** (2014), pp. S237–S266, <https://doi.org/10.1137/130919210>.
- [18] R. Kumar, C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, *Efficient matrix completion for seismic data reconstruction*, *Geophysics* **80** (2015), pp. V97–V114, <https://doi.org/10.1190/geo2014-0369.1>.
- [19] M. Maraschini, R. Dyer, K. Stevens, and D. Bird, *Source separation by iterative rank reduction - theory and applications*, in *74th EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2012) pp. 1–5, <https://doi.org/10.3997/2214-4609.20148370>.
- [20] H. Wason, R. Kumar, F. J. Herrmann, and A. Y. Aravkin, *Source separation via svd-free rank minimization in the hierarchical semi-separable representation*, in *SEG Technical Program Expanded Abstracts 2014* (Society of Exploration Geophysicists, 2014) pp. 120–126.
- [21] J. Cheng and M. D. Sacchi, *Separation and reconstruction of simultaneous source data via iterative rank reduction*, *Geophysics* **80** (2015), pp. V57–V66, <https://doi.org/10.1190/geo2014-0385.1>.
- [22] M. Bekara and M. Van der Baan, *Local singular value decomposition for signal enhancement of seismic data*, *Geophysics* **72** (2007), pp. V59–V65, <https://doi.org/10.1190/1.2435967>.
- [23] S. Trickett and L. Burroughs, *Prestack rank-reduction-based noise suppression*, *CSEG Recorder* **34** (2009), pp. 24–31.
- [24] N. Moldoveanu, *A singular value decomposition algorithm for the attenuation of high energy towed-streamer noise*, *First Break* **29** (2011), pp. 101–108, <https://doi.org/10.3997/1365-2397.29.12.55902>.
- [25] Y. Chen, Y. Zhou, W. Chen, S. Zu, W. Huang, and D. Zhang, *Empirical low-rank approximation for seismic noise attenuation*, *IEEE Transactions on Geoscience and Remote Sensing* **55** (2017), pp. 4696–4711, <http://dx.doi.org/10.1109/TGRS.2017.2698342>.
- [26] M. Zhang, Y. Liu, H. Zhang, and Y. Chen, *Incoherent noise suppression of seismic data based on robust low-rank approximation*, *IEEE Transactions on Geoscience and Remote Sensing* **58** (2020), pp. 8874–8887, <https://doi.org/10.1109/TGRS.2020.2991438>.
- [27] A. M. Alfaraj, R. Kumar, and F. J. Herrmann, *Automatic statics and residual statics correction with low-rank approximation*, in *80th EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2018) pp. 1–5, <https://doi.org/10.3997/2214-4609.201801107>.

- [28] A. Alfaraj, M. Almubarak, and F. Herrmann, *Correcting for short-wavelength statics with low rank approximation*, in *81st EAGE Conference and Exhibition*, 1 (European Association of Geoscientists & Engineers, 2019) pp. 1–5.
- [29] G. H. Golub and C. Reinsch, *Singular value decomposition and least squares solutions*, in *Linear algebra* (Springer, 1971) pp. 134–151.
- [30] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, *Psychometrika* **1** (1936), pp. 211–218.

6

NEAR- AND SUB-SURFACE MODEL ESTIMATION WITH JOINT MIGRATION INVERSION

In chapters 2 and 5, we proposed data-driven rank-reduction-based methods to correct for the near-surface weathering layers effects of 2D and 3D data. The methods can be efficiently applied prior to imaging and inversion of land seismic data. Despite being effective for accurate subsurface model estimation, near-surface correction remains a simple approximation of the actual near-surface. A more optimal goal is to estimate the near- and sub-surface models, rather than only estimating the latter while approximating the former. To do so, we can use joint migration inversion (JMI), which we briefly introduce.

If the starting near-surface model for JMI is not accurate, which is usually the case, the subsurface model will also become inaccurate. With JMI, the long- and short-wavelength near-surface effects are inherently separated with the estimation of low- (background velocity) and high-wavenumber (reflectivity) models, respectively. We utilize this property to focus on the effect of rapid variations of the near-surface, which manifest themselves in the high wavenumber reflectivity model. In the presence of a simple near-surface, the reflectivity update exhibits low-rank structure, which gets destroyed when the near-surface becomes complex.

By using low-rank image updates, we can lower the influence of the complex near-surface at early stages of the inversion. To allow for more accurate near-surface estimation, we use a multi-scale approach. We demonstrate the results of our proposed method on data modelled on a complex near-surface model.

This chapter is a modified version of the proceeding "A. M. Alfaraj, and D. J. Verschuur, Mitigating the Near-Surface Effect in Velocity and Reflectivity Estimation with Multi-Scale Low-Rank Approximated Image Updates, in 82nd EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2021)

6.1. INTRODUCTION

Land seismic data acquired on top of near-surface weathering layers suffer from undesired wave propagation effects, mostly resulting in time-shifts commonly called statics [1]. This is due to the nature of the unconsolidated near-surface, which can be heterogeneous, rapidly changing, season-dependent and composed of low-velocity material, ranging from 250 to 1000 m/s [2]. Compensating for near-surface anomalies is an important step because if the near-surface model contains errors, deeper subsurface images are usually distorted [3]. To overcome the effect of the weathering layers, a velocity model of the near-surface is first estimated to derive the long-wavelength component of the near-surface effects. As might be expected, this solution is not complete because the near-surface model does not account for the base of weathering, rapid changes in elevation and weathering velocity [1]. At a later stage of the processing workflow, the short-wavelength component of near surface effects is estimated and applied to form the final near-surface correction solution.

While statics correction methods are conventionally used in seismic data processing, they can be also utilized in data pre-processing prior to subsurface models estimation in automatic full wavefield methods. For example, [4] state that full waveform inversion (FWI) fails to estimate a velocity model in the presence of residual reflection statics, which requires residual statics correction prior to FWI. Along similar lines, [5] applies small statics to compensate for elevation differences between a flat surface used in the inversion and the elevation of sources and receivers. [6] utilize guided wave inversion to improve the vertical resolution of the near-seabed model and use the estimated acoustic velocity model as a constrained in further runs of multi-parameter FWI. In their application to multi-component land data pre-processing for FWI, [7] extracts the short-wavelength statics component from the field data and applies its inverse to the modelled synthetic data before matching both datasets. [8] also perform residual statics correction prior to FWI. Recently, [9] show that elastic FWI of the near-surface fails to converge to an accurate solution. The authors state that short-wavelength near-surface correction is the most essential step for land seismic data analysis.

However, as has been outlined by [10] as well as [11], statics correction is a convenient way of handling near-surface anomalies, which effectively should be replaced by velocity analysis of the near-surface. Moreover, prior statics correction if not handled properly may destroy valuable information, e.g. the far offset information if the surface consistency assumption is not met. Additionally, conventional methods require long processing workflows and they focus on enhancing primaries. Therefore, multiples may get destroyed, which may create erroneous inversion results as multiples are also modelled with full waveform techniques. [12] overcomes these limitations. Despite being effective for bypassing the weathering layers, a more optimal approach is to incorporate the near-surface model estimation as part of the inversion.

Joint migration inversion (JMI) aims at estimating velocity and reflectivity models by utilizing all orders of internal multiples [13, 14]. Compared with FWI [15], JMI's propagation operators explain travel-time information (scattering-free), while its reflectivity operators explain amplitude information (travel-time free) making the inversion less non-linear [16]. Ideally, we would want the near-surface anomalies to be solely handled by the inversion engine. However, if the starting near-surface model is not accurate, i.e. not

detailed, which is usually the case for a complex near-surface, the subsurface model will also be inaccurate. Conventionally, short-wavelength near-surface correction is used to remove details that cannot be explained with the inversion. However, that is not optimal due to the aforementioned reasons. In this work, we propose to incorporate a low-rank-based process that takes a similar role to short-wavelength near-surface correction, but also allows for estimating accurate near- and sub-surface models.

6.2. METHODOLOGY

Joint migration inversion is a data fitting procedure that estimates velocity and reflectivity models by iteratively solving the following objective function [13, 14]:

$$\min_{\mathbf{m}} \sum_{i=1}^{N_s} \sum_{\omega} \frac{1}{2} \|\mathbf{p}_i^{-\text{mod}}(\mathbf{m}) - \mathbf{p}_i^{-\text{obs}}\|_2^2, \quad (6.1)$$

where $\|\cdot\|_2$ is the l_2 -norm, N_s is the number of sources, $^-$ denotes up-going wavefields and \mathbf{p}^{mod} and \mathbf{p}^{obs} are, respectively, the modelled and observed wavefields in the frequency domain (ω) at the surface. The model parameters \mathbf{m} with respect to which the objective function is minimized are the velocity (\mathbf{v}) and reflectivity (\mathbf{r}) models. A flowchart summarizing the steps followed to estimate these models is shown in Figure 1, which divides the JMI process into two loops: (i) the upper one for updating the velocity model and (ii) the lower one for updating the reflectivity model. To reduce the non-linearity, these two model parameters are updated differently in a way that the velocity update is based on wavefields traveling in the same direction and is scattering-free. On the other hand, the reflectivity update is based on wavefields traveling in opposite directions and is travel-time free. At the same time, they influence each other since after each model update, newly forward modelled wavefields are generated. For further details, the reader is referred to [13, 14, 16].

In the presence of near-surface anomalies, it is important to update the near-surface model first prior to updating the deeper parts of the model in order to avoid obtaining erroneous subsurface models [3]. As we expect a low-wavenumber (smooth) velocity model from JMI, we should not worry about the long-wavelength effect of the near-surface. It is the short-wavelength component that we need to compensate for. Rather than relying on pre-processing to overcome the short-wavelength near-surface effect, we propose to include that as part of the JMI process.

[12] use a model-independent low-rank-based near-surface estimation and correction in the midpoint-offset domain. Their method requires a transform domain that distinguishes between the singular values decay of data with and without the near-surface effects. In the JMI case, the short-wavelength features are expected to be captured in the high-wavenumber reflectivity update. To further investigate that, we model synthetic data with full wavefield modelling [17] using a complex rapidly changing near-surface model (Figure 6.2(a)). The phase and amplitude of the data are distorted (Figure 6.2(c)). The gather mainly contains short wavelength-like effects in contrast to data modelled with a simple near-surface (Figures 6.2(b) and 6.2(d)). By comparing the first image updates of these two datasets (Figures 6.3(a) and 6.3(b)), we observe that the singular values decay of the latter is much slower than the former (Figure 6.3(d)). Therefore, the image-

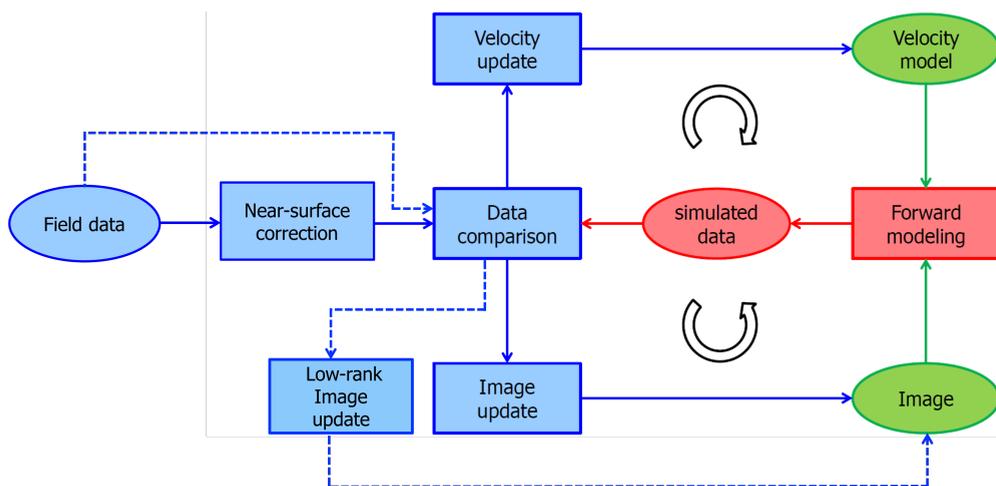


Figure 6.1: Conventional JMI flowchart (solid lines) and the proposed modifications (dashed lines).

