# Diving wave tomography:
# Velocity modelling using first arrival traveltime

Amatul Syafi Abdul Basit[1,*], Abdul Rahim Md Arshad[1], Arulini Permalu[2]

[1] Centre of Excellence in Seismic Imaging & Hydrocarbon Prediction (CSI), Universiti Teknologi PETRONAS,
32610 Bandar Seri Iskandar, Perak, Malaysia
[2] Department of Geoscience, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia
* Corresponding author email address: amatulsyafi@gmail.com

**Abstract:** In hydrocarbon exploration, information carried by diving waves and post-critical reflections that are used to reconstruct the long-to-intermediate wavelength of the subsurface is an integral part of successful velocity model building. Diving wave tomography (DWT) is one of the tools for shallow velocity assessment particularly when seismic data has poor signal-to-noise ratio (SNR) with complex geologic settings where no clear reflector is present. Considering the relationship between velocity with time and space, the output from tomography plays a crucial role to align data between time and depth domain and produce a reliable image of the deeper structure where hydrocarbon reservoir is typically located. In geophysics, tomography is primarily used to correct seismic trace alignment to produce a reliable stack section. In advanced imaging it is used as an initial model for waveform inversion in an integrated workflow. In the post-processing stage, it is used to correct the misfit between well logs and seismic data and is crucial for the quantitative analysis of rock physics. In this paper, we focus on tomography and its working principle on near-surface velocity modelling. We restricted our workflow to 2D synthetic data simulating the shallow gas occurrence that is prominent in the offshore Malay Basin to demonstrate how tomography works in velocity reconstruction. Results from synthetic and real data example shows that DWT can recover local large-scale structure and improved stacked data, considering no other seismic data and constraint from well data is included in the iterative process.

**Keywords:** Traveltime tomography, diving wave, shallow gas, forward modelling, traveltime inversion, Malay Basin

**Abstrak:** Dalam penerokaan hidrokarbon, maklumat yang dibawa oleh gelombang menerjun dan pantulan pasca-kritikal adalah sangat penting dalam pembinaan model halaju dalam bumi. Tomografi gelombang menerjun (DWT) adalah salah satu teknik penilaian model halaju cetek terutamanya apabila data seismik di kawasan kajian mempunyai nisbah isyarat-hingar (SNR) yang lemah dengan tetapan geologi yang kompleks di mana tiada pembias yang jelas. Hasil model dari tomografi memainkan peranan penting untuk menyelaraskan data antara domain masa dan ruang, supaya gambaran seismik untuk proses tafsiran struktur takungan hidrokarbon di bahagian dalam bumi boleh lebih dipercayai. Dalam bidang geofizik, tomografi digunakan sebagai model awal untuk inversi seismik, untuk membetulkan jajaran seismik semasa pemprosesan data, dan untuk penyelarasan antara log telaga dan seismik yang akan digunakan dalam analisis kuantitatif seterusnya. Dalam makalah ini, kami akan fokus kepada prinsip kerja tomografi dan aplikasinya terhadap data 2D sintetik dan data nyata dari Lembangan Melayu. Hasil dari tomografi menunjukkan bahawa DWT dapat memperbaiki hasil susunan seismik dan memulihkan struktur berskala besar dalam gambaran seismik bumi.

**Kata kunci:** Tomografi, gelombang pembiasan, gas cetek, model seismik, inversi seismik, Lembangan Melayu

## VELOCITY IN SEISMIC DATA PROCESSING

Seismic velocity is defined as the rate of change of wave displacement with respect to time. Displacement represents the relative change in position of rock particles with respect to time as the seismic wave travels through the subsurface from source to receiver. The correlation is expressed as the following average velocity equation,

$$v = \frac{(\Delta u(R,t))}{\Delta t} \quad [m/s] \tag{1}$$

Where u is the displacement function of position vector $R$ with respect to time $t$ (Pujol, 2003). By definition, velocity is essential in linking both time-domain and space-domain data in seismic processing and interpretation. In seismic data analysis, velocities can be measured directly from borehole measurements or derived indirectly from seismic data itself. Borehole measurements are obtained during or after drilling operation and it comes in the form of vertical seismic profile (VSP), check shot survey and sonic logs to acquire acoustic wave velocity or shear wave velocity from ocean bottom surveys. Through well-seismic-tie, borehole

measurements calibrate and improve the odds of hydrocarbon predictions by significantly elevating the reliability of seismic data for both qualitative and quantitative analysis (Nanda, 2016a). Different types of velocity, e.g. interval, average and root-mean-square (RMS) velocity can be derived by cross-referencing and integrating both seismic and well data acquired from field survey.

In the offshore marine environment, subsea topography and complex subsurface structures result in velocity anomalies that directly affect the arrival time of seismic wave detected by hydrophones. Geologically, seismic velocity is affected by these anomalies may arise from the rapid change in lithology, differential compaction from overburden pressure and structural discontinuity such as fault, fracture or even carbonate pore system. As pointed by Nanda (2016b), common pitfalls associated with velocity are the push-down and pull-up effect caused by the discrepancy between recorded traveltime and the actual depth of the structure. It often appears as pseudo-structures such as false syncline or anticline in a seismic section; or perhaps as fault shadow that may blur the line between normal and listric fault. What we see in a seismic image may not be the direct representation of subsurface geology. Precautionary steps are implemented in the oil and gas industry during and after the seismic processing stage, where velocity information act as the crux of those steps, including but not limited to the iterative-based tomography, normal moveout (NMO), static correction, migration and seismic-well-tie (Yilmaz, 2001). In upstream drilling operations, accurate velocity modelling is particularly important as to avoid losses incurred should there be errors in the depth estimation of reservoir pay zone and volumetric analysis for hydrocarbon production.

In the processing stage, the only reliable velocity is the velocity that gives the ideal stack (Yilmaz, 2001). For a relatively simple geological setting, one can first achieve optimal stack through velocity analysis qualitatively using constant velocity gather or quantitatively based on the semblance analysis. The latter involves the estimation of NMO velocities at discrete common midpoint (CMP) interval, beginning with large interval of few kilometers down to tens of meters. This velocity is then used for migration which involves the repositioning of seismic wavelet so that the maximum amplitude coincides with reflectors or in geological sense, the interface of rock layers. This two-step workflow is commonly known as pre-stack time migration (PreSTM). For a more complex geology such as salt diapirs and thrust fold belt, velocity model is obtained through tomography for pre-stack depth migration (PreSDM). Considering the origin of tomography in radiology, the term itself generally applied to methods that estimate properties of model through inversion (Nolet, 2008). Another method is theoretically an extension of tomography, known as full waveform inversion (FWI), which not only considers the kinematics of seismic data, but also considers the amplitude and phase of the seismic waveform to produce a high-fidelity velocity model (Virieux & Operto, 2009).

