

Acquisition design for direct reflectivity and velocity estimation from blended and irregularly sampled data

Nakayama, Shotaro; Blacquièrè, Gerrit; Ishiyama, Tomohide

DOI

[10.1111/1365-2478.12841](https://doi.org/10.1111/1365-2478.12841)

Publication date

2019

Document Version

Final published version

Published in

Geophysical Prospecting

Citation (APA)

Nakayama, S., Blacquièrè, G., & Ishiyama, T. (2019). Acquisition design for direct reflectivity and velocity estimation from blended and irregularly sampled data. *Geophysical Prospecting*, 67(8), 2127-2146. <https://doi.org/10.1111/1365-2478.12841>

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.

Acquisition design for direct reflectivity and velocity estimation from blended and irregularly sampled data

Shotaro Nakayama^{1,2*}, Gerrit Blacquière¹ and Tomohide Ishiyama^{2,3}

¹Civil Engineering and Geosciences, Delft University of Technology, Building 23, Stevinweg 1, 2628 CN Delft, The Netherlands, ²INPEX Corporation, Akasaka Biz Tower 5-3-1, Akasaka, Minato-ku, Tokyo 107-6332, Japan, and ³ADNOC, P.O. Box 898, Abu Dhabi, United Arab Emirates

Received January 2019, revision accepted July 2019

ABSTRACT

Blended acquisition along with efficient spatial sampling is capable of providing high-quality seismic data in a cost-effective and productive manner. While deblending and data reconstruction conventionally accompany this way of data acquisition, the recorded data can be processed directly to estimate subsurface properties. We establish a workflow to design survey parameters that account for the source blending as well as the spatial sampling of sources and detectors. The proposed method involves an iterative scheme to derive the survey design leading to optimum reflectivity and velocity estimation via joint migration inversion. In the workflow, we extend the standard implementation of joint migration inversion to cope with the data acquired in a blended fashion along with irregular detector and source geometries. This makes a direct estimation of reflectivity and velocity models feasible without the need of deblending or data reconstruction. During the iterations, the errors in reflectivity and velocity estimates are used to update the survey parameters by integrating a genetic algorithm and a convolutional neural network. Bio-inspired operators enable the simultaneous update of the blending and sampling operators. To relate the choice of survey parameters to the performance of joint migration inversion, we utilize a convolutional neural network. The applied network architecture discards suboptimal solutions among newly generated ones. Conversely, it carries optimal ones to the subsequent step, which improves the efficiency of the proposed approach. The resultant acquisition scenario yields a notable enhancement in both reflectivity and velocity estimation attributable to the choice of survey parameters.

Key words: Acquisition, Optimization, Reflectivity, Velocity, Genetic algorithm, Neural network.

INTRODUCTION

During the last decade, blended acquisition has realized the industry's ambition towards efficient and cost-effective seismic operations that still attain the required data quality (Beasley, Ronald and Jiang 1998; Berkhout 2008; Bouska 2010; Abma *et al.* 2012; Nakayama *et al.* 2015). Furthermore, the enhancement in the survey productivity contributes to minimizing health, safety and environment (HSE) exposure in the field.

In general and where applicable, separation of overlapping source-wavefields comes to the initial step to process blended data rather than processing them directly (Moore *et al.* 2008; Lin and Herrmann 2009; Mahdad, Doulgeris and Blacquière 2011; Kontakis and Verschuur 2015). Deblended data subsequently enable us to use standard processing and interpretation practices.