6

domain forms a potential transform domain that we can utilize as shown in dashed lines of Figure 1 to obtain image-updates without the short-wavelength near-surface effects (Figure 6.3(c)). However, this low-rank approximated gradient removes details from the near-surface model, which needs to be incorporated at a later stage. To do so, we use a multi-rank scale strategy.

6.2.1. RANK SELECTION

Selecting the rank used in obtaining the image updates may not be a trivial task. The JMI algorithm runs the inversion using frequency bands from low to high frequencies. We utilize this property and propose to select the rank according to the frequency content. We start with a low-rank at low frequencies and increase it with increasing the frequency content. This is a natural way of selecting the rank since higher frequency content means more details and therefore more variability in the model updates. That requires higher rank to include important model updates or low rank to remove these updates. By using a multi-scale approach, where we use a low-rank at initial iterations and increase it at later iterations, we increase the effect of the near-surface gradually. Consequently, we can reduce the effect of the near-surface at the start of the inversion, and include more details as iterations progress. Following this criteria, we are able to estimate accurate near- and sub-surface velocity and reflectivity models.

6.3. RESULTS AND DISCUSSION

In this section, we apply JMI and our proposed algorithm to the synthetic data shown in Figure 6.2(c). Even though the subsurface model is simple (Figures 6.4(c) and 6.4(d)), the estimated reflectivity and velocity models via JMI contain erroneous anomalies at around 350 m depth and between -100 and 100 m lateral position (Figures 6.4(e) and 6.4(f)). This

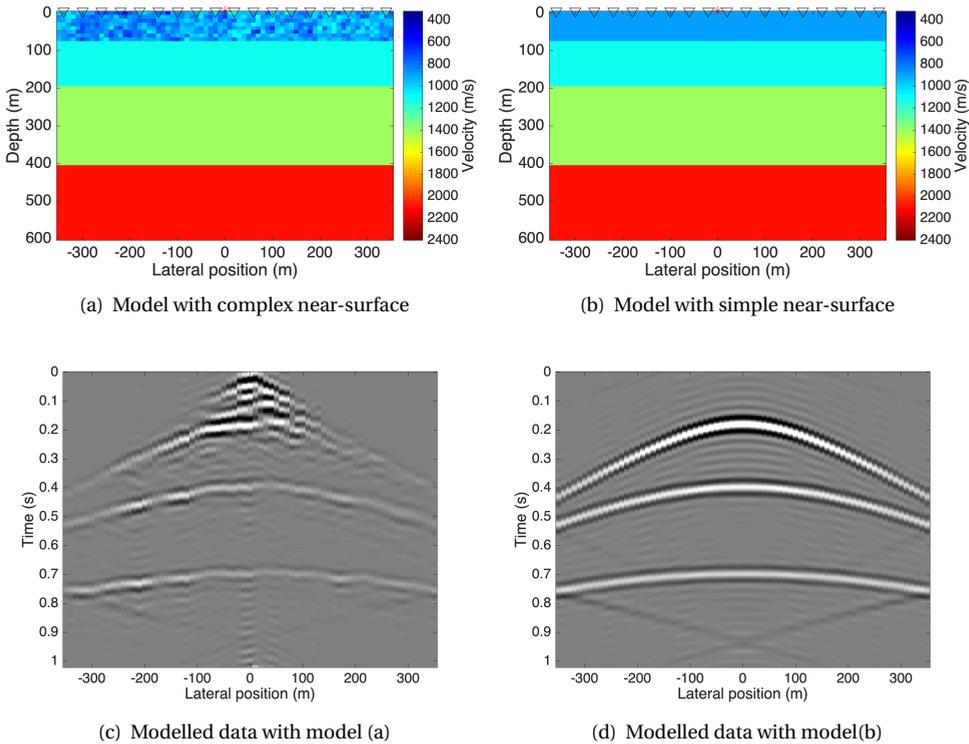


Figure 6.2: Velocity models with (a) complex near-surface and (b) simple near-surface. Synthetic data modelled with (c) complex and (d) simple near-surface.

error is attributed to the fact that the near-surface model in the first iterations is not accurate enough. As a result, the algorithm tries to compensate for that with inserting an anomaly in the subsurface. Once we apply the algorithm with our proposed modifications (Figure 1), we obtain more accurate models as shown in Figures 6.4(g) and 6.4(h).

6.4. CONCLUSIONS

The estimation of subsurface models using land seismic data is complicated by the presence of near-surface weathering layers. To obtain undistorted subsurface models, the near-surface model needs to be accurate to begin with. Despite being effective, accounting for rapid variations in the near-surface properties with near-surface correction is a simplification of the complex near-surface. We propose an alternative solution, which is to allow the velocity estimation engine part of JMI to find a smooth near-surface velocity model while accounting for the high-frequency short-wavelength features by using low-rank approximated image updates in a multi-scale fashion. In this way, we are able to estimate accurate near- and sub-surface velocity and reflectivity models as demonstrated by the synthetic data example.

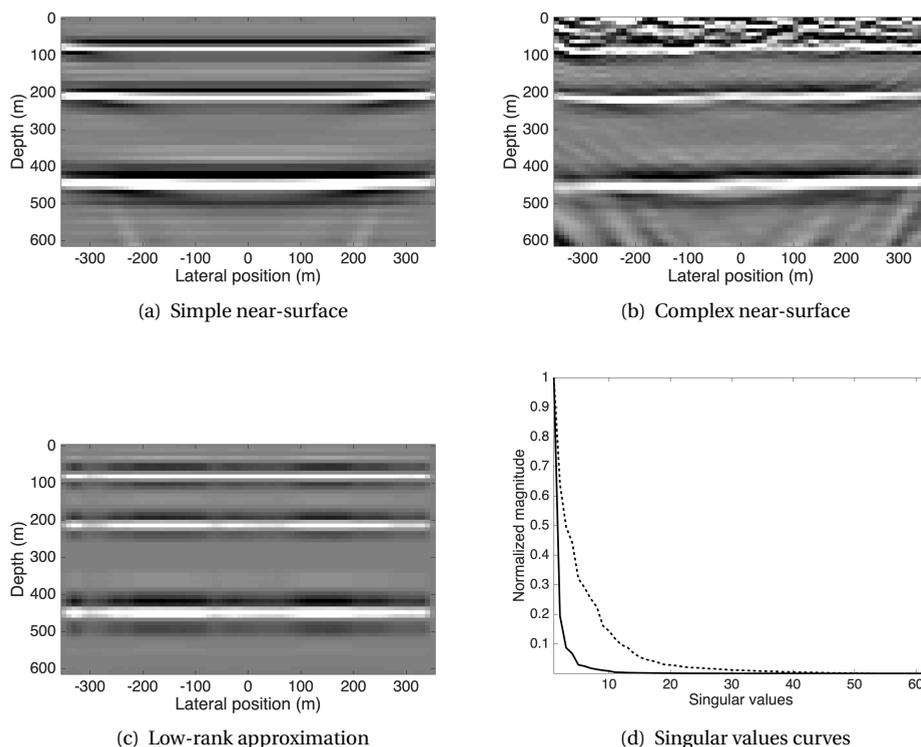


Figure 6.3: First image updates of data modelled: (a) with a simple near-surface layer and (b) with complex near-surface. (c) Low-rank approximated image update. (d) The singular values decay curves of (a) and (b) in solid and dashed, respectively.

6.5. ACKNOWLEDGEMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his PhD studies.

REFERENCES

- [1] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <http://dx.doi.org/10.1190/1.9781560801580>.
- [2] R. E. Sheriff and L. P. Geldart, *Seismic velocity*, in *Exploration Seismology* (Cambridge University Press, 1995) pp. 107–144, 2nd ed.
- [3] I. F. Jones, *Tutorial: Incorporating near-surface velocity anomalies in pre-stack depth migration models*, *First Break* **30** (2012), pp. 47–58, <https://doi.org/10.3997/1365-2397.2011041>.
- [4] A. Vassiliou, R. Meek, R. Stromberg, and L. Woog, *3D acoustic full waveform inversion of 3-D seismic data and PSDM: A near-surface Midland Basin*

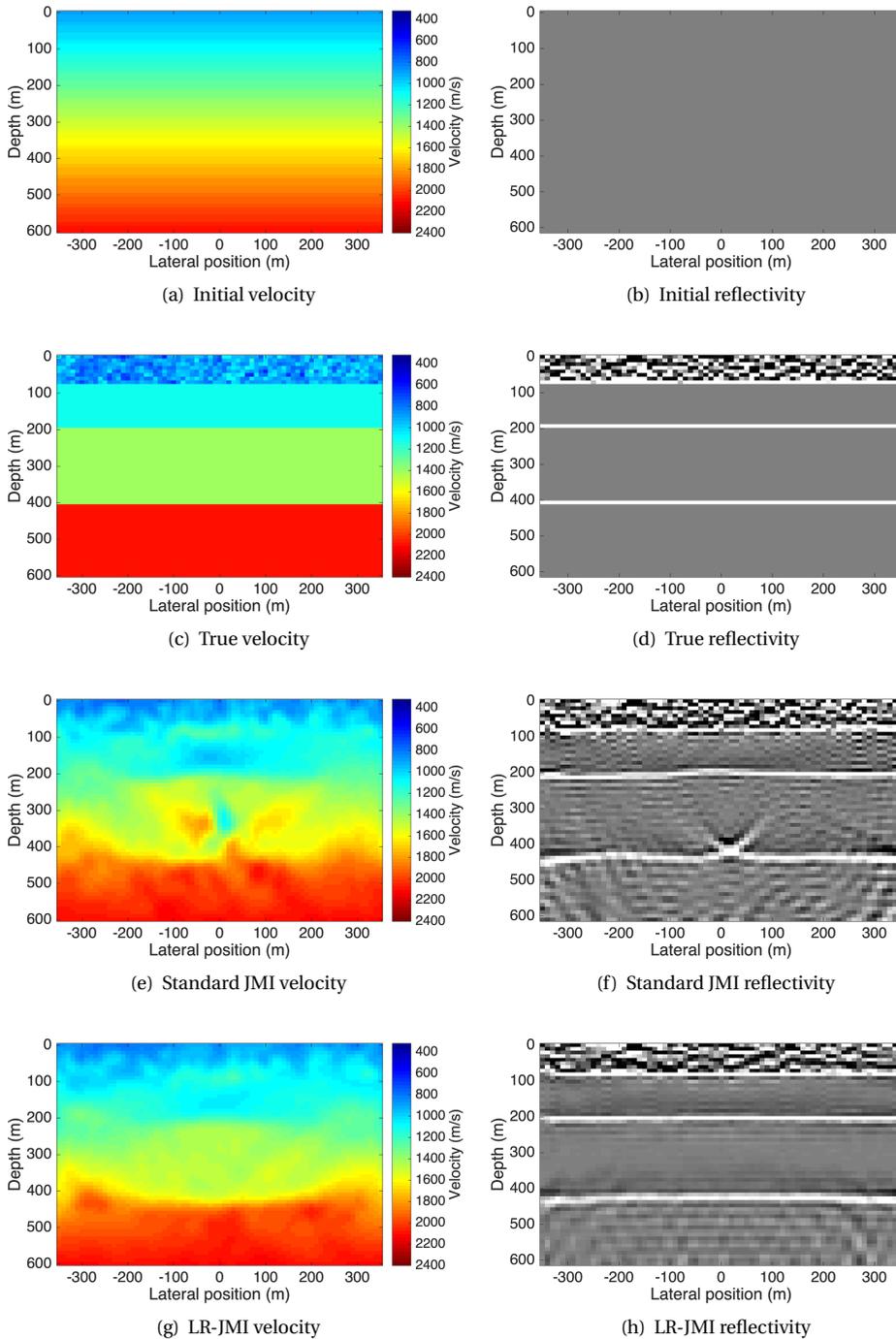


Figure 6.4: The initial (a) velocity and (b) reflectivity models. The true (c) velocity and (d) reflectivity models. The standard JMI estimated (e) velocity and (f) reflectivity models. The estimated (g) velocity and (h) reflectivity with our proposed method.