## SEISMIC IMAGING THROUGH TOMOGRAPHY
### Concept of diving wave

In the imaging process, near-surface velocity model is said to be essential to image deeper structure (Taillandier *et al.*, 2009). Depending on the purpose of modelling, near surface constitutes depth of up to 100 m for uphole surveys and mining engineering while for hydrocarbon exploration, near surface may exceed 100 m in depth. In such cases, near surface modelling act as a mean to focus an image and to eliminate traveltime anomalies (Bridle, 2016). There are methods available to resolve these anomalies: elevation correction to reduce the effect of topography (Widess, 1946), refraction statics associated with the base of weathering layer using plus-minus method (Hagedoorn, 1959) or generalized reciprocal method (Palmer, 1981). Eventually as algorithms and computer hardwares develop, the concept of diving wave became crucial in tomography, and together with other information contained in seismograms, used for FWI which employs full wave equation modelling especially for area as complex as Marmousi and the Valhall oil field in the North Sea (Virieux & Operto, 2009).

The linear relationship of velocity with distance travelled by seismic wave (ref. Equation 1) implied a constant gradient medium that isn't always the best representation of Earth's subsurface (Levin, 1996). As such, refraction theory governed by Snell's Law and represented by layered model is inadequate to best describe the non-linear behavior of first arrivals caused by geological complexity of near surface structure. This may lead to incorrect static correction and inaccurate raypath prediction, and by extension inaccurate traveltime calculation. Small error in the initial stage of velocity modelling can consequently introduce large error in subsequent iteration, thus severely deteriorates imaging quality at depth. Turning ray model, commonly known as diving wave was proposed to accommodate non-linearity, albeit imperfectly, and act as an approximation to describe the vertical velocity gradient within a rock layer and to accommodate strong lateral velocity variation across distances (Zhu *et al.*, 1992). The concept of diving wave allows for more accurate approximation with a less stringent underlying assumptions compared to conventional refraction mainly for two reasons:

1) First break tends to be more visible than hyperbolic events in shot gather
2) First break need not to be linear

This is so that even when scattering or attenuation is severe, DWT may still be able to estimate weathering and velocities without geological framework or priori information (Stefani, 1995). Tests with synthetic and real data both had demonstrated that DWT yields better near surface velocity with more fidelity between observed and predicted model compared to refraction tomography (Zhu *et al.*, 1992). DWT is also known to be used in an integrated

workflow with reflection and surface wave tomography as well (Song *et al.*, 2014; Duret *et al.*, 2016).

## Concept of seismic tomography

In a conventional seismic reflection survey, the limited acquisition offset and frequency bandwidth of the source leads to the poor sensitivity of seismic image to large and intermediate wavelength (Virieux & Operto, 2009). For a relatively simple geological setting, velocity macromodel is generated using kinematic information before projecting corresponding seismic through migration. In a more complex setting, iterative approach for velocity modelling was proposed involving the minimisation of misfit function between measured and modeled data, i.e. tomographic inversion (Tarantola, 1984). The process involves the inversion of traveltime to image local-scale structure, all based on few classes of data recorded in acquisition process, namely: normal incidence reflection, wide angle reflection and refraction, teleseismic and local earthquake (Rawlinson & Sambridge, 2003). The relationship between traveltime and wave velocity (ref. Equation 1), together with the relative ease of extracting such information from seismogram are the incentive behind the routine use of traveltime tomography in the industry. The general workflow of tomography in seismic data processing is shown in Figure 1.

In tomography, forward modelling refers to the process of predicting the result of measurement given a complete description of a physical system, i.e. model parameter. Inversion on the other hand, is the process of using actual result of measurements to infer the values of parameters that describe a system (Tarantola, 2005).



**Figure 1:** A simplified workflow of tomographic inversion. In a real data application, observed traveltime is directly obtained from pre-processed shot gather, otherwise it is computed in forward modelling process.

A simple expression given by Rawlinson & Sambridge (2003), $d = g(m)$ describes the basis of tomography, where the measured travel time $d$ can be predicted for a given source-receiver array by line integration through a model, $m$. The model is described based on the elastic properties of the subsurface e.g. velocity and density. Through iterations, tomographic inversion seeks to minimize the difference between observed data ($d_{obs}$) and initial model estimation ($m_0$) until termination criteria is satisfied. As part of the tomographic inversion, the perturbation model is obtained by cross-correlation of forward-propagated wave with back-propagated residual wave, which is then added to the initial model. The updated model is then used as an input for subsequent iteration(s).
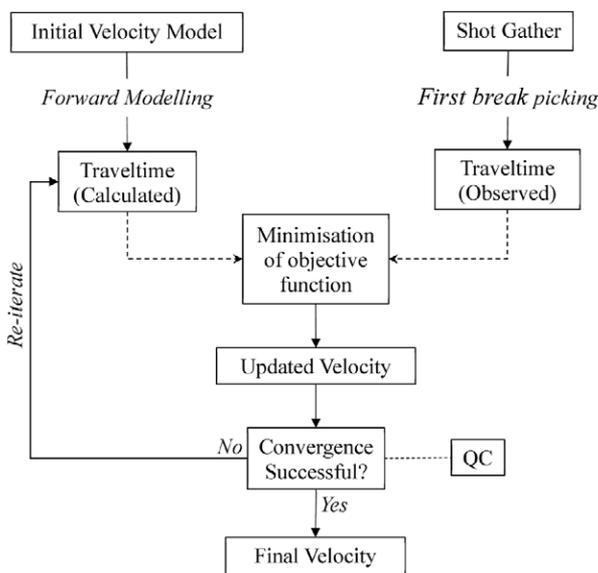
## Traveltime calculation in forward modelling

As the first half of the workflow, traveltime calculation for a given velocity field can be done either by ray tracing or by computing the numerical solutions of the eikonal equation. In ray tracing, raypath trajectory can be described using the following kinematic ray equation,

$$\frac{d}{ds}\left[\frac{1}{v(x)}\frac{dx}{ds}\right] = \nabla\frac{1}{v(x)} \qquad (2)$$