While conventional seismic surveys aim ideally at regular and dense sampling to satisfy the Shannon–Nyquist theorem, further easing the spatial-sampling requirements contributes to the business aspect, and reduces the environmental

*E-mail: s.nakayama@tudelft.nl; shotaro.nakayama@inpeco.jp

footprint. For example, compressive sensing has recently been successfully implemented in the industry (Hennenfent and Herrmann 2008; Herrmann 2010; Mosher, Kaplan and Janiszewski 2012). This technique allows for a non-uniform spatial sampling along with fewer measurements in the field. Subsequent signal recovery from compressive measurements is generally posed as an underdetermined inversion problem. To solve it, compressive sensing exploits the sparsity of signals in a proper transform domain, allowing a small number of coefficients to sufficiently represent the signals that one attempts to obtain. The improvement in spatial-sampling efficiency, coupled with optimal data recovery, potentially enables us to design seismic surveys that satisfy demanding geophysical objectives at an affordable cost.

An alternative approach to handle blended data is to apply imaging without deblending (Tang and Biondi 2009; Verschuur and Berkhouit 2011; Berkhouit, Blacquièrè and Verschuur 2012; Chen *et al.* 2015; Caporal, Blacquièrè and Davydenko 2018). Instead of directly migrating blended data, these studies utilized a least-squares migration (LSM) scheme. The formulation of the imaging problem as a least-squares problem enables LSM to iteratively minimize the misfit between the real and the modelled blended data, which consequently produces the subsurface reflectivity without the separation of blended wavefields. Insufficient spatial sampling in the acquisition often induces aliasing noise in migration. A common practice to address this issue is to apply data reconstruction prior to imaging, yet alternatively, LSM can be used for this purpose. Nemeth, Wu and Schuster (1999) showed that the technique is capable of suppressing migration artefacts and producing images with balanced amplitudes and high resolution even when the input data suffer from coarse and/or irregular spatial sampling. Full-waveform inversion (FWI) is capable of using acquired seismic data to retrieve the subsurface properties that determine the seismic wavefield in an iterative manner (Tarantola 1984), that is, blended records can be directly used in FWI, which attempts to minimize the misfit between observed data and forward-modelled blended data (Florez, Mantilla and Ramirez 2016). Due to the heavy computational burden in FWI, the concept of blending is considered as one of the crucial strategies to improve its efficiency, even when the acquired data were not blended. It combines individual shot records into supershot records which are then simulated during waveform forward modelling and the residuals are back-propagated for gradient calculation. This implementation ideally accelerates the inversion process with a factor of the number of shots assembled in the supershot record.

A fundamental drawback in estimating reflectivity or velocity directly from blended data is the crosstalk noise that arises from the interference of overlapping wavefields from multiple sources. One way to minimize this issue is to apply encoding to the sources to be blended both for LSM and FWI, see, for example, the studies by Krebs *et al.* (2009), Boonyasiriwat and Schuster (2010), Ben-Hadj-Ali, Operto and Virieux (2011), Jeong *et al.* (2013) and Huang and Schuster (2012). In these studies, different encoding schemes were implemented, such as the use of random time shifts, frequency scheduling, amplitude encoding, selection of source locations used for the inversion or combinations thereof. However, the inadequacy in spatial sampling still deteriorates inversion results (Ben-Hadj-Ali *et al.* 2011; Aldawood, Hoteit and Alkhalifah 2014). Boonyasiriwat and Schuster (2010) applied a random distribution of sources to minimize the crosstalk noise in FWI. Wang *et al.* (2014) confirmed the improvement of FWI results with irregularly decimated data compared to regularly decimated data. These studies infer that designing a survey incorporating irregularity potentially leads to effective reflectivity and/or velocity estimation when one aims to directly process blended and irregularly sampled data.

This paper, hence, introduces a survey-design workflow that iteratively optimizes the survey parameters related to both blending and spatial sampling of detectors and sources, leading to satisfactory reflectivity and velocity estimation via joint migration inversion (JMI). JMI is a method that iteratively estimates a high-resolution subsurface reflectivity model as well as a migration velocity model by estimating two independent operators responsible for reflection and propagation (Staal and Verschuur 2013; Berkhouit 2014b). We extend the standard JMI scheme to handle blended and irregularly sampled data, which is then incorporated into the proposed survey-design scheme. The workflow utilizes errors in reflectivity and velocity estimates from the JMI process for given survey design. They are assigned to its objective function and are subsequently input into a survey-parameter update system based on the integration of a genetic algorithm (GA) and a convolutional neural network (CNN). Stochastic operators in the GA that imitate the theory of natural evolution allow for simultaneous update of the blending and sampling operators towards optimum JMI results. The implementation of the CNN aids the population management in our GA by selecting optimum designs while discarding suboptimal ones among newly generated solutions from genetic operators. Only classified designs in the CNN are fed into the subsequent step that involves the evaluation of the objective function through JMI. Since JMI is computationally expensive, this use of an