- study*, in *SEG Technical Program Expanded Abstracts 2015* (2015) pp. 1242–1245, <http://dx.doi.org/10.1190/segam2015-5914651.1> .
- [5] A. Sedova, G. Royle, O. Hermant, M. Retailleau, and G. Lambaré, *High-resolution land full waveform inversion—a case study on a data set from the sultanate of Oman*, in *79th EAGE Conference and Exhibition 2017* (2017) <https://doi.org/10.3997/2214-4609.201701163> .
- [6] L. Zhang, N. Zimmermann, A. Bullock, T. Khlusovich, Ø. Runde, and T. Skørve, *Anisotropic earth model building using multi-survey full waveform inversion with ocean-bottom cable and streamer data*, in *79th EAGE Conference and Exhibition 2017* (2017) <https://doi.org/10.3997/2214-4609.201701234> .
- [7] R. Cova, B. K. Law, and K. Innanen, *Multicomponent land data pre-processing for FWI: a benchmark dataset*, in *CSEG Geoconvention Expanded Abstracts 2018* (Canadian Society of Exploration Geophysicists, 2018).
- [8] J. Sheng, C. Reta-Tang, F. Liu, A. Vázquez Cantú, and A. Cabrales Vargas, *Full-waveform inversion and FWI imaging for land data*, in *Second International Meeting for Applied Geoscience & Energy* (2022) pp. 827–831, [10.1190/image2022-3749887.1](https://doi.org/10.1190/image2022-3749887.1) .
- [9] Ö. Yilmaz, K. Gao, M. Delic, J. Xia, L. Huang, H. Jodeiri, and A. Pugin, *A reality check on full-wave inversion applied to land seismic data for near-surface modeling*, *The Leading Edge* **41** (2022), pp. 40–46, <https://doi.org/10.1190/tle41010040.1> .
- [10] J. Ronen and J. F. Claerbout, *Surface-consistent residual statics estimation by stack-power maximization*, *GEOPHYSICS* **50** (1985), pp. 2759–2767, <https://doi.org/10.1190/1.1441896> .
- [11] M. Cox, *Static corrections for seismic reflection surveys* (Society of Exploration Geophysicists, 1999) <https://doi.org/10.1190/1.9781560801818> .
- [12] A. M. Alfaraj, D. Verschuur, and F. J. Herrmann, *Low-rank-based residual statics estimation and correction*, *Geophysics* **88** (2023), pp. V215–V231, <https://doi.org/10.1190/geo2022-0415.1> .
- [13] X. Staal and D. Verschuur, *Joint migration inversion, imaging including all multiples with automatic velocity update*, in *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013* (2013) <https://doi.org/10.3997/2214-4609.20130375> .
- [14] A. Berkhout, *An outlook on the future of seismic imaging, Part III: Joint migration inversion*, *Geophysical Prospecting* **62** (2014), pp. 950–971, <https://doi.org/10.1111/1365-2478.12158> .
- [15] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, *Geophysics* **74** (2009), pp. WCC1–WCC26, <https://doi.org/10.1190/1.3238367> .

- [16] D. Verschuur, X. Staal, and A. Berkhout, *Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering*, *The Leading Edge* **35** (2016), pp. 1037–1046, <https://doi.org/10.1190/tle35121037.1> .
- [17] A. G. Berkhout, *Review Paper: An outlook on the future of seismic imaging, Part I: forward and reverse modelling*, *Geophysical Prospecting* **62**, pp. 911–930, <https://doi.org/10.1111/1365-2478.12161> .

7

CONCLUSIONS AND RECOMMENDATIONS

Throughout the dissertation, we demonstrate that low-rank-based methods successfully overcome the complexity of the weathering layers near the surface. In chapters 2 and 3, we propose near-surface correction solutions that can be used prior to imaging and inversion of periodically- and densely-sampled land data. For randomly subsampled data influenced by the weathering layers, we propose a joint reconstruction technique in chapter 4. In chapter 5, we extend the near-surface correction to 5D data. These methods should replace existing techniques as they require minimal pre-processing and no knowledge of the subsurface velocity model. Since the ultimate goal is to directly estimate a near- and sub-surface velocity model at once, we propose to use multi-rank-scale image updates to achieve that with joint migration inversion in chapter 6. In this chapter, we draw general conclusions and provide recommendations for future research.

The complex weathering layers near the surface pose challenges for accurate estimation of the subsurface models. In the INTRODUCTION, we summarize these challenges, which we overcome by achieving the dissertation's objectives. Next, we present the conclusions.

7.1. CONCLUSIONS

The following subsections summarize the dissertation's conclusions:

7.1.1. BYPASSING THE WEATHERING

Currently, the near-surface model can be estimated with different data segments. Using direct arrivals, diving and surface waves, smooth models with reasonable resolution can be estimated. However, their resolution is not high enough to capture the rapid variations in the weathering layers near the surface. Therefore, data-driven short-wavelength near-surface correction is essential to overcome the limitations of acquisition and model building methods. It is currently an irreplaceable step prior to velocity model building of land data [1], e.g. with full waveform inversion (FWI) [2, 3]. However, conventional short-wavelength statics estimation methods are not completely data-driven as they need a velocity model to obtain NMO-corrected or migrated gathers [4, 5]. Since the weathering layers influence subsurface velocity estimation, and vice-versa, where the velocity model, iterative velocity and statics estimation are usually necessary. Moreover, attenuation of multiple reflections, or at least selection of a window that only contains primaries — as they are the aligned reflectors after NMO correction or migration —, which may not always be feasible, is required. Due to relying on the surface-consistency assumption, which is not always strictly satisfied, conventional residual statics estimation methods are usually complemented with an additional non-surface-consistent correction step. The aforementioned requirements make short-wavelength statics estimation and correction efforts- and time-consuming, especially for 3D data.

To avoid these challenging requirements, we propose model-independent low-rank-based near-surface estimation and correction. For 2D data (Chapter 2), we show that the midpoint-offset-frequency domain creates favorable processing conditions. In that domain, seismic data exhibit low-rank structures, which get destroyed by the weathering layers as they render coherent energy incoherent. For 3D data (Chapter 5), the x - y -frequency domain meets these conditions. To promote the low-rank structure, we use low-rank approximation. However, thresholding the singular vectors and associated singular values result in loss of data, i.e. the amplitude versus offset is not preserved. To avoid that, we cross-correlate low-rank approximated data with the original data to estimate time-shifts that remove the weathering layers effect. Since the proposed methods are model-independent, we bypass the need of many processing steps, e.g. velocity estimation and multiple attenuation. We also avoid the need of separate surface- and non-surface-consistent statics as they are estimated at the same time. Given the advantages of the low-rank-based near-surface estimation and correction, which we demonstrate on synthetic and field data (Chapters 2, 3 and 5) they become favorable compared to existing techniques. While the proposed methods can handle densely and periodically sampled data, reconstruction of subsampled data affected by the weathering layers remains chal-

lenging.

7.1.2. DATA RECONSTRUCTION

When seismic data is subsampled, e.g. with a compressive sensing acquisition design, in the presence of a complex near-surface, the combined weathering and subsampling effects become formidable. Increasing the coherency of the data is essential for its reconstruction. The conventional approach of near-surface correction followed by interpolation depends on knowledge of the subsurface model [6]. That requires velocity estimation and multiple attenuation, which are difficult to perform from densely sampled data influenced by the weathering layers. Adding data subsampling makes these tasks harder.

In Chapter 4), we propose a joint reconstruction method that simultaneously corrects for the near-surface and interpolates the data. That way, we render incoherent energy coherent to obtain accurately reconstructed densely sampled statics-free data, which we demonstrate on synthetic and field data. The proposed reconstruction shares the same advantages of the low-rank-based model-independent near-surface correction mentioned in the previous subsection. Since low-rank approximation is used for near-surface correction, we also use it as a good initial solution to the rank-minimization optimization problem. Consequently, we decrease the number of iterations required to reconstruct the data. Since both near-surface correction and interpolation are performed in the midpoint-offset domain, we avoid the transformations back and forth to the source-receiver domain. Therefore, the joint reconstruction additionally increases the computational efficiency.

7.1.3. NEAR- AND SUB-SURFACE MODEL ESTIMATION

The previous solutions are geared towards data preparation for subsurface model estimation. They can be essential prior to model building, e.g. with FWI. However, rather than bypassing the weathering layers and removing their effects, it would be optimal to estimate them. Conventionally, we separate the near-surface from the sub-surface model because of the challenge to estimate the former. However, they should be treated equally during model inversion. That requires accurate knowledge of the near-surface as it influences estimation of the subsurface, which is the reason behind the separation in the first place. If we examine the near-surface model, we notice that it contains smooth features and rapidly changing features. That's why near-surface correction is divided into long- and short-wavelength. Joint migration inversion (JMI) offers to estimate a low-wavenumber velocity model and a high-wavenumber reflectivity model. That is an inherent separation of the long- and short-wavelength effects. Therefore, we can estimate both the near- and sub-surface models simultaneously. However, if the starting model is inaccurate, which can be the case for a complex near-surface, the subsurface model will also be inaccurate.

We utilize the fact that rapid variations in the near-surface model lead to high wavenumber features, which end up affecting the reflectivity model of JMI. Since it is hard to estimate an accurate near-surface model at the start of the inversion, we need to mitigate the near-surface effect on the subsurface model at that stage. To do so, we propose in Chapter 6 to use low-rank-approximated image updates. In the reflectivity domain, the near-surface leads to incoherency and subsequently, slow singular values decay, which otherwise would be rapidly decaying. That allows for accurate low-rank

approximation. Using a multi-scale approach, we allow more details of the near-surface to enter the solution at later stages of the inversion, which is effective in estimating the near- and sub-surface models simultaneously.

7.2. RECOMMENDATIONS

The proposed rank-based methods show their potential in achieving the objectives stated in the INTRODUCTION. Using a low-rank-based approach, we perform model-independent near-surface correction and reconstruction of subsampled data influenced by the weathering layers. Since these methods are model-independent and require minimal data preparation, they can be used as quick and efficient processing steps prior to FWI or JMI to allow for estimation of accurate subsurface models. Note that conventional methods are geared towards primaries, which can destroy multiples. Since FWI and JMI model the multiples, the subsurface model may not be accurate. On the other hand, the proposed methods enhance primaries and multiples. To enable estimation of the near- and sub-surface models from data affected by a complex near-surface, we use low-rank image updates during JMI. The proposed methods in general use a multi-rank-scale approach, which makes the effect of rank selection on the results minimal. By exploiting relationships amongst different frequencies, we allow for low-rank approximation of the high frequencies, which is otherwise challenging. In the next subsections, we provide recommendations for future research.