Where $s$ is the arc length of wavefronts and $v(x)$ is velocity at point $x$. Equation 2 works based on Snell's law at grid boundaries and the Fermat's principle for a given source and receiver array. The eikonal equation on the other hand, governs the traveltime from a fixed shot position in an acoustic, isotropic and heterogeneous media. It can be written in two-dimensional form as follows,

$$\left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial z}\right)^2 = \frac{1}{v^2(x,z)} \qquad s^2 = \frac{1}{v^2} \qquad (3)$$

Where $t$ is the traveltime and $s^2$ is the slowness-squared or the inverse of velocity. Ray tracing can handle multi-arrivals due to its simplicity and robust applicability, making it a widely used method in body wave tomography. Major drawback is that the interpolation usually suffers from shadow zones in complex velocity media (Vidale, 1990; Aki & Richards, 2009). To overcome this issue, the numerical method derived from Equation 3 is preferred. Methods such as finite-difference is able to describe wave motion in a discretized media with spatial variation of elastic property. Despite being computationally demanding, it became the norm in the industry on account of the perpetual advance in microprocessor technology to break down complex problems into simpler forms. Mathematically, finite-difference works by approximating the solution to the eikonal equation at nodal point of each grid in a mesh using Taylor series expansion. Two popular finite-differencing method is the fast-marching method (FMM) (Sethian & Popovici, 1999) and the fast-sweeping method (FSM) (Zhao, 2005). FMM
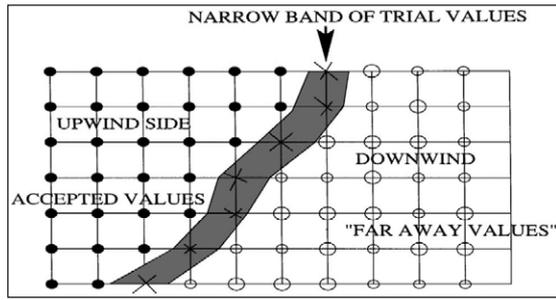
**Figure 2:** The schematic of upwind scheme of finite difference method in a discretised grid mesh. The narrowband contain the active nodes for which the approximation to the eikonal is updated through iterations and forms the propagating wavefront with time. Image extracted from Sethian & Popovici (1999).

employs the upwind scheme whereby the traveltime field propagating in the downwind direction is solved within the narrowband based on the known value in the upwind direction (Figure 2). FSM on the other hand, avoids the use of narrowband and group propagating wavefront in a prefixed number of direction from the source, making it dependent on the refractive index of the medium. Both method aims to achieve optimal efficiency and relies on iterative update to compute first-arrivals in a given media based on the recovered causality (Li & Fomel, 2013).

**Traveltime inversion**

The second half of the workflow is the least-square ($l_2$) minimization of the cost function, aka the objective function in a loop of tomographic inversion. One can solve it explicitly as linear analytical problem or take the numerical method approach whereby the implicit nonlinear systems of the optimization problem is solved iteratively. The latter is often preferred in engineering applications since it yields an approximation for complex multivariable functions. In traveltime tomography, inversion aims to reach the global minima of traveltime data residual. Convergence is said to be achieved when minima is reached i.e. the difference between predicted and calculated traveltime is optimally small. Derivative-based method is one of the way to compute the gradient of cost function so that it descent to minimum point. Commonly known as gradient-based search method, it assumes the function is continuous and differentiable at least up to second-order derivative.

As explained by Arora (2017), the gradient-based method can be generally expressed in vector and component form as follows,

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)} \quad k = 0, 1, 2.... \quad i = 1,2...n \quad (4)$$

$$x_i^{(k+1)} = x_i^{(k)} + \Delta x_i^{(k)} \quad (5)$$

$$\Delta x^{(k)} = \alpha_k d^{(k)} \quad (6)$$

Where $x_i$ is the component of function $x$, $i$ is the $i^{th}$ variable in the function $x$, $k$ is the number of iteration with $x^{(0)}$ being the initial estimate or priori, $\Delta x^{(k)}$ is the change in estimate or perturbation at $k^{th}$ iteration. At each iteration, the $\Delta x^{(k)}$ can be determined if step size $\alpha_k$ and the descent direction of gradient $d^{(k)}$ is known. In order to satisfy the optimality condition whereby a point along the cost function is indeed the verifiable minimum point, the following condition for the descent direction $d^{(k)}$ must be held true in general,

$$\alpha_k(c^{(k)} \bullet d^{(k)}) < 0 \qquad c^{(k)} = \nabla f(x^{(k)}) \quad (7)$$

Where $c^{(k)}$ is the gradient of cost function $x$ at point $x^{(k)}$. The step size $\alpha_k$ here is always positive scalar value and can be omitted from the descent condition regardless. Recall that the rule of a negative dot product implies a cosine angle of more than 90° between the two vectors, hence Equation 7 can be geometrically described as the descending gradient going in the opposite direction, away from both vector $c^{(k)}$ and $d^{(k)}$ stepping at $\alpha_k$ in each iteration.

Now that the descent condition is set, there are two well-known method of finding the search direction $d^{(k)}$ – the steepest-descent and the conjugate gradient method. The former being the simplest, oldest known numerical method for unconstrained problems introduced by Cauchy in 1847. It seeks the direction in which the gradient of cost function decreases most rapidly. The direction of steepest descent and descent condition is expressed as follows,

$$d_i^k = -c_i^k = -\frac{\partial f(x^k)}{\partial x_i} \qquad i = 1, 2...n \quad (8)$$

$$(c \bullet d) = -\|c\|^2 < 0 \qquad k = 0, 1, 2... \quad (9)$$

Conjugate gradient on the other hand, is a modified steepest-descent method that significantly improves the rate of convergence by adding a scaled direction $\beta_k$ in all subsequent iteration after the first (Hestenes & Stiefel, 1952). The conjugate direction and descent condition is expressed as follows,

$$d_i^k = -c_i^k + \beta_k d^{k-1} \qquad i = 1, 2...n \quad (10)$$

$$(c \bullet d) = -\|c\|^2 + \beta_k (c \bullet d^{k-1}) \quad k = 1, 2... \quad (11)$$

Substituting $d$ from Equation 10 into Equation 7 would satisfy the descent condition of less than zero. Unlike steepest-descent method, the descent direction of conjugate gradient is not orthogonal to each other, rather it cuts diagonally towards possible minima thus considerably reducing the number of iterations needed to convergence. Due to this reason alone conjugate gradient is always preferrable over steepest-descent method (Arora, 2017).