antecedent classifier effectively prevents wasteful computation incurred to suboptimal solutions, which consequently enhances the performance of the overall survey-design scheme. We also incorporate the dispersed source array (DSA) concept (Berkhout 2012a) in the proposed survey-design scheme. This technique utilizes a plurality of source types, each having a dedicated narrow bandwidth. This permits each narrowband source to be independently deployed to satisfy its own spatial-sampling criteria determined by its frequency range. Caporal *et al.* (2018) extensively discussed its advantages in terms of geophysical, operational and environmental perspectives.

In our acquisition design, the aim is to find the optimum blending and spatial-sampling schemes, contributing to efficiency, economics and HSE, that satisfy the geophysical objectives without relying on an expensive survey acquired in an unblended and regularly well-sampled fashion. Our main objective in this paper is twofold: (1) to explore the effect of the choice of survey parameters on the performance of JMI and (2) to illustrate the proposed survey-design workflow. Numerical examples provide the results of our approach that aims to enhance the JMI results by simultaneously updating different survey parameters involved in DSA acquisition.

SURVEY-DESIGN FRAMEWORK

Using the matrix notation proposed by Berkhout (1982), the monochromatic expression of seismic data is given by

$$\mathbf{P}(z_d; z_s) = \mathbf{D}(z_d)\mathbf{X}(z_d, z_s)\mathbf{S}(z_s). \quad (1)$$

Each column and row of the data matrix, $\mathbf{P}(z_d; z_s)$, correspond to a common shot record and a common detector gather, respectively, each being measured by detectors at depth z_d and sources at depth z_s . The entire seismic data volume can be stored by the collection of $\mathbf{P}(z_d; z_s)$ for each frequency component. Matrices $\mathbf{D}(z_d)$ and $\mathbf{S}(z_s)$ are the detector and source matrix, respectively. The columns in $\mathbf{D}(z_d)$ represent the spatial detector locations whereas the rows in $\mathbf{S}(z_s)$ are the spatial source locations. Matrix $\mathbf{X}(z_d, z_s)$ is the Earth transfer operator containing the subsurface response including propagation effects and (multiple) reflections. The elements of each matrix contain the amplitude and phase information for one frequency component. In the following, we assume the detectors and sources to be located at the same depth, that is, $z_s = z_d = z_0$. According to Berkhout (2014a), $\mathbf{X}(z_0, z_0)$ can be approximated by

$$\mathbf{X}(z_0, z_0) = \sum_{m=1}^{M_d} \underline{\mathbf{W}}^-(z_0; z_m) \mathbf{R}^\cup(z_m) \underline{\mathbf{W}}^+(z_m; z_0), \quad (2)$$

with

$$\begin{aligned} \underline{\mathbf{W}}^-(z_0; z_m) &= \mathbf{W}^-(z_0; z_1) \prod_{n=1}^{m-1} [\mathbf{I} + \delta\mathbf{T}^-(z_n)] \mathbf{W}^-(z_n; z_{n+1}) \\ \underline{\mathbf{W}}^+(z_m; z_0) &= \mathbf{W}^+(z_m; z_{m-1}) \prod_{n=m-1}^1 [\mathbf{I} + \delta\mathbf{T}^+(z_n)] \mathbf{W}^+(z_n; z_{n-1}), \end{aligned} \quad (3)$$