7.2.1. 3D RECONSTRUCTION

In Chapter 4, we propose a joint reconstruction of 2D data in the midpoint-offset domain. It is based on the low-rank-based near-surface correction (Chapter 2) and rank-minimization-based interpolation (Chapter 4). Since we develop the 3D model-independent near-surface correction in the $x - y$ domain (Chapter 5), extension of the proposed joint reconstruction scheme to 3D data becomes feasible.

7.2.2. ADDING BACK THE NEAR-SURFACE EFFECTS

Near-surface correction improves the estimation of subsurface models. After that, it may be feasible to focus on estimation of the near-surface model. That can be achieved with a similar approach to that proposed in Chapter 6, but on the data side. By gradually inserting back the removed near-surface effects to the data, more details to the near-surface model can be added.

7.2.3. NEAR- AND SUB-SURFACE MODEL ESTIMATION

In Chapter 6, we estimate the near- and sub-surface models simultaneously using multi-scale low-rank image updates. We test the method on synthetic data generated by a complex near-surface model. The method needs to be tested on field data. Verification of the results may require in-situ measurements of the near-surface properties.

7.2.4. A TARGET IN THE NEAR-SURFACE

The target of our proposed methods in Chapters 2 to 5 lies beneath the near-surface weathering layers. For shallow near-surface applications such as construction and place-

ment of wind turbines, the shallow subsurface becomes the main target. However, that subsurface can also be affected by a shallower near-surface that exhibits rapid variations of properties. When the model resolution can not capture these variations, they will influence the model beneath. Therefore, the proposed methods can also be implemented on data with high frequency content to estimate high resolution near-surface models.

7.2.5. ACCURATE MODELING

The near-surface is complex. Not only it affects the phase, but also the amplitudes, see Chapter 6. The acoustic approximation of the modelled wavefields can be sufficient for travel-time inversion. However, for accurate phase and amplitude inversion, elastic wavefield modelling is important. That can be achieved with elastic FWI and JMI. However, data conditioning before the inversion can lead to erroneous results. For example, denoising can remove the noise, but also part of the signal. In this case, the amplitudes are no longer reliable for elastic inversion. On the other hand, model-building methods may not be able to handle noisy data. Therefore, extra care must be taken during data preconditioning to prepare the data.

7.2.6. CHANGING THE WORKFLOW

Currently, short-wavelength near-surface correction and interpolation are some of the last steps in the seismic processing workflow. Since the proposed methods in Chapters 2 to 5 are model-independent, they can be implemented early in the processing workflow. That can lead to improvements in the processing steps that come afterwards. For example, denoising estimates coherent data from incoherent data. The higher the initial coherency, the better the denoising performance. Since the proposed methods are robust to realistic noise, up to a certain extent as shown by the field data examples, they may improve the denoising process that comes afterwards. However, high noise levels may influence the performance of the proposed methods, which requires further research. Other processing steps that can benefit from early near-surface correction and interpolation are model-based long-wavelength statics estimation, e.g. travel-time tomography, and methods that require coherent input data such as transform-based methods, e.g. sparse deconvolution and multiple attenuation.

7.2.7. MARINE DATA

Marine data collected in very shallow water may get influenced by mud roll and statics. When receivers are placed at the ocean bottom, they also can be affected by rapid surface elevation variations and changes of the near ocean bottom properties, which can be resolved with the proposed model-independent near-surface correction. In rough water conditions, towed sources and receivers at the surface will not be on a flat surface. If their positions are rapidly changing, that may equate to rapid surface elevation variations on-shore. The latter can be treated with short-wavelength statics correction. However, the former is also influenced by sailing sources and receivers, which are stationary in the case of land data. Since the proposed 2D and 3D methods (Chapters 2, 4 and 5) can estimate non-surface-consistent statics, they may improve the quality of marine data influenced by rough water conditions. However, if those conditions lead to dynamic effects, they can be challenging to overcome with the proposed methods.

REFERENCES

- [1] Ö. Yilmaz, K. Gao, M. Delic, J. Xia, L. Huang, H. Jodeiri, and A. Pugin, *A reality check on full-wave inversion applied to land seismic data for near-surface modeling*, *The Leading Edge* **41** (2022), pp. 40–46, <https://doi.org/10.1190/tle41010040.1> .
- [2] A. Tarantola, *Inversion of seismic reflection data in the acoustic approximation*, *Geophysics* **49** (1984), pp. 1259–1266, <https://doi.org/10.1190/1.1441754> .
- [3] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, *Geophysics* **74** (2009), pp. WCC1–WCC26, <https://doi.org/10.1190/1.3238367> .
- [4] O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, 2001) <https://doi.org/10.1190/1.9781560801580> .
- [5] J. Xu, H. Zhang, and J. Zhang, *A robust surface-consistent residual phase correction method based on migrated gathers*, *Exploration Geophysics* **49** (2018), pp. 336–344, <https://doi.org/10.1071/EG17017> .
- [6] D. Trad, *Five-dimensional interpolation: Recovering from acquisition constraints*, *Geophysics* **74** (2009), pp. V123–V132, <https://doi.org/10.1190/1.3245216> .

A

FULL WAVEFORM INVERSION OF ONE-WAY WAVEFIELDS AT THE OCEAN BOTTOM

Conventional seismic full waveform inversion (FWI) estimates a velocity model by matching the two-way predicted data with the observed data. The non-linearity of the problem and the complexity of the Earth's subsurface, which results in complex wave propagation may lead to unsatisfactory inversion results. Moreover, having inaccurate overburden models may lead to erroneous velocity models. Therefore, estimation of the water layer velocity (overburden for ocean bottom data) is important to creating a subsurface velocity, which we demonstrate for land data influenced by complex weathering layers in chapter 6. Even though the water velocity is usually assumed to be known in the marine scenario, we show that a slight deviation from the actual velocity can lead to cycle-skipping of full waveform inversion, which consequently results in erroneous subsurface models. In order to obtain accurate velocity models from ocean bottom data, we simplify the FWI problem by decomposing the total two-way observed and predicted wavefields into one-way wavefields at the receiver locations using acoustic wavefield decomposition. We then propose to fit the less-complex one-way wavefield, which is the down-going wavefield to obtain a velocity model that contains valuable information about the overburden. We use this inverted velocity model as a better initial model for the inversion using the other more-complex one-way wavefield. We demonstrate our proposed scheme on acoustic non-inverse crime synthetic data produced from the Marmousi2 model. The proposed method provides improved inversion results compared with conventional FWI due to properly accounting for separate model updates from the up- and down-going wavefields.

This appendix is a modified version of the proceeding "A. M. Alfaraj, R. Kumar, and F. J. Herrmann, Seismic waveform inversion using decomposed one-way wavefields, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2018."

A.1. INTRODUCTION

Seismic Full Waveform Inversion (FWI) estimates a high resolution velocity model by minimizing the difference between the two-way observed data and predicted data [1, 2]. Due to the non-linearity of the problem and due to the complexity of seismic data, several authors came up with variations of FWI that include, but not limited to data-correlation based FWI [3], wavefield reconstruction inversion [4], adaptive waveform inversion [5] and tomographic FWI [6] to obtain better subsurface models and mitigate the local minima problem. Apart from changing the FWI formulation, different types of regularizers have been proposed such as Tikhonov [7], total variation (TV) [8], asymmetric TV [9] to circumvent the non-uniqueness problem. [10, 11] suggest simplifying the objective function by using important portions of the data as input for wave-equation inversion. [12] uses damping to extract early arrivals. Following a similar approach, but with a wave-equation based step, we propose to perform wavefield decomposition at the receiver locations to obtain up- and down-going wavefields and then use proper one-way wavefields in the inversion.

The Earth's subsurface is a heterogeneous medium that leads to complex wave propagation. Using approximations in the physics of wave propagation used in FWI that does not take into account certain effects may result in unsatisfactory velocity models. Moreover, using wrong overburden models, i.e water velocity and water-bottom depth for marine acquisition scenario, also may lead to erroneous velocity models. In this work, we propose to simplify the inversion problem by decomposing the complicated two-way wavefields using acoustic wavefield decomposition into one-way wavefields and emphasize these data separately during the inversion to obtain separate model updates.

For acquisition scenarios where the receivers are placed at a distance below the the sources, i.e receivers placed at the ocean bottom while sources are at the surface, the down-going wavefield is a transmitted wavefield that travels between the sources and receivers. This makes it contain valuable information about the medium it travels through that we use to obtain accurate water-bottom depth and correct for errors in the water velocity. We utilize the fact that for ocean bottom data, the one-way down-going wavefield contains less complicated data, namely the direct arrival, surface-related multiples and some of the diving waves, compared with the up-going wavefield. Therefore, we propose to start the inversion using the down-going wavefield for few iterations to obtain a velocity model that we use as an initial model for the inversion using the more-complicated up-going wavefield to obtain the final velocity model. Doing so, we first update the shallow part of the model that contains in our case the water and water-bottom depth, which are important for getting accurate velocity models. Using our proposed method, we are able to properly account for model updates from up- and down-going wavefields separately leading to improved velocity models. For the case when the up-going wavefield does not contain the down-going wavefield, a variation of the method can be used such as minimizing the difference between the up- and down-going wavefields simultaneously.

In the following section, we explain the theory of wavefield decomposition. Next, we modify the FWI formulation to use the decomposed one-way wavefields. Finally, we demonstrate our proposed method on a subset of the Marmousi2 model, [13].

A.2. WAVEFIELD DECOMPOSITION

To perform wavefield decomposition before waveform inversion, we need to decompose the two-way seismic data into its one-way components. Using wavefield decomposition [14], we can decompose the measured two-way wavefields \mathbf{q} at a certain depth level into one-way wavefields \mathbf{d}^\pm using a decomposition matrix \mathbf{N} in the frequency-wavenumber (f-k) domain as follows:

$$\mathbf{d}^\pm = \mathbf{N}\mathbf{q}. \quad (\text{A.1})$$

For the acoustic case, where we perform decomposition in a two-dimensional acoustic media, the two-way wavefields are the acoustic pressure \mathbf{p} and vertical particle velocity \mathbf{v}_z components, while we chose the one-way wavefields to be the up- and down-going acoustic pressure components. In this case, the decomposition matrix \mathbf{N} is dependent on the properties of the acoustic media just above the decomposition depth level. The acoustic decomposition equation in the f-k domain becomes:

$$\begin{pmatrix} p^+ \\ p^- \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\omega\rho}{2k_z} \\ \frac{1}{2} & \frac{-\omega\rho}{2k_z} \end{pmatrix} \begin{pmatrix} p \\ v_z \end{pmatrix}, \quad (\text{A.2})$$

where ω and ρ are the angular frequency and density, respectively. The superscripts $-$ and $+$ indicate up- and down-going wavefields, respectively. The vertical wavenumber k_z is defined as follows:

$$k_z^2 = \frac{\omega^2}{c_p^2} - k_x^2, \quad (\text{A.3})$$

where c_p and k_x are the P-wave velocity and horizontal wavenumber, respectively.