## APPLICATION OF DWT ON SYNTHETIC DATA

Based partly on the reproducible Madagascar script by Li *et al.* (2013) on first-arrival traveltime tomography, synthetic data is used to simulate diving wave propagation through a low velocity anomaly by means of finite-difference forward modelling based on a modified eikonal solver of Equation 3. In the inversion step, the minimisation of misfit function is solved directly as nonlinear optimization problem to compute the conjugate gradient of the cost function using the adjoint-state method (Plessix, 2006).

For the simulation, a 2D velocity model with constant velocity gradient and Gaussian anomaly is generated (Figure 3). The synthetic model is measured at $4 \times 2.5$ km² dimension with sampling grid of 0.008 km. The constant velocity gradient model is expressed as

$$v_z = v_0 + kz \tag{12}$$

Where $v_z$ is velocity at depth $z$ and $k$ is the gradient constant. Velocity at zero depth is 1.5 km/s and $k = 3.0$. For demonstration on traveltime calculation (Figure 4), the following double-square root (DSR) eikonal equation from Li *et al.* (2013) is used,

$$\frac{\partial t}{\partial z} = -\sqrt{\frac{1}{v^2(z,s)} - \left(\frac{\partial t}{\partial s}\right)^2} - \sqrt{\frac{1}{v^2(z,r)} - \left(\frac{\partial t}{\partial r}\right)^2} \tag{13}$$

Where $s$ and $r$ are the source and receiver location along offset $x$ respectively. The negative component of $s$ and $r$ implies the downward pointing of the velocity (or slowness) vectors. The boundary condition given is when $t = 0$, $x = s = r$. Concurrently, ray tracing is used to demonstrate the trajectory of diving waves through both the constant velocity gradient model and the anomaly model (Figure 3), where the effect of low-velocity zone on traveltime can be seen and appeared as shadow zone.

The next step is to run inversion using the adjoint-state technique coupled with gradient-based approach to solve non-linear system and minimise least square ($l_2$) misfit of the cost function. For a given velocity field, the cost function $E$ of data residual is defined as,

$$E(w) = \frac{1}{2}(t - t^{obs})^T(t - t^{obs}) \qquad w \equiv \frac{1}{v^2} \tag{14}$$

Where $w$ is slowness-squared, $t^{obs}$ and $t$ is the observed traveltime and calculated traveltime from the shot-indexed eikonal equation respectively. The transpose of the data residuals implies the adjoint-state computation of the gradient vector. In a linear system, the gradient vector of Equation 14 can be expressed as,

$$\nabla_w E \equiv \frac{\partial E}{\partial w} = \sum_{S=1}^{nx}(Js)^T(t - t^{obs}) \qquad s = 1, 2, ..., nx \tag{15}$$
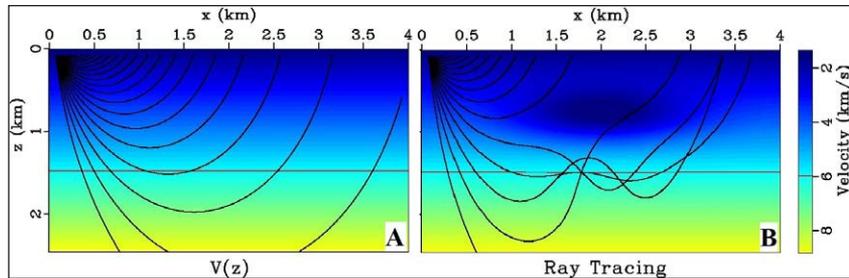


**Figure 3:** The initial velocity model (A) and the anomaly model (B) are overlaid with ray trajectories from shot point (0,0). The presence of low velocity anomaly simulating gas deflects projected ray paths and causes irregularities in traveltime arrival on the receiver. In (B), the area with the absence of traveltime response at offset between 1.6 km to 2.9 km is known as the 'shadow zone'.
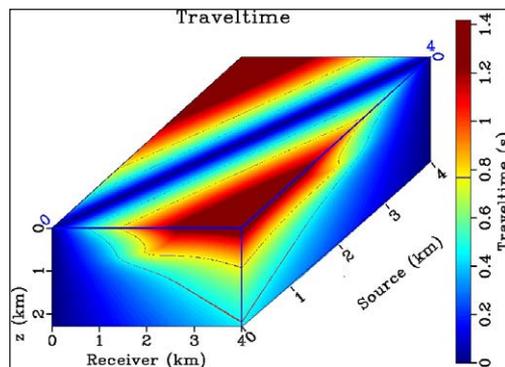


**Figure 4:** The forward modelling of anomaly model by modified eikonal solver with FMM (ref. Eq. 13). Note that the boundary condition is when traveltime is zero, $x = s = r$. The delay in traveltime in the centre is caused by the low velocity anomaly present in the model in Figure 3.

Where $s$ is the number of shot along $x$-axis and $J$ denotes the Fréchet derivative (or Jacobian matrix) at $s^{th}$ shot, which is the partial derivative of state variable, traveltime ($t$) with respect to model parameter, slowness-squared ($w$).

Inversion is run with 5 number of linear iteration, 10 number of conjugate gradient iteration and Tikhonov reguralisation (Neubauer, 1989). The misfit values for the Gaussian anomaly model and Marmousi model are shown in Table 1 and Table 2 respectively, and the convergence rates are shown in Figure 5. From the inversion result seen in Figure 6A – 6C, it is clear that the boundary of the Gaussian anomaly is not well-defined. However, a certain degree of its velocity structure is able to be recovered with lowest velocity ±1.5 km/s is centred around coordinate (2.0, 0.8), thus showing consistency with the true velocity model. This result is also consistent with the Marmousi example where large-scale structures such as the dipping and strong lateral velocity change are recoverable within the limit of the source-receiver array and depth up to 1500 m (Figures 6D – 6F).

In this study, the main limitation is the available computational resources required to run both forward modelling and inversion, which is highly dependable on processing power, and storage if study area covers tens of kilometre square. For a full scale Marmousi model measuring at 9.204 × 3.004 km², finite-difference forward modelling on local machine alone would require hours of processing time and produces at least 150 GB worth of shot gather and wavefront snapshot for each shot, depending on the design of source-receiver array and discretisation parameter. Ideally, parallel computation of gradient for each independent shot is preferable. This can be achieved by using supercomputer to distribute tasks to clusters of processors, thus significantly improve efficiency. Otherwise, the model is restricted to within 4 × 4 km² dimension for convenience purposes. Furthermore, the synthetic simulation is deprived of any priori as opposed to real data application, where at least the second-pass of velocity analysis using semblance method would give some degree of information on stacking velocity for the algorithm to begin reconstruction.