where $\mathbf{W}^-(z_{m-1}; z_m)$ and $\mathbf{W}^+(z_m; z_{m-1})$ account for the upward and downward propagation, respectively, between depth levels z_{m-1} and z_m . They are transposed versions of each other. Parameter M_d indicates the maximum number of depth levels. Matrix $\mathbf{R}^\cup(z_m)$ is the down-up reflectivity operator responsible for scattering at depth, z_m . The terms $[\mathbf{I} + \delta\mathbf{T}^\mp(z_m)]$ are the upward and downward transmission operators at depth, z_m , respectively. Hence, matrices $\underline{\mathbf{W}}^\mp$ include both propagation and transmission effects. According to Berkhout (2014a), the assumption of angle- and frequency-independent reflectivity along with negligible mode conversion allows for the following approximations to be used:

$$\begin{aligned} \mathbf{R}^\cup(z_m) &\approx -\mathbf{R}^\cap(z_m) \\ \delta\mathbf{T}(z_m)^\mp &\approx \mp\mathbf{R}(z_m), \end{aligned} \quad (4)$$

where $\mathbf{R}^\cap(z_m)$ is the up-down reflectivity operator. For notational simplicity, the depth index and the superscripts indicating the directions of wave propagation and reflection are hereinafter omitted.

As for the design of \mathbf{D} and \mathbf{S} , our primary focus is on the effect of spatial sampling. We, therefore, assume the application of a point detector and a point source, each having a delta-functioned response. This simply makes all the non-zero elements of \mathbf{D} and \mathbf{S} one. In addition, the distributions of zero and non-zero elements in \mathbf{D} and \mathbf{S} depict the spatial locations of the detectors and sources, respectively. In an ideal acquisition geometry, that is, carpet detectors and sources, \mathbf{D} and \mathbf{S} become identity matrices. In this case, \mathbf{P} equals \mathbf{X} , and, consequently, optimum retrievals of \mathbf{R} and \mathbf{W} from \mathbf{P} are expectantly attainable. However, in practice, operational and economic constraints lead to imperfections in the spatial sampling, which means that \mathbf{D} and \mathbf{S} are no longer identity matrices. This eventually results in the erroneous estimation of \mathbf{R} and \mathbf{W} . Therefore, we need to properly design survey parameters to obtain reliable estimates of \mathbf{R} and \mathbf{W} , which is targeted in our acquisition design strategy.

Berkhout (2008) proposed the theoretical framework of source blending by introducing a blending matrix, Γ . Based

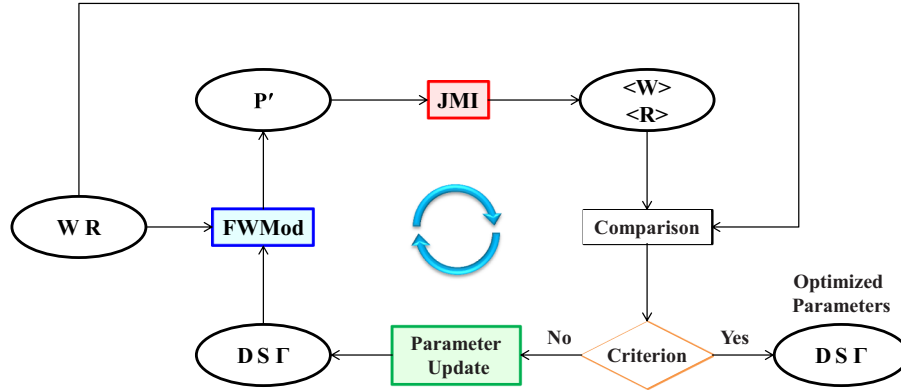


Figure 1 The survey-design workflow. The iterative scheme aims to output the blending and sampling operators that can provide the optimum deblending and reconstruction quality. The workflow starts with \mathbf{R} and \mathbf{W} . The forward process in the blue-filled step derives \mathbf{P}' using FWMod. The inverse process in the red box then estimates $\langle \mathbf{R} \rangle$ and $\langle \mathbf{W} \rangle$ from \mathbf{P}' using JMI. The procedure stops if the misfits between \mathbf{R} and $\langle \mathbf{R} \rangle$ as well as between \mathbf{W} and $\langle \mathbf{W} \rangle$ are sufficiently small or the maximum number of iterations is reached. If not, \mathbf{D} , \mathbf{S} and Γ are updated in the green box and subsequently used in the next iteration.