The properties required for acoustic decomposition are roughly known as the receivers are placed in the ocean for an ocean bottom acquisition scenario. [15] shows that the decomposition operators require only a rough estimate of these properties. From the two-way wavefields and using equation A.2, we compute the up- and down-going one-way wavefields that we use in our proposed waveform inversion scheme.

A.3. WAVEFORM INVERSION USING ONE-WAY WAVEFIELDS

In this section, we present our proposed algorithm that utilizes the one-way wavefields in waveform inversion. The method is a two step approach that we use to update the overburden first before updating the deeper subsurface. We utilize the property that acoustic wavefield decomposition at the ocean bottom contains pure down-going wavefield while the up-going wavefield contains both the up-going and down-going wavefields. This is because we perform the decomposition just above the ocean bottom. As a result, we can perform two separate inversions of the down- and up-going wavefields, which allows using the inverted velocity model from the former as in improved initial model for the latter. We modify the conventional two-way full waveform inversion objective function from [1, 2]:

$$\min_{\mathbf{m}} \sum_{i=1}^{n_s} \frac{1}{2} \|\mathbf{d}_i^{\text{pred}}(\mathbf{m}) - \mathbf{d}_i^{\text{obs}}\|_2^2, \quad (\text{A.4})$$

where $\mathbf{m}, \mathbf{d}_i^{\text{pred}}$ and $\mathbf{d}_i^{\text{obs}}$ are the velocity model, the predicted and observed data corresponding to the i th source respectively, to:

$$\min_{\mathbf{m}} \sum_{i=1}^{n_s} \frac{1}{2} \|\mathbf{d}_i^{\pm \text{pred}}(\mathbf{m}) - \mathbf{d}_i^{\pm \text{obs}}\|_2^2, \quad (\text{A.5})$$

where the superscript \pm indicates one-way down or up-going wavefields. We carry out the same optimization procedure used to solve equation A.4, [16] to solve equation A.5. We note that we keep the acoustic parameters required for decomposition fixed so that they do not influence the gradient. In order to obtain the one-way wavefields, we apply wavefield decomposition to the observed and predicted data at the receiver level. We first use the down-going wavefield in the inversion in order to obtain the correct overburden parameters, which are the water velocity and water-bottom depth in our case. We use the obtained velocity model from the down-going wavefield inversion as an input for the next inversion step that fits the more-complicated up-going wavefield which provides the final velocity model. In the next section, we show different inversion results using our proposed method and compare it with conventional FWI.

A.4. NUMERICAL RESULTS AND DISCUSSION

We demonstrate our proposed algorithm on non-inverse crime synthetic data using part of the Marmousi2 model, (figure A.2(a)). We model 2D acoustic synthetic data, (Figure A.1(a)), using time-domain finite difference modelling [17]. The acoustic data is parametrized by the P-wave velocity and density. The sources are placed near the ocean surface while the receivers are placed at the ocean bottom. The source and receiver intervals are 200m and 12.5m, respectively.

We perform P-wave velocity inversion in the time-domain using a constant-density wave-equation, see [16] for more details, starting from a good initial model, (Figure A.2(b)). Figure A.2(c) shows the result of running conventional FWI of the two-way acoustic data using the limited-memory projected quasi-Newton algorithm [18]. We impose lower and upper bound constraints on the velocity model. Using the same initial model, (Figure A.2(b)), we run our proposed algorithm on the two-way data as follows:

Algorithm A.1: Inversion of one-way wavefields

- 1 Compute the two-way predicted data \mathbf{d}^{pred} using the initial model
 - 2 Decompose \mathbf{d}^{pred} and \mathbf{d}^{obs} using equation A.2, (Figures A.1(b), A.1(c))
 - 3 Invert using the down-going wavefield: $\mathbf{m} = \underset{\mathbf{m}}{\text{argmin}} \sum_i^{n_s} \frac{1}{2} \|\mathbf{d}_i^{+\text{pred}}(\mathbf{m}) - \mathbf{d}_i^{+\text{obs}}\|_2^2$
 - 4 Use \mathbf{m} as initial model, invert using the up-going wavefield:

$$\mathbf{m} = \underset{\mathbf{m}}{\text{argmin}} \sum_i^{n_s} \frac{1}{2} \|\mathbf{d}_i^{-\text{pred}}(\mathbf{m}^+) - \mathbf{d}_i^{-\text{obs}}\|_2^2.$$
-

The velocity model obtained from the down-going wavefield inversion has the correct water-bottom depth compared with the initial model as well as minor model updates,

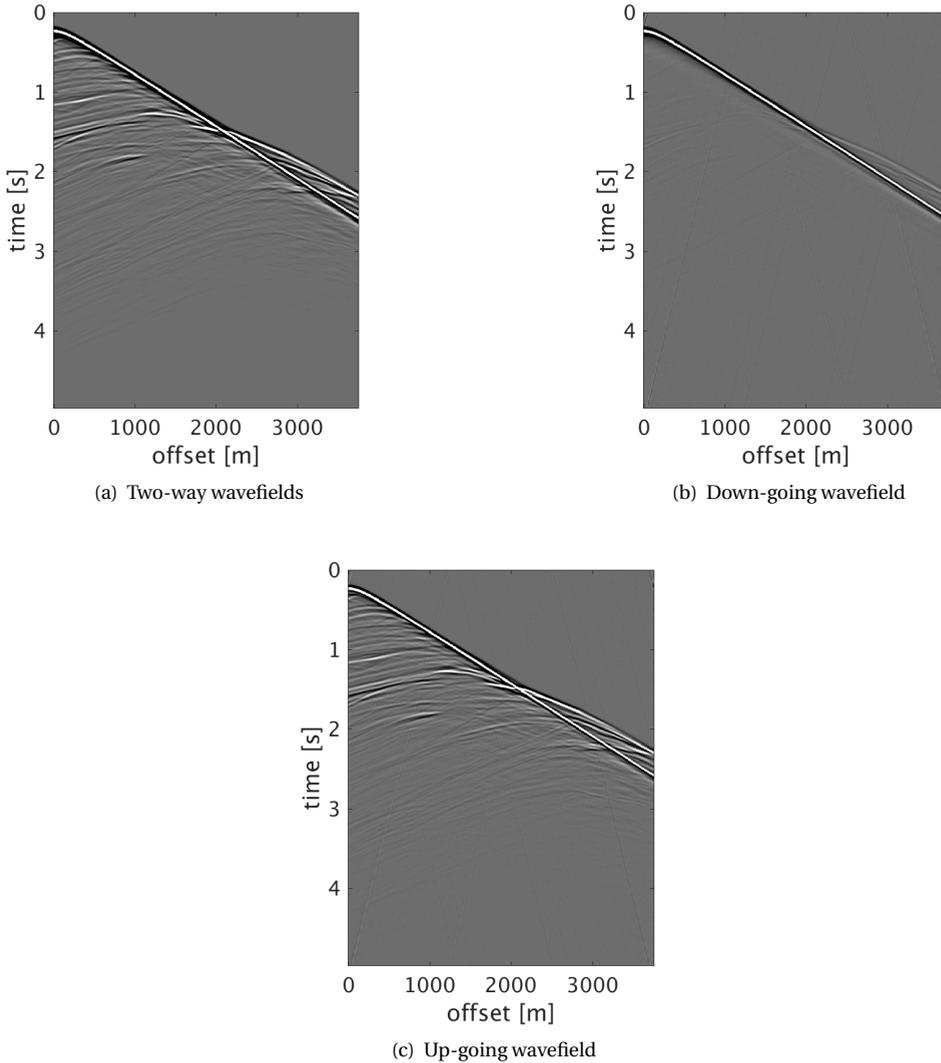


Figure A.1: (a) Two-way hydrophone data, (b) decomposed down- and (c) up-going acoustic pressure at the ocean bottom.

(Figure A.2(d)). Compared with the model obtained from conventional acoustic FWI of the total data shown in Figure A.2(c), the velocity model obtained from our proposed method, (Figure A.3(d)), is better resolved and the different layers and reservoirs can be easily interpreted.

When the starting model is poor, the data becomes cycle-skipped, which leads to unsatisfactory velocity models. To mitigate the cycle skipping problem, we perform multi-

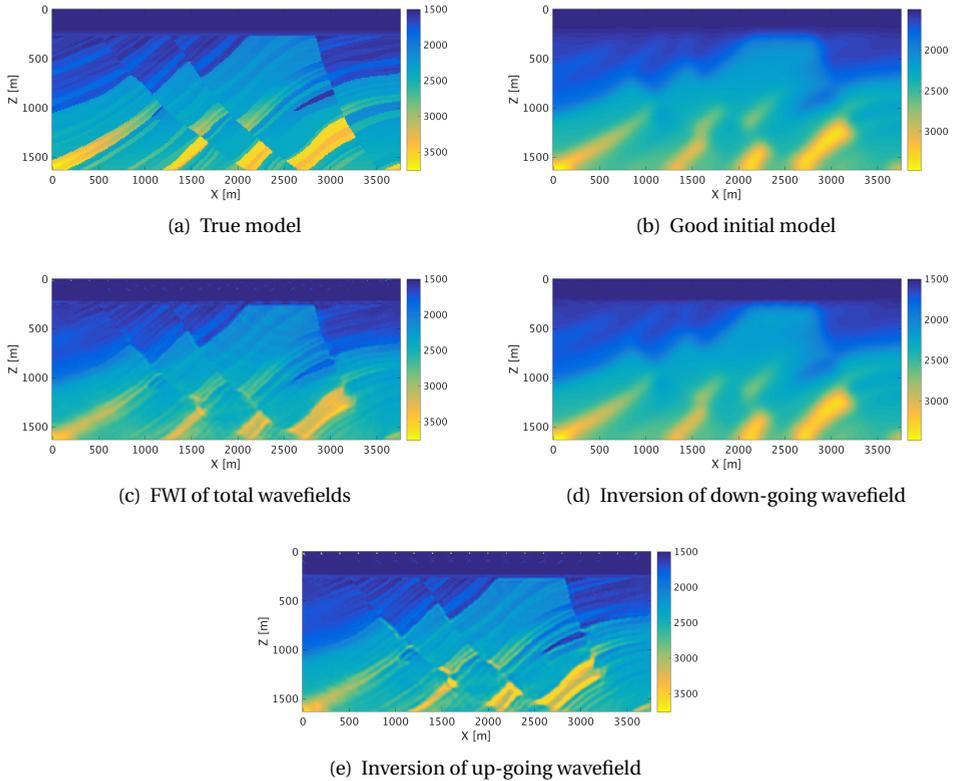


Figure A.2: (a) Part of the true p -wave Marmousi2 model, (b) good initial velocity model, velocity models obtained from (c) FWI of the two-way wavefields, and waveform inversion of the (d) down- and (e) up-going wavefields.

scale inversion, [19], starting from a poor initial model, (Figure A.3(a)). We use frequencies starting from 3.5 Hz to 43 Hz. We increase the frequency bandwidth with 3 Hz after running 10 iterations at each frequency band. Comparing the multi-scale conventional FWI results, (Figure A.3(b)), with the results obtained from our proposed algorithm combined with multi-scale inversion, (Figure A.3(d)), we observe that the latter is a better velocity model with well-defined structures in the shallow and deeper parts of the model.