**Table 1:** The least-square misfit at each iteration of tomographic inversion for the Gaussian anomaly model.

| Number of iterations | $L_2$ |
|---|---|
| 1 | 1.0000 |
| 2 | 0.4230 |
| 3 | 0.3865 |
| 4 | 0.3044 |
| 5 | 0.2148 |

**Table 2:** The least-square misfit at each iteration of tomographic inversion for the Marmousi model.

| Number of iterations | $L_2$ |
|---|---|
| 1 | 1.0000 |
| 2 | 0.3759 |
| 3 | 0.3424 |
| 4 | 0.2346 |
| 5 | 0.0674 |
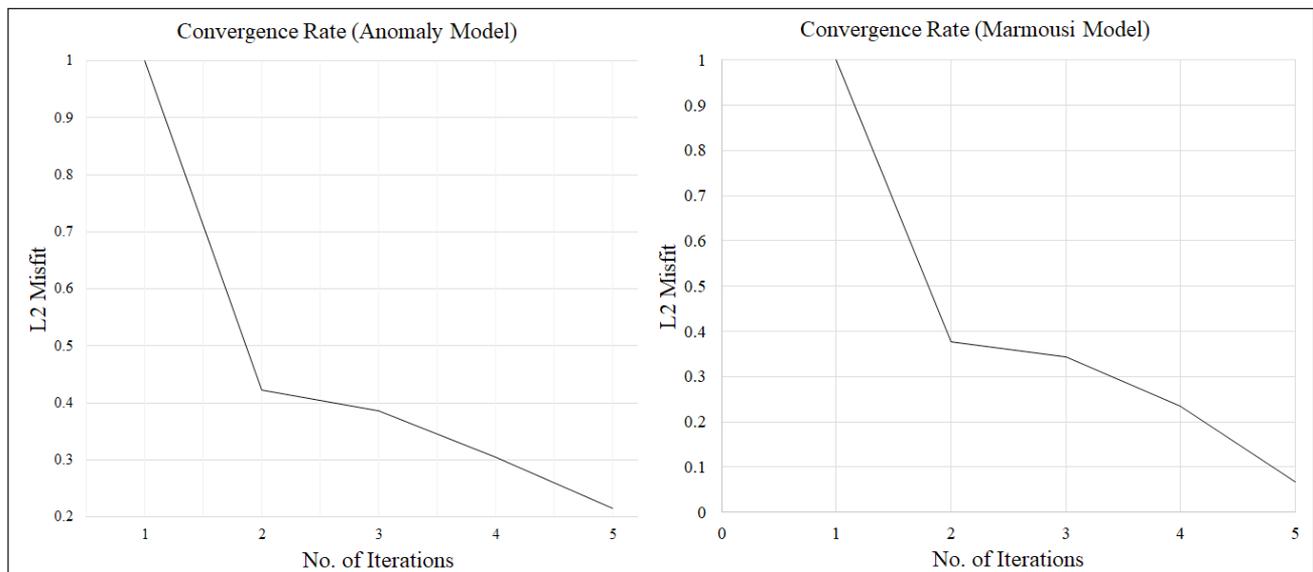


**Figure 5:** The reduction in the $L_2$ misfit over 5 iterations of the anomaly model from Figure 6A – 6C and Marmousi model from Figure 6D – 6F. No significant drop can be observed after the fifth iteration, which indicates that a global minimum is reached. If the final velocity model resembles the original model, convergence can be considered successful, otherwise parameter adjustments and/or additional constraints are needed to solve the inverse problem.
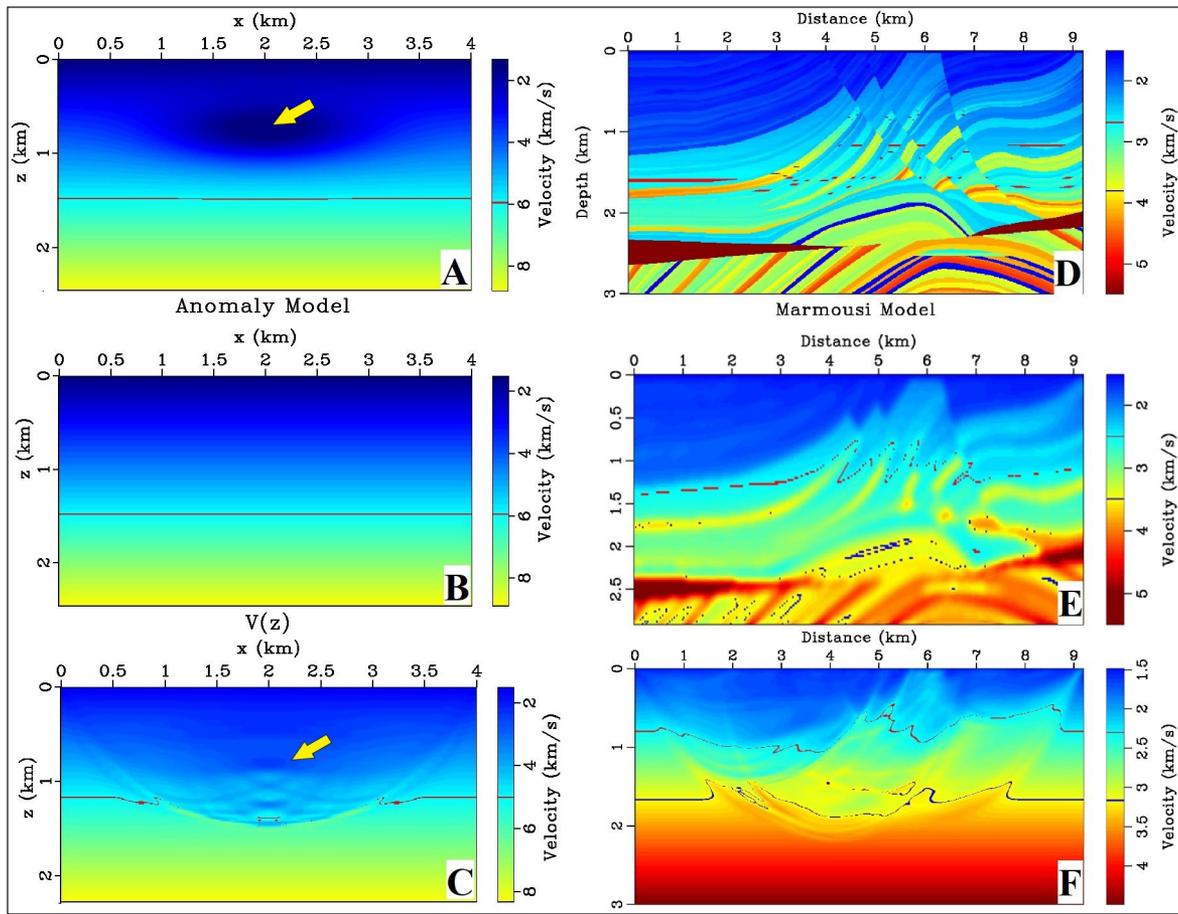
**Figure 6:** Gaussian anomaly model (A) with low velocity region centred at (2.0, 0.8) as indicated by the arrow, a constant gradient velocity model (B), used as a priori for the tomographic inversion and the final velocity model (C). The synthetic Marmousi model (D) commonly used for blind testing in geophysics and the smoothed version (E) used as reference for the analysis, and the final velocity model (F) after the DWT inversion.