on equation (1), the forward model of blended data, \mathbf{P}' , can be expressed as

$$\mathbf{P}' = \mathbf{P}\Gamma = \mathbf{D}\mathbf{X}\mathbf{S}\Gamma. \quad (5)$$

Each column and row of Γ represent a blended experiment and a source location, respectively. One column of Γ contains the blending code of the sources involved in one blended experiment, and element i, j for frequency ω is given by

$$\gamma_{i,j} = a_{i,j}(\omega) \exp[-j\phi_{i,j}(\omega)], \quad (6)$$

where $a_{i,j}$ and $\phi_{i,j}$ are the amplitude term and phase term corresponding to the i th source of the j th blended experiment. Any type of source signature can be encoded in equation (6). In the case of simple time delays $\tau_{i,j}$, the corresponding phase shift can be expressed as $\phi_{i,j} = \omega\tau_{i,j}$. Our forward-modelling scheme is also applicable to the dispersed source array (DSA) concept, where a plurality of narrowband sources with different centre frequencies accounts for the total blended wavefield. Being formulated in the frequency domain, equations (2) and (5) imply that starting with \mathbf{R} and \mathbf{W} , various blending and spatial-sampling schemes can be modelled by designing \mathbf{D} , \mathbf{S} and Γ .

Figure 1 illustrates the proposed survey-design workflow to find optimum \mathbf{D} , \mathbf{S} and Γ , while we assume \mathbf{R} and \mathbf{W} , which contain all relevant information of the subsurface, to be available as prior information. Such information is available, for example, at the development or production stage, or when data from a previous acquisition and appropriate well logs are available. It means that the approach is subsurface

dependent. The blue-filled step in Fig. 1 corresponds to the forward process to generate blended data \mathbf{P}' from \mathbf{R} and \mathbf{W} along with the current survey parameters, \mathbf{D} , \mathbf{S} and Γ . This forward-modelling engine is called full-wavefield modelling (FWMod) (Berkhout 2014a). In FWMod, scattering is assumed to take place at each depth level via the reflection operator \mathbf{R} , whereas in between each level the wavefields propagate via the propagation operator \mathbf{W} .

The red-filled step is the inverse process, that is, joint migration inversion (JMI), to obtain $\langle \mathbf{R} \rangle$ and $\langle \mathbf{W} \rangle$ from \mathbf{P}' , where the angled brackets $\langle \rangle$ and $\langle \rangle$ indicate estimation. JMI iteratively derives both reflectivity and velocity models using FWMod as its forward engine (Berkhout 2014b). Therefore, in this study, the same FWMod engine is used for both blue-filled and red-filled steps in Fig. 1. It should be noted that proving the viability of the JMI algorithm is not the objective of this study. Our main focus is on exploring and understanding the effect of acquisition design on the performance of JMI. Hence, we utilize this idealized condition in which the choice of survey parameters can be a major factor determining the quality of estimated subsurface properties. The implementation of FWMod and JMI in our study is based on Staal and Verschuur (2013) and Staal (2015), which updates and estimates one single reflectivity parameter per grid point in \mathbf{R} and one single slowness parameter per grid point in \mathbf{W} . We consider the underlying assumptions of acoustic wavefield behaviour along with angle- and frequency-independent reflectivity made in this implementation to be valid in this study. Staal (2015) suggested the JMI framework can be extended to blended acquisition. In this study, we formulate the objective function in JMI as

$$J_{\text{JMI}} = \sum_{\omega} \|\Delta \mathbf{P}'\|_2^2 = \sum_{\omega} \|\mathbf{P}' - \langle \mathbf{P}' \rangle\|_2^2. \quad (7)$$