Another case where our proposed method is beneficial is when the ocean water velocity is unknown or contain errors. The water velocity can change with depth because of variations in temperature and salinity. Assuming a fixed water velocity or running the inversion with a minimum water velocity of 1500 m/s can lead to erroneous inversion results as shown in Figure A.4(a). We should note that in this example we constrain the minimum water velocity to the true minimum water velocity, and use a good initial model for the inversion (Figure A.2(b)). Alternatively, we use the down-going wavefield to invert for the water velocity starting from a constant velocity model. We then incorporate this

model with the good initial velocity model, (Figure A.2(b)), to obtain the model shown in Figure A.4(b). This example clearly shows the importance of having an accurate overburden model, which we obtain by performing a relatively very small number of waveform inversion iterations using the down-going wavefields, in order to be able to update the deeper parts of the model.

We demonstrate the benefits of using our proposed method on several examples to obtain improved, higher resolution and more accurate velocity models compared with conventional FWI. In terms of computational costs, it is equivalent to the computational costs of FWI, but requires less total number of iterations at least for the examples we present. It is in our future plan to use our proposed method to estimate the S-wave velocity model from elastically decomposed data, estimate the source wavelet and investigate the cycle skipping problem.

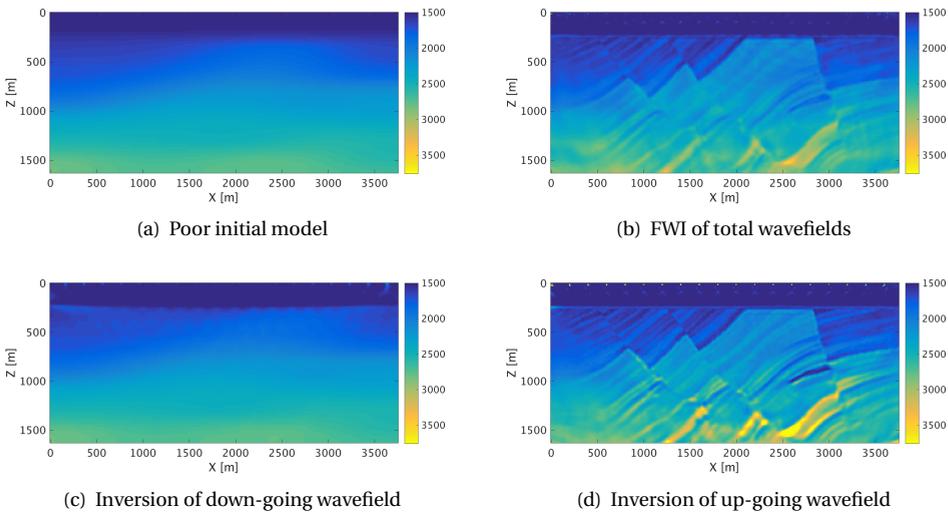


Figure A.3: (a) Poor initial velocity model, velocity models obtained from (b) FWI of the total wavefield, and inversion of the (c) down- and (d) up-going wavefields.

A.5. CONCLUSIONS

We present a method that utilizes the decomposed acoustic one-way wavefields in acoustic waveform inversion. The method simplifies the FWI problem by matching the one-way decomposed wavefields rather than the total complex two-way wavefields. We implement the method by first performing waveform inversion of the less-complicated down-going wavefields followed by waveform inversion of the more-complicated up-going wavefields using the results of the first inversion as a better initial model for the second one. We demonstrate the successful application of our proposed method on several examples using non-inverse crime acoustic synthetic data produced from the Marmousi2 model. In the case when the initial water velocity was inaccurate, FWI fail to provide a satisfactory

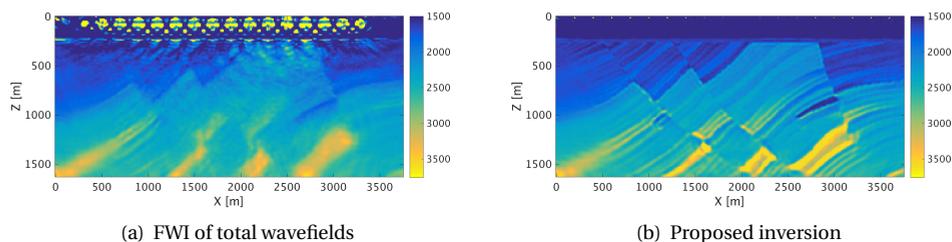


Figure A.4: Illustration of starting the inversion with an erroneous overburden (water velocity) model: (a) FWI of the total wavefield, (b) our proposed inversion.

model while the proposed method does not because it estimates the water velocity using the down-going wavefield. In all the examples, our proposed method provide improved, higher resolution and more accurate inversion results compared with conventional FWI by separately accounting for model updates from the up- and down-going wavefields.

A.6. ACKNOWLEDGEMENTS

Ali M. Alfaraj extends his gratitude to Saudi Aramco for sponsoring his Ph.D. studies.

REFERENCES

- [1] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005) <https://epubs.siam.org/doi/pdf/10.1137/1.9780898717921>.
- [2] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, *GEOPHYSICS* **74** (2009), pp. WCC1–WCC26, <https://doi.org/10.1190/1.3238367>.
- [3] T. Van Leeuwen and W. Mulder, *Velocity analysis based on data correlation*, *Geophysical Prospecting* **56** (2008), pp. 791–803, <https://doi.org/10.1111/j.1365-2478.2008.00704.x>.
- [4] T. Van Leeuwen and F. J. Herrmann, *Mitigating local minima in full-waveform inversion by expanding the search space*, *Geophysical Journal International* **195** (2013), pp. 661–667, <https://doi.org/10.1093/gji/ggt258>.
- [5] M. Warner and L. Guasch, *Adaptive waveform inversion-fwi without cycle skipping-theory*, in *76th EAGE Conference and Exhibition 2014* (2014) <https://doi.org/10.1190/geo2015-0387.1>.
- [6] B. Biondi and A. Almomin, *Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion*, *Geophysics* **79** (2014), pp. WA129–WA140, <https://doi.org/10.1190/geo2013-0340.1>.

- [7] G. H. Golub, P. C. Hansen, and D. P. O’Leary, *Tikhonov regularization and total least squares*, *SIAM Journal on Matrix Analysis and Applications* **21** (1999), pp. 185–194, <https://doi.org/10.1137/S0895479897326432>.
- [8] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, *Physica D: nonlinear phenomena* **60** (1992), pp. 259–268, [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- [9] E. Esser, L. Guasch, T. van Leeuwen, A. Y. Aravkin, and F. J. Herrmann, *Total variation regularization strategies in full-waveform inversion*, *SIAM Journal on Imaging Sciences* **11** (2018), pp. 376–406, <https://doi.org/10.1137/17M111328X>.
- [10] Y. Luo and G. T. Schuster, *Wave equation inversion of skeletalized geophysical data*, *Geophysical Journal International* **105** (1991), pp. 289–294, <http://dx.doi.org/10.1111/j.1365-246X.1991.tb06713.x>.
- [11] K. Lu, J. Li, B. Guo, L. Fu, and G. Schuster, *Tutorial for wave-equation inversion of skeletonized data*, *Interpretation* **5** (2017), pp. SO1–SO10, <https://doi.org/10.1190/INT-2016-0241.1>.
- [12] Y. Choi and T. Alkhalifah, *Frequency-domain waveform inversion using the phase derivative*, *Geophysical Journal International* **195** (2013), pp. 1904–1916, <http://dx.doi.org/10.1093/gji/ggt351>.
- [13] G. S. Martin, R. Wiley, and K. J. Marfurt, *Marmousi2: An elastic upgrade for marmousi*, *The Leading Edge* **25** (2006), pp. 156–166, <https://doi.org/10.1190/1.2172306>.
- [14] C. Wapenaar, P. Herrmann, D. Verschuur, and A. Berkhout, *Decomposition of multi-component seismic data into primary p- and s-wave responses*, *Geophysical Prospecting* **38** (1990), pp. 633–661, <https://doi.org/10.1111/j.1365-2478.1990.tb01867.x>.
- [15] A. M. Alfaraj, D. Verschuur, and K. Wapenaar, *Near-ocean bottom wavefield tomography for elastic wavefield decomposition*, in *SEG Technical Program Expanded Abstracts 2015* (2015) pp. 2108–2112, <http://dx.doi.org/10.1190/segam2015-5918497.1>.
- [16] P. A. Witte, M. Louboutin, K. Lensink, M. Lange, N. Kukreja, F. Luporini, G. Gorman, and F. J. Herrmann, *Full-waveform inversion - Part 3: optimization*, *The Leading Edge* **37** (2018), pp. 142–145, (The Leading Edge), <https://doi.org/10.1190/tle37020142.1>.
- [17] J. W. Thorbecke and D. Draganov, *Finite-difference modeling experiments for seismic interferometry*, *Geophysics* **76** (2011), pp. H1–H18, <https://doi.org/10.1190/geo2010-0039.1>.
- [18] M. Schmidt, E. Berg, M. Friedlander, and K. Murphy, *Optimizing costly functions with simple constraints: A limited-memory projected quasi-newton algorithm*, in *Artificial Intelligence and Statistics* (2009) pp. 456–463.
- [19] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent, *Multiscale seismic waveform inversion*, *Geophysics* **60** (1995), pp. 1457–1473, <https://doi.org/10.1190/1.1443880>.

ACKNOWLEDGEMENTS

THE JOURNEY IS WHAT MATTERS

"The journey is what matters". I couldn't agree more with my beloved wife's perspective. The journey with all that it carried, whether ups or downs, happy or sad moments, successes or failures has been wonderful. There has always been and there will always be moments where we wish that did or didn't happen. In the end, we always see a bright future ahead. How can it not be?

Even though we look towards the future, it is time that we stop and look backwards as "the past determines the future". The people I encountered throughout my life influenced its direction. My deepest appreciation goes to each and everyone who has been part of my journey. This part of my book is dedicated you, as without you, I wouldn't be where I'm today.

THE BEGINNING

Allow me to start with my supervisor and promoter **Dr. Eric Verschuur**. Dear Eric, I have learned a lot from you at a scientific, academic, professional and personal level. It has been an honour to work with you. You are a source of inspiration and motivation to everyone around you. Despite the vast knowledge you carry, you are humble, kind and supportive. I cannot thank you enough for everything that you taught me explicitly or just by being a great role model.

We go back a long way. The first time we met was around 12 years ago during the inauguration of Aramco's research centre at TU Delft. Later that night, we had dinner at the Des Indes hotel together with **Prof. Berkhout**, who I sincerely thank for establishing the **Delphi** Consortium, **Dr. Panos Kelamis**, who I highly appreciate for introducing me to Delphi, and Aramco's management **Mr. Samer Alashgar** and **Dr. Mustafa Alali**, who provided me with immense support. During the event, your characteristics were hardly unnoticeable. I immediately realized that I wanted to do my master's thesis with you. Thank you for giving me that opportunity. My experience was great, so I wanted to continue my PhD under your supervision. However, I was too late to apply as the available PhD seat was already taken. That seemed to be the end of our time together at TU Delft (to be continued...).

BEAUTIFUL BRITISH COLUMBIA

The journey moved to beautiful British Columbia, where I joined my co-promoter **Prof. Felix Herrmann** at the Seismic Laboratory for Imaging and Modeling (**SLIM**) at the University of British Columbia (**UBC**) in Vancouver. Dear Felix, thank you for giving me the chance to join your group, work with you and learn from you. It has truly been a privilege. Your ability to capture emerging technologies from different fields and adapt them for the seismic community is impressive. During my time at UBC, we ventured around differ-

ent topics. The scattering transform, which I never heard of before joining UBC, seemed promising. But in the end, we had to switch to data reconstruction, deblending and FWI of marine multicomponent data. Our work on reconstruction of multicomponent marine data was recognized as one of the top papers presented at the SEG annual meeting. Thanks to your course "Imaging and estimation with wavelets" that sparked my interest in the topic, and your critical feedback, especially on the writing. Even though you mentioned that my writing is good, you still gave me plenty of corrections, which took it to the next level.