## APPLICATION OF DWT ON OFFSHORE MALAY BASIN

A single seismic line used to run DWT is extracted from a 3D seismic volume acquired from the offshore Northern Malay Basin, where siliciclastic reservoirs are prone to natural-gas condensates with varying $CO_2$ concentration, both of organic and inorganic origin (Madon *et al.*, 2006). In a real data application, the tomographic inversion workflow differs slightly in preparing the initial velocity model. Whilst synthetic testing uses a constant gradient velocity model (ref. Equation 12), the initial model from 2D seismic line is generated based on the first arrival time picked at common-depth-point (CDP) gathers, which had undergone processing and preliminary velocity analysis.

In SeisSpace ProMAX, tomography works by simulating diving waves through the initial model. To cover the entire stretch of 17.3 km line, the shot and receiver interval is set at 12.5 m and 37.5 m respectively, totaling 1384 shot points each with 480 active receivers. After diving wave simulation, the predicted traveltimes are subtracted from the actual picked traveltimes to produce residuals (ref. Equation

14). Using all the projected ray paths and residuals in the inverse sparse matrix, the slowness field $w$ is adjusted to reduce misfit at each iteration. The final velocity model is then used for time migration and Kirchhoff depth migration to view the stacked response (Figure 7A – 7D).

In the PreSTM section (Figure 8), the presence of gas is indicated by time sagging, phase distortion as well as amplitude and frequency loss due to seismic attenuation. After the inversion, seismic traces located near the gas zones are time-shifted -20 ms to -30 ms, improving lateral continuity. In the depth migrated section shown in Figure 9, corrected moveout from the inversion relatively flatten CDP gathers at far offset, which resulted in more focused reflective events that are repositioned about 20 m upwards.

## CONCLUSION

Under ideal circumstances, such that a layer cake model or a simple geologic setting, the arrival times of a seismic shot gather would constitute first arrivals and hyperbolic curves as the reflective event. With the presence of low velocity anomaly in the shallow subsurface, the traveltimes recorded would
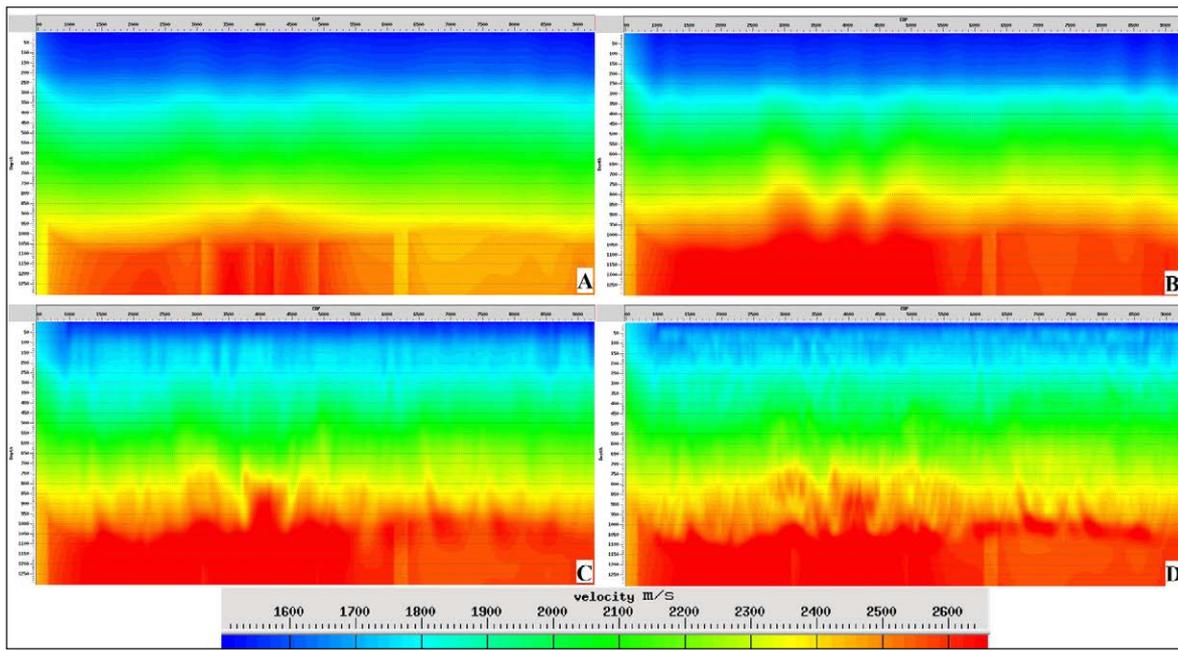
**Figure 7:** The initial model (A) derived from the first arrival time picked on seismic shot gathers and the first (B), the third (C), and fourth iteration (D) of traveltime inversion on the 2D line. At each iteration, diving wave simulation is done to generate traveltime for residual calculation. The low-velocity zone on the shallow region is refined at each iteration and the final velocity model is then used for migration.
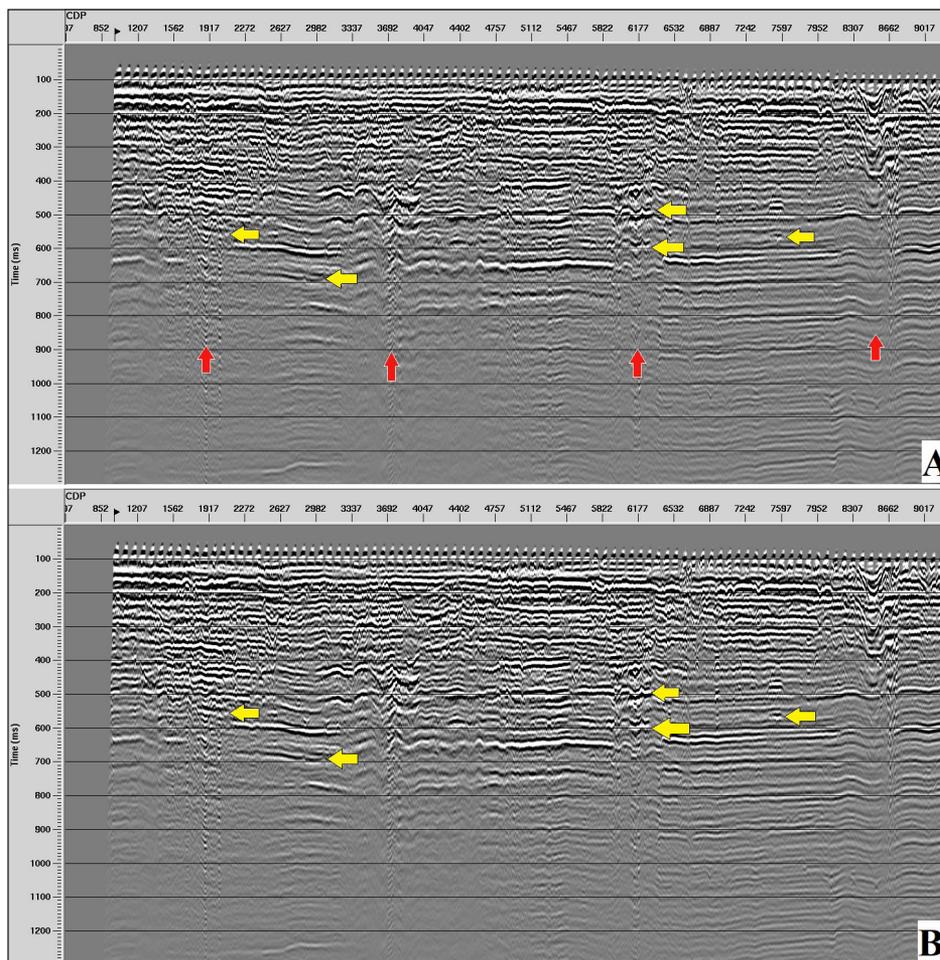