Here, we feed \mathbf{P}' directly into the JMI process. We perform JMI on the spatially regular and well-sampled detector grid determined by \mathbf{X} . Each blended and irregularly sampled shot record can be stored on this grid, yet the information is only present at active detector locations defined by \mathbf{D} . Note that this choice means that our solution is spatially irregular in the sense that not all grid points contain detectors. However, all detectors coincide with grid points. By taking the grid size in \mathbf{X} sufficiently small, the consequences of this choice can be largely mitigated. Together with the current reflectivity and slowness estimates, FWMod simulates each blended shot record based on the given \mathbf{S} and Γ at the regular, well-sampled detector grid. By selecting traces from the modelled shot records according to \mathbf{D} , $\langle \mathbf{P}' \rangle$ can be obtained. This subsequently allows us to compute the residual between \mathbf{P}' and $\langle \mathbf{P}' \rangle$ as described in equation (7). The JMI procedure iteratively minimizes the data misfit via a gradient descent scheme. At each iteration, reflectivity and propagation operators are updated by cross-correlating wavefields, that is, the back-propagated residual wavefield and the forward-modelled wavefield, in the opposite directions and in the same direction, respectively (Verschuur, Staal and Berkhouit 2016). In our case, empty traces corresponding to detector locations that are not present in \mathbf{D} are also used in the back-propagation in JMI along with the forward-propagation of blended source-wavefield determined by \mathbf{S} and Γ .

The overall acquisition-design scheme attempts to minimize the residue between \mathbf{R} and $\langle \mathbf{R} \rangle$ as well as \mathbf{W} and $\langle \mathbf{W} \rangle$:

$$\mathbf{j} = \begin{bmatrix} J_{\mathbf{R}} \\ J_{\mathbf{W}} \end{bmatrix} = \begin{bmatrix} \sum_{\omega} \|\Delta \hat{\mathbf{R}}\|_2^2 \\ \sum_{\omega} \|\Delta \hat{\mathbf{W}}\|_2^2 \end{bmatrix} = \begin{bmatrix} \sum_{\omega} \|\hat{\mathbf{R}} - \langle \hat{\mathbf{R}} \rangle\|_2^2 \\ \sum_{\omega} \|\hat{\mathbf{W}} - \langle \hat{\mathbf{W}} \rangle\|_2^2 \end{bmatrix}, \quad (8)$$

where \mathbf{j} is the objective-function vector containing errors in $\hat{\mathbf{R}}$ and $\hat{\mathbf{W}}$. The term $\hat{\mathbf{W}}$ represents a scalar velocity field in the space-depth domain derived from a propagation operator, and $\hat{\mathbf{R}}$ is a scalar reflectivity field converted to the space-time domain such that any undesired effects from errors in $\langle \mathbf{W} \rangle$ on $J_{\mathbf{R}}$ can be avoided.

The green-filled step in Fig. 1 updates the blending and sampling operators that are subsequently carried into the next iteration. The procedure stops once the objective function is sufficiently small, or the maximum number of iterations is exceeded. Therefore, our approach iteratively computes the acquisition design parameters \mathbf{D} , \mathbf{S} and Γ that minimize the objective-function vector, meaning that optimum reflectivity

and velocity estimates can be obtained. With the proposed survey-design scheme, the variation of the JMI results is assumed to be attributable fundamentally to the choice of survey parameters.