While discussing my work with **Dr. Andrey Bakulin**, he raised the question: how can the methods we developed be applied onshore? While it seemed logical to stay away from land data as everyone around me did, it sparked my interest due to its challenging nature. Therefore, we began to work in that direction. However, due to a political conflict, there was an order to move out of Canada, which marked the end of my time at UBC. That also coincided with your move to Georgia Tech. It was not easy, but after some time, I realized that there is a bright future ahead.

At UBC, I had the pleasure of meeting and working with many colleagues that I would like to acknowledge. **Miranda Joyce**, thank you for your assistance with my move to UBC and my subsequent requests. **Henryk Modzelewski**, thank you for your support with organizing the cluster and for keeping up with my swap-intensive jobs. Thank you **Prof. Eldad Haber** for your inverse theory course. **Prof. Michael Bostock** and **Prof. Eldad Haber**, thank you for your input on my research. My colleagues at SLIM and UBC, **Rajiv, Yiming, Thibaut, Haneet, Luz, Zhilong, Rongrong, Shashin, Ben, Oscar, Bas, Devin, Keegan, Felix O, Ali, Curt, Emmanouil, Marie, Verena** and **Gustavo**, thank you for enlightening me on various topics during the seminar talks and for the good times. **Rajiv**, I highly appreciate your support upon joining the group. It was a pleasure to work with you. **Audrey Van Slyck, Prof. Roger Beckie**, and **Prof. Michael Bostock**, thank you for your assistance with my move out of UBC. I also would like to thank the sponsors of the **SINBAD** Consortium for the discussions and feedback during the meetings.

While in Vancouver, we were fortunate to make friends for life. **Marisol, Mario, Emma, and David**, thank you for being part of our family and for allowing us to join yours. **Leen & Yasmeen** grew up in your house as your children and for that, we are forever in debt to you. **Berta**, thank you for becoming part of our family. **Nizar**, it was nice to meet you every now and then in Vancouver. I'm grateful for the **UBC Tennis Club** and my hitting partners **Masato, Pom** and **Nathan**. My brother **Abdullah**, thank you for visiting us far from home; it was lovely. My brother **Hussain**, it was nice having you with us during our last summer in Vancouver. Taking the English course was worth it; only few months are left before you earn your B.Sc. degree, also in Geophysics! Acadia park, our neighbours, and the city of Vancouver, thank you for welcoming us and for making us feel home. Beautiful memories we look back to with fondness.

A NEW BEGINNING

Destiny brought me back to Delft! Before leaving Canada, I contacted Eric, explained the situation and applied for a PhD position. This time, luckily, a slot was available. Thank you **Eric** for giving me another chance to re-join you, and thanks to **Felix** for accepting to be my co-promotor. We started a new journey that we thought would be a short one. Well,

it wasn't short enough, but also not long enough. To make it interesting, we decided to focus on the challenges of land data. We started to work on JMI of land data. I was a fan of rank-based methods, thanks to **Felix** for sparking my interest, so we continued in that direction, which resulted in this dissertation. I'm grateful for your guidance, support and mentorship. With that, we went from initial ideas, all the way to robust methods applied to 2D and 3D field data. Whenever I encountered a problem, you had the solution ready. Your enthusiasm and energy make it possible to overcome any obstacle. Also thank you for being easy to approach and fun to talk to. I fairly enjoyed all the meetings, conversations, dinners and our latest Crown lounge visit.

The journey this time also wasn't so smooth. Around one year after starting at TU Delft, COVID-19 struck out of the blue. The lockdowns and the work from home started. The situation was not easy, but we managed to survive with online meetings. Despite not ideal, you made the meetings almost equivalent to physical meetings. In a short time, you moved the 3D acoustic imaging in air research practicum from physical to online. Then supervising the students was an easy task for me, as I also managed to learn from your online meeting skills. I hope to also learn more from your impressive efficiency and organizational skills. Your schedule is usually full, yet you always find time to finish a lot of tasks and meet with everyone. Perhaps one day you will write a book about it. It'll definitely be a bestseller!

Dear members of the doctoral committee: **Prof. Bernd Rieger, Prof. Mirko van der Baan, Prof. Martin van Gijzen, Dr. Deyan Draganov and Dr. Clément Kostov**, I extend my gratitude to you for accepting the invitation to join my committee. Your willingness to participate is sincerely appreciated, and I consider it a great honour. I'm grateful for your feedback. Thank you **Mirko** and **Clément** for your thorough revision and the critical questions. Your contributions resulted in improving my dissertation greatly.

I'm grateful for the sponsors of the **Delphi** consortium for their interactions, feedback and socializing during the Delphi meetings in Houston and The Hague. The challenging questions they asked were translated into ideas for improving my research. Organizing these meetings is not an easy task. The lion's share goes to **Eric**. Anyone else can be absent, except him. Thank you **Eric** for everything you do to make those meetings smooth and successful. Also thank you for making Delphi shine, which I'm sure will continue to do so under your leadership. Thank you **Gerrie** for the amazing organization of the venue, breakfasts, lunches and dinners. Thank you **Dr. Gerrit Blacquiere** for the organization of the second day of the meeting. Everyone associates that day and the blue book with you. I'm sure you are now enjoying your retirement and sailing trips. **Dr. Deyan Draganov, Dr. Koen van Dongen, Prof. Evert Slob** thank you for your inspirational talks on interferometry, medical imaging, EM and Marchenko. **Eric, Gerrie, Gerrit, Deyan, Koen, Evert**, the side talks with you were always fun.

At TU Delft, I enjoyed taking a number of courses at that added to my knowledge and broadened my horizons. Thank you **Dr. Jan Thorbecke** for the high performance computing course. Thank you **Eric** for the multiples course. Thanks to the **Graduate School** for offering courses related to management, leadership, teamwork and cross-cultural communication. I also appreciate the the weekly seminars at the Department of **Imaging Physics**, first at the acoustical imaging group on medical imaging and then at the computational imaging group on microscopy. Thank you **Eric, Prof. Sjoerd Stallinga, Prof.**

Bernd Rieger, Dr. Jeroen Kalkman, Angela and **Annelies** for organizing the seminars, the discussions and feedback. **Dr. Tristan van Leeuwen**, than you for the organizing the Dutch Inverse Problems Meeting.

I would like to acknowledge the department's support system. **Angela**, thank you for your support with my move to the Netherlands and for dealing with all the subsequent requests, while always showing a smiling face. We used to talk about tennis, but unfortunately, we didn't have the chance to play together. The department days out were lots of fun, thanks to you and **Annelies** for making all the arrangements. **Henry**, thank you for your help with setting up my machine and account. Also thank you for the help with the lab equipment. **Ronald** thank you for your support with the cluster requests. My colleagues at the ImPhys department: **Jos, Peter, Isabel, Qingru, Wenxiu, Loes, Alim, Kote, Huangcheng**, thank you for the side talks and the good times.

Jelle of Proefschriftspecialist, thank you for your assistance with printing my dissertation.

SAUDI ARAMCO

You may have noticed that most of my chapters end with acknowledgements to **Saudi Aramco**. I'm forever in debt to **Saudi Aramco** for sponsoring my PhD studies and counting the duration of my PhD towards my service years. The time I spent abroad influenced the trajectory of my life. It gave me and my family the chance to experience different cultures and systems. It allowed us to grow not only scientifically, but also professionally and personally. Without Saudi Aramco's support, all of that wouldn't have been possible.

Thanks to my department, the **EXPEC Advanced Research Center** for the unwavering support. Thank you **Dr. Ali Meshari** and **Dr. Ahmed Eidan** for your support and encouraging words. Thank you **Dr. Panos kelamis** for introducing me to the Delphi consortium. Thank you **Dr. Mustafa Alali** for your continuous support during my ADP. The previous chief technologists of GPT, **Dr. Abdulaziz Almuheidib, Dr. Maher Almarhoon, Dr. Khalid Rufaii, Dr. Ali Almomin** and **Dr. Nasher BenHasn**, thank you for always offering your assistance. The current chief technologist, my dear friend **Dr. Abdulrahman Alshuhail**, thank you for your endless support. Your advice and suggestions influenced the directions I took. I look forward to working with you on new frontiers. **Dr. Andrey Bakulin**, thank you for keeping me connected to EXPEC ARC with meetings and continuous communication. My current and previous colleagues at **Delft GRC: Ali, Diego, Rolf, Roald, Rob, Yimin, Apostolos, Jewoo, Hannes, Mikhail, Marcin and Paul**, thank you for your hospitality during the centre visits. Meeting you every now and then makes me feel home.

Throughout my PhD, I interacted with **Aramco Overseas** offices in London and The Hague. They were always there, checking on us and offering their assistance. I would like to thank the previous head of professional development in London **Turki Ayed** for his support, especially during the COVID-19 pandemic. I wish you all the best in your new role, I'm sure you will excel. I also would like to sincerely thank my advisors at Aramco Overseas **Aggie** and **Marc** for taking care of my requests in a timely manner. You were always there whenever I needed you. Knowing that I have your support put me at ease throughout my studies. The meetings we had on your trips to the Netherlands were fruitful. Dear **Marc**, your support, especially during the last stretch of the program was instrumental. I thoroughly enjoyed the conversations we had. I believe that you are truly a role model advisor.

I hope to meet you again in the future, to give you one of those KLM Houses.

I am grateful to Saudi Aramco's higher management, in particular **Mr. Faisal Al-Hajji** and **Mr. Sami Al-Murshid** for meeting with us during their trips to the Netherlands and for giving me the opportunity to take part in their activities. Thank you for appointing me the representative of Saudi Aramco employees during focus group meetings with **Mr. Al-Hajji**. It was also an honour for me to be the MC Saudi Aramco **Board of Directors** dinner in the Netherlands. Such high-level events exposed me to a completely different management league and enriched my range of experiences.

I sincerely thank the Ambassador of Saudi Arabia in The Netherlands **H.E.Mr. Ziad Al Atiyah** for fostering a wonderful environment for us in the Netherlands. Thank you for always offering your support, for being kind and humble. I cannot thank you enough for accepting the invitation of my daughters to visit their class. Talking to them at a young age, and giving them the chance to interview you will forever stay in their minds. I also would like to thank you for always inviting us to the embassy's events, where we feel at home.

TO MY FRIENDS

During my time at TU Delft, I was lucky enough to work with colleagues that became close friends. My dear **Delphi 0.0: Dong, Leo, Siamak, Joost, and Anyu**, it was a pleasure to share the office with you. I learned a lot from you throughout all the discussions we had. You made the office a fun place to work at. The lunches and coffee breaks made the day go by quickly. Well, we still had fun outside office hours over dinners and boardgames. Thank you **Leo** for introducing us to the world of boardgames. Despite the competitive spirit during those games, they were so much fun. **Dong**, thank you introducing me to the Chinese cuisine. We cannot have enough of those Chinese, Persian and Middle Eastern dinners. Well, I wish we could go for one right now.