**Figure 8:** The time migrated section before (A) and after (B) incorporating the final velocity model derived from DWT. Major gas chimneys indicated by red arrow causes severe reduction in SNR and sagging. Improvement particularly in areas indicated by yellow arrow in terms of lateral continuity and flatten events.
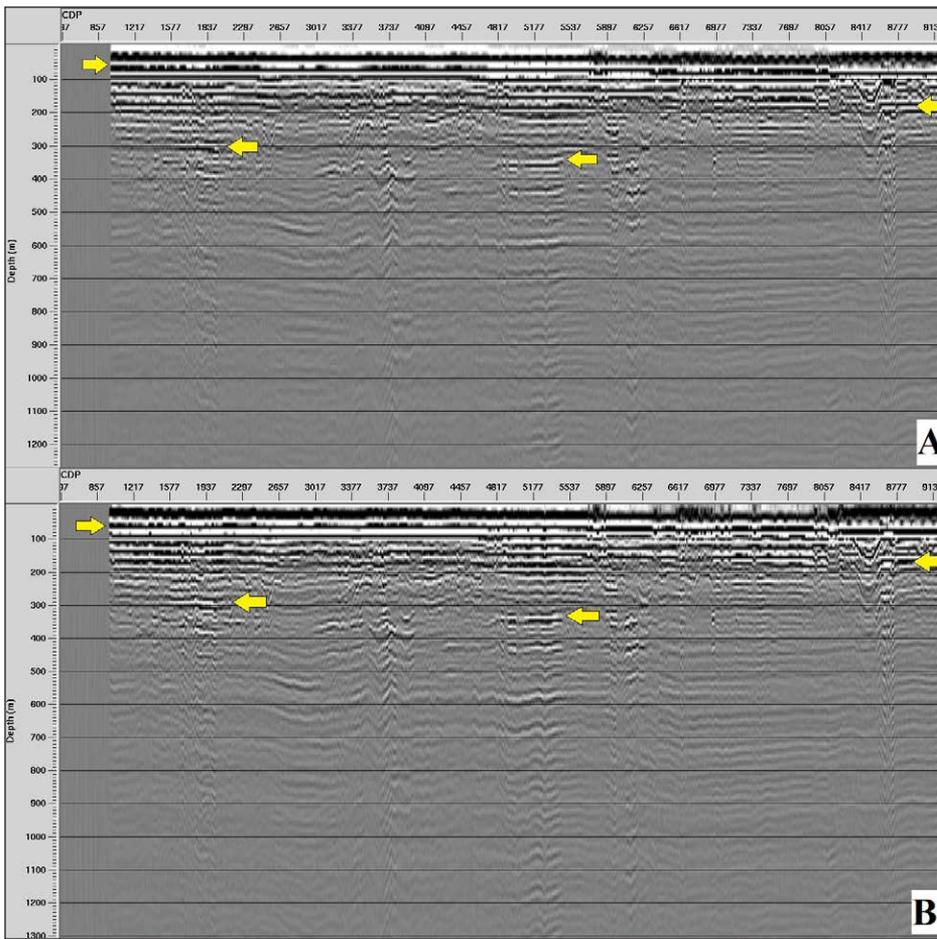
**Figure 9:** The seismic section before (A) and after (B) depth migration, which shows phase shift in the shallow region between 0 – 100 m depth, shifting seismic event up about 20 m and improved seismic continuity at places indicated by yellow arrow.

carry non-linearities in the traveltime-depth gradient. These non-linearities would affect the seismic imaging of deeper region if not resolved, rendering the resulting seismic section less reliable for subsequent geological interpretations and rock physical analysis. DWT was introduced to estimate shallow velocity model and minimize the effects of near-surface structures. Although DWT is only able to recover large-scale structure, the diving wave component is still an integral part of imaging because it carries crucial shallow subsurface information. Often there are ways to improve the resolution of final velocity model of tomography. For example, DWT can be used in conjunction with other forms of tomography that includes other information in the seismogram, such as reflected wave and surface wave to image structures from shallow to deeper region, improving the reliability of the overall seismic stack section after PreSDM. Well log data is also known to help further constraining the inversion process by compensating for the absence of low frequency content in the seismic amplitude spectrum. Now with the help of advancing technology, the computationally demanding FWI is becoming a routine in producing high-fidelity velocity model that is accurate enough to visualize discernable geological structure even before migration, and tomography helps in producing the priori model to be used in FWI.