F230 felt like home for me, and you were like my brothers. I remember the day you fasted to show your support during Ramadan. It wasn't easy, but you did it. **Leo**, you surprised me with fasting for the whole week. **Siamak**, thank you for your generosity with always sharing relevant research and events with us. I wouldn't have attended the amazing **Dutch Inverse Problems Meeting** if you hadn't encouraged me to. Thank you **Dong** for bringing Chinese presents and candy from your trip to China. I will forever cherish the memories we created over the years. We will definitely create more over the coming years too. We are all waiting for the invitation to go to China for your wedding **Dong**. I'm not sure, as he always says, if we will also go to **Australia** for someone's else wedding.

Dear **Delphi 1.0: Dieter, Billy, Camille, Andrea, Azin, Sverre, Aydin, and Mohammed**, it was a pleasure to get to know you and learn from you. Despite being in the building next door, the occasional lunches we had kept us connected. Thank you for being caring, supportive and fun to be around. **Andreas (Delphi 0.5?)** and my ex-Delphi colleagues, **Shan, Mikhail, Jan-Willem, Aparajita, Tiexing, Bouchaib, Siddarth, Hussain, Shotaro, Nick, Junhai, Aayush, Lele, and Silvia**, thank you for the good times. It was a pleasure to discuss many ideas with you and learn from you. My dear Delphi friends, you made the Delphi meetings in Houston and The Hague, the workshops, the EAGE and SEG conferences special. Until we meet again, keep up the momentum.

Prof. Tariq Alkhalifah, I enjoyed our day-out during your casual visit to the Netherlands. Thank you for enlightening me on the capabilities of machine learning. It did not

end up in my dissertation, but it's on my to-do list.

Outside of my work environment, I had the chance to make friendships that I cherish. **Ahmed, Rayan, Mahmoud, Fouad, Yasser, Mohammed, Hamza, Abdullah, Abdulrahman, Nawaf, Ibrahiem, Rami, Nawaf, Hassan, Saleh and Hamad**, thank you for the good times and for your insights on topics outside of my field. It was interesting to learn about policy management, architecture, 6G, urban planning and business administration. **Dr. Ahmed Falimban** thank you for organizing the amazing social events that allowed us to connect every now and then. DSTV Obvious, Stijn, Owain and Joeri, thank you for organizing the tennis hitting sessions.

I'm grateful to **The British School in The Netherlands** for providing my twin daughters with proper education and fostering a safe learning environment. Leen & Yasmeen started their learning journey at your school, and it was nothing but great. Thanks to all the amazing teachers for their hard work. Through the school, we also had the chance to become friends with families from diverse backgrounds that we appreciate. Thank you **Ingrid, Christian and Ivett** for the good times. We are also grateful to **Zein** for the after school care. As working parents, their support was crucial for us to carry on with our tasks. **Our neighbours**, especially **Martin, Ria and Yangling**, thank you for making us feel home.

We extend our gratitude to the cities of Delft and The Hague for welcoming us. My connection with the Netherlands started during layovers at Schiphol Airport between Dammam and Houston when I was pursuing my bachelor's degree. Despite not leaving the airport, I fell in love with the country. I eventually visited it with my wife during our first trip as a couple. But I never imagined that I would live in the Netherlands. I guess felling in love with it had an influence. In this country, I learned to appreciate Art after my visit to the Rijksmuseum, which became one of my favourite places to visit. For the first time, my daughters had the chance to plant tulips that kept growing every year. We cherish the beautiful memories we have created in the Netherlands, and we are grateful for that.

My friends back home, my brothers **Ayman, Hassan, Saud, Saleh and Hani**, thank you for your messages and phone calls during my time abroad. Despite being away for a long time, we kept communicating with each others. I'm grateful for having you in my life.

I would like to take a moment to thank my tennis coach **Sherif Qadry** who taught me how to play tennis since I was 7. Playing tennis alongside pursuing my Ph.D. provided a healthy work-life balance. Dear coach, thank you for teaching me the skills to play tennis at a high level. I cherish the memories we created during the trainings and tournaments. You were not only a coach, but also part of our family. I look forward to getting back to the courts with you and sharpening my tennis skills.

TO MY FAMILY

My dear **family**, I owe you the world. Your endless love keeps me motivated and fills my heart with gratitude. Thank you for your unwavering support. Without you, I wouldn't be where I'm today.

Dear **Mom and Dad**, I want to express my deepest gratitude for all that you have provided me. You are the source of all my achievements. Your sacrifices paved the way for my success. I wouldn't be in the position I am today without your invaluable guidance,

boundless love, unwavering support, and encouragement. Thank you for visiting us in the Netherlands, which we thoroughly enjoyed it. Rowing boats in the canals, feeding goats in the farm and playing football in the playground are memories we will forever cherish. Even though the trip was short, it was worth every minute. Going with me to school to pick up Leen & Yasmeen reminded me of the days when you dropped me off and picked me up from school. That was the start of my journey. Thank you **Dad** for sending us to the best schools, and always wanting the best for us. Thank you **Mom** for spending your afternoons following up with our learning tasks and making sure we are healthy. Thank you from the bottom of my heart for being the pillars of my life.

I'm deeply grateful to my **grandmas**, who I dearly miss. They showered me with love and care while I grew up in their houses. I was saddened by the passing of my grandma during my PhD. Dear grandma, despite living a challenging life, you were an exemplary role model, teaching invaluable lessons through your actions. During our phone conversations, you always asked when I was returning back home. That alone made me feel your love as if I were sitting next to you. Rest in peace dear grandma. Your memory forever will be cherished.

My dear **Mother and Father-in-law**, your constant encouragement and support have been instrumental in this journey. Thank you for your constant calls to check on us. We are grateful for your love and guidance. Thank you also for hosting us and for your generosity during our visits back home. We deeply appreciate your kindness and welcoming spirit. Thank you for being a source of inspiration and strength throughout this academic endeavour.

I extend my heartfelt gratitude to my dear **brothers and sisters-in-law, uncles, aunts, and cousins** for their unwavering love and support throughout my journey. Thank you for checking on us with phone calls and messages every now and then. Thank you for the joyous moments and shared laughter. Knowing that I have your support makes me comfortable taking risks throughout my journey.

My dear brother **Abdullah**, I'm fortunate to have you in my life. You have been my go-to person whenever I needed help, and you never let me down. Thank you so much for your endless love and support. It's always fun times whenever we go out together. Also thank you for visiting us in the Netherlands. We were happy to have you, especially Leen & Yasmeen who felt the joy of being with you. We are proud of your tennis and padel accomplishments nationally and internationally. Just an advice, improve your tennis skills as I'm coming back soon.

My dear sister **Dr. Fatimah**, we are grateful for having a sister in our family. Thank you for your love and kindness. We enjoyed visiting you when you were in Poland. Thank you for showing us around Wroclaw and for your hospitality. Despite the health complications and challenging times you faced in Poland, you made it with great success. Even though it was tough, yet you always showed a smiling face. We are so proud of you.

The younger generation, my dear brothers **Hussain, Ahmed, and Hassan** thank you for your support and love. I like to listen to your advices and suggestions as you think bigger than your age. Well, sometimes you don't make sense, but most of the times, you do. Your visit to the Netherlands was joyful. Hussain, thank you for showing me around Denver during the SEG meeting. We are already proud of your achievements and I'm sure you will all make us even more proud in the future.

My beloved wife **Dr. Sukainah Alfaraj, MD.**, I'm grateful to have you in my life. We've shared nearly 12 wonderful years together. Throughout this journey, you have always encouraged me and supported me to achieve my goals without hesitation. My dearest **Suka**, I'm forever in debt to you for the freedom you gave me to pursue my dreams. Thank you for keeping up with my busy schedule and shouldering many of the responsibilities. Without your support, I wouldn't have been able to reach where I'm today. Thank you for your patience with rescheduling our vacation trips when I needed more time to work. I hope that the spectacular mountains, lakes and beaches made up for the change of plans. You probably noticed that I included MD. following your name. That's not only because it's the title for medical doctors, but also to make space for your soon-to-be extra title (Ph.D.). I give up on the discussion and admit that Dr. is the right title for medical doctors :-). I'm proud of you and honored to be your husband.

My beloved twins **Yasmeen & Leen**, you are my joy and happiness. You are smart and beautiful shining stars. I'm grateful for having you in my life and I'm proud of you. Since your birth, I have been enjoying every stage of your journey. Btw, thank you for waiting to be born just few days after I arrived in the US following the defence of my master's degree. And now, thank you for your understanding during my busy schedule. I'm grateful for the COVID-19 lockdowns as they allowed me to spend more with you. I had the chance to start teaching you tennis and we prepared sushi together. You definitely enjoyed that period. Well, you also made the lockdowns fun for me. Now that I'm near the end of my PhD, I hope I'll be able to spend more time with you in the future. The future is bright.

Ali M. Alfaraj

CURRICULUM VITÆ

Ali M. Alfaraj

23-09-1987 Born in Dammam, Saudi Arabia

EDUCATION

2019 – 2024 PhD candidate
Department of Imaging Physics, Delft University of Technology, The Netherlands

2012 – 2015 M.Sc. in Applied Geophysics
TU Delft, The Netherlands
ETH Zurich, Switzerland
RWTH Aachen, Germany

2006 – 2010 B.Sc. in Geophysics (cum laude)
University of Houston, The United States of America

RESEARCH EXPERIENCE

2010 – present Geophysicist
EXPEC Advanced Research Center, Saudi Aramco, Saudi Arabia

2015 – 2018 Researcher
University of British Columbia, Canada

LIST OF PUBLICATIONS

JOURNAL PAPERS

1. **A. M. Alfaraj**, D. J. Verschuur, and F. J. Herrmann, Low-rank-based residual statics estimation and correction, *Geophysics* 88 (2023), pp. V215–V231.
2. **A. M. Alfaraj**, D. J. Verschuur, and F. J. Herrmann, Seismic data reconstruction in complex near-surface regimes, undergoing review, *Geophysics*.
3. **A. M. Alfaraj**, D. J. Verschuur, and F. J. Herrmann, 3D model-independent rank-based near-surface correction, undergoing review, *Geophysics*.
4. D. Zhang, D. J. Verschuur, M. Davydenko, Y. Chen, **A. M. Alfaraj**, and S. Qu, Local primary-and-multiple orthogonalization for leaked internal multiple crosstalk estimation and attenuation on full-wavefield migrated images, *Geophysics* 86 (2021), pp. A7–A13.

CONFERENCE PAPERS

1. **A. M. Alfaraj**, and D. J. Verschuur, Towards 3D near-surface correction without NMO — A rank-based approach, in 84th EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2023)
2. **A. M. Alfaraj**, and D. J. Verschuur, Land Near-Surface Pre-Processing for RTM and LS-RTM, in 83rd EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2022)
3. **A. M. Alfaraj**, D. J. Verschuur, and F. J. Herrmann, Residual statics correction without NMO — A rank-based approach, in First International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists, 2021.
4. **A. M. Alfaraj**, and D. J. Verschuur, Mitigating the Near-Surface Effect in Velocity and Reflectivity Estimation with Multi-Scale Low-Rank Approximated Image Updates, in 82nd EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2021)
5. **A. M. Alfaraj**, D. J. Verschuur, and F. J. Herrmann, Non-surface-consistent short-wavelength statics correction for dense and subsampled data: A rank-based approach, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2020.
6. **A. M. Alfaraj**, M. Almubarak, F. J. Herrmann, Correcting for short-wavelength statics with low rank approximation, in 81st EAGE Annual Conference & Exhibition, (European Association of Geoscientists & Engineers, 2019.)

7. **A. M. Alfaraj**, R. Kumar, and F. J. Herrmann, Seismic waveform inversion using decomposed one-way wavefields, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2018.