## AUTHOR CONTRIBUTIONS
ASAB, as the main author, drafted the manuscript, carried out the synthetic data simulation with the expert guidance of ARMA, and the real data application with the help of AP.

## CONFLICT OF INTEREST
The authors have no conflicts of interest to declare that are relevant to the content and material of this article.

## REFERENCES
Aki, K. & Richards, P.G., 2009. Quantitative seismology. University Science Books, Herndon, Virginia. 742 p.

Arora, J.S., 2017. Introduction to optimum design (4th ed.). Elsevier Academic Press. 931 p.

Bridle, R., 2016. Near-surface modeling and imaging. The Leading Edge, 35(11), 938–939. https://doi.org/10.1190/tle35110938.1.

Duret, F., Bertin, F., Garceran, K., Sternfels, R., Bardainne, T., Deladerriere, N. & Le Meur, D., 2016. Near-surface velocity modeling using a combined inversion of surface and refracted P-waves. The Leading Edge, 35(11), 946–951. https://doi.org/10.1190/tle35110946.1.

Hagedoorn, J.G., 1959. The plus-minus method of interpreting seismic refraction sections. Geophysical Prospecting, 7(2), 158–182. https://doi.org/10.1111/j.1365-2478.1959.tb01460.x.

Hestenes, M.R. & Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standars, 49(6), 409–436.

Levin, F.K., 1996. Anatomy of diving waves. Geophysics, 61(05), 1417–1424. https://doi.org/10.1190/1.1444066.

Li, S. & Fomel, S., 2013. Kirchhoff migration using eikonal-based computation of traveltime source-derivatives. Geophysics, 78(4), S211–S219.

Li, S., Vladimirsky, A. & Fomel, S., 2013. First-break traveltime tomography with the double-square-root eikonal equation a. Geophysics, 78(6), U89–U101.

Madon, M., Yang, J.-S., Abolins, P., Abu Hassan, R., Yakzan, A. & Zainal, S.B., 2006. Petroleum systems of the Northern Malay Basin. Bulletin of the Geological Society of Malaysia, 49, 125–134. https://doi.org/10.7186/bgsm49200620.

Nanda, N.C., 2016a. Seismic data interpretation and evaluation for hydrocarbon exploration and production. In: Nanda, N.C., (Ed.), Seismic data interpretation and evaluation for hydrocarbon exploration and production. Springer, Cham. 304 p. https://doi.org/10.1007/978-3-319-26491-2.

Nanda, N.C., 2016b. Seismic pitfalls. In: Nanda, N.C., (Ed.), Seismic data interpretation and evaluation for hydrocarbon exploration and production. Springer, Cham. 304 p. https://doi.org/https://doi.org/10.1007/978-3-319-26491-2_12.

Neubauer, A., 1989. Tikhonov regularisation for non-linear ill-posed problems: Optimal convergence rates and finite-dimensional approximation. Inverse Problems, 5(4), 541–557. https://doi.org/10.1088/0266-5611/5/4/008.

Nolet, G., 2008. A breviary of seismic tomography. Cambridge University Press, New York. xiv+344 p.

Palmer, D., 1981. An introduction to the generalized reciprocal method of seismic refraction interpretation. Geophysics, 46(11), 1508–1518. https://doi.org/10.1190/1.1441157.

Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophysical Journal International, 167(2), 495–503. https://doi.org/10.1111/j.1365-246X.2006.02978.x.

Pujol, J., 2003. Elastic wave propagation and generation in seismology. Cambridge University Press, New York. 462 p. https://doi.org/10.1017/CBO9780511610127.

Rawlinson, N. & Sambridge, M., 2003. Seismic traveltime tomography of the crust and lithosphere. In: R. Dmowska, (Ed.), Advances in Geophysics, v. 46. Elseview, Amsterdam, London. 282 p. https://doi.org/10.1016/S0065-2687(03)46002-0.

Sethian, J. A. & Popovici, A.M., 1999. 3-D traveltime computation using the fast marching method. Geophysics, 64(2), 516–523. https://doi.org/10.1190/1.1444558.

Song, M., Lv, X., Han, X.L., Wang, D., Yang, S. & Hinz, C.E., 2014. Combining diving-wave tomography and prestack reflection tomography for complex depth imaging — A case study from mountainous western China. The Leading Edge, 33(8), 868–874. https://doi.org/10.1190/tle33080868.1.

Stefani, J.P., 1995. Turning-ray tomography. Geophysics, 60(6), 1917–1929.

Taillandier, C., Noble, M., Chauris, H. & Calandra, H., 2009. First-arrival traveltime tomography based on the adjoint-state method. Geophysics, 74(6), WCB1–WCB10. https://doi.org/10.1190/1.3250266.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics, 49(8), 1259–1266. https://doi.org/10.1190/1.1441754.

Tarantola, A., 2005. Inverse problem theory (2nd ed.). Society for Industrial and Applied Mathematics. 339 p. https://doi.org/10.1137/1.9780898717921.

Vidale, J.E., 1990. Finite-difference calculation of traveltimes in three dimensions. Geophysics, 55(5), 521–526. https://doi.org/https://doi.org/10.1190/1.1442863.

Virieux, J. & Operto, S., 2009. An overview of FWI in exploration geophysics. Geophysics, 74(6), WCC127–WCC152. https://doi.org/10.1190/1.3238367.

Widess, M.B., 1946. Effect of surface topography on seismic mapping. Geophysics, 11(3), 362–372. https://doi.org/10.1190/1.1437260.

Yilmaz, Ö., 2001. Seismic data analysis: Processing, inversion, and interpretation of seismic data (Vol. 1). Society of Exploration Geophysicists, Tulsa. 1028 p. (2nd ed.; S.M. Doherty, (Ed.)). https://doi.org/10.1190/1.9781560801580.

Zhao, H., 2005. A fast sweeping method for eikonal equations. Mathematics of Computation, 74(250), 603–627.

Zhu, X., Sixta, D. & Angstman, B., 1992. Tomostatics: Turning-ray tomograph + static corrections. The Leading Edge, 11(12), 15–23. https://doi.org/10.1190/1.1436864.