SURVEY-PARAMETER UPDATE

Designing survey parameters in an irregular fashion requires extensive efforts as it inherently makes the parameter selection problem huge, unlike a conventional survey that generally employs regularly positioned detectors and sources along with a spatially uniform source signature. To manage the large problem space, we presented a survey-design scheme utilizing a genetic algorithm (GA) (Nakayama *et al.* 2019). Together with a constraint on the parameter space, a so-called repeated encoding sequence, this scheme aimed at deriving survey parameters allowing for optimum deblending and data reconstruction quality. In this study, we extend the approach to obtain blending and sampling operators that optimize joint migration inversion (JMI) results. Owing to its ability to manage optimization problems with non-convexity, non-differentiability, the existence of many local minima and large problem space, GAs have successfully been implemented in various application domains (Monteagudo and Santos 2015; Perez-Liebana *et al.* 2015; Bak, Rask and Risi 2016; Davies *et al.* 2016; Scirea *et al.* 2016). However, a standard GA inherently and inevitably needs to evaluate all the solutions to obtain their objective-function values. The evaluation of suboptimal solutions, that is, those that do not contribute to the next generation, potentially makes GAs computationally expensive.

A number of studies have shown the integration of metaheuristics and neural networks. In spite of recent successes in neural networks (Hinton *et al.* 2012; Leung *et al.* 2014; Schmidhuber 2015), advanced architectures have become deeper and more complex. Designing network architectures is still a challenging task due to the existence of numerous parameters. In this respect, a typical approach is to explore and design efficient network architectures using GAs (Ahmadizar *et al.* 2015; Xie and Yuille 2017). The other way around, Kramer (2016) outlined ways to support and accelerate evolutionary algorithms with the learning process. In various engineering domains, determining the objective function of a single candidate takes a long time to execute due to the use of finite element analysis and computational fluid dynamics simulations. To deal with such computationally expensive optimization problems, a viable option is to use a model that can approximate the outcomes from the fitness evaluations at relatively low computation cost and to limit the

number of real evaluations (Tenne and Goh 2010). Kramer (2017) illustrated a workflow to embed a supervised learning model into a GA where the real fitness evaluations are performed only to the solutions that fulfil criterion in the predictive model. This way of integration is easily recognizable in different engineering applications (Marcelin 2004; Sreekanth and Datta 2011; Ibaraki, Tomita and Sugimoto 2015; Sato and Fujita 2016; Sakaguchi *et al.* 2018). Exploiting a strategy from these engineering problems is certainly valuable to our survey-design problem because a large computation effort to calculate the objective function inherently restricts the number of fitness evaluations in our case. Hence, this study implements a convolutional neural network (CNN), which aims primarily to aid the population management in our GA and subsequently reach a satisfactory solution with affordable turn-around time.

As described in Algorithm 1, the proposed approach updates a set of parameter vectors representing \mathbf{D} , \mathbf{S} and Γ . The vectors for the p th individual solution in the g th generation are given by

$$\mathbf{c}_{g,p} = [\mathbf{d}_{g,p}, \mathbf{s}_{g,p}, \boldsymbol{\gamma}_{g,p}]^T, \quad (9)$$

with

$$\mathbf{d}_{g,p} \in \{0, 1\}^{l_d}, \quad \mathbf{s}_{g,p} \in \{0, 1\}^{l_s}, \quad \boldsymbol{\gamma}_{g,p} \in \{0, 1\}^{l_s \times l_g}. \quad (10)$$

Here, $\mathbf{d}_{g,p}$ and $\mathbf{s}_{g,p}$ are binary vectors indicating detector sampling and source sampling operators for the p th individual solution in the g th generation, each of which can be expressed as

$$\mathbf{d}_{g,p} = \text{diag}[\mathbf{D}_{g,p}^T \mathbf{D}_{g,p}] \quad (11)$$

and

$$\mathbf{s}_{g,p} = \text{diag}[\mathbf{S}_{g,p} \mathbf{S}_{g,p}^T]. \quad (12)$$

Their dimensions, l_d and l_s , are equal to the number of detectors and sources in \mathbf{X} . As described previously, the assumption of point detectors and point sources allows distributions of zero and non-zero elements to explicitly indicate their spatial coordinates. With the use of delta-functioned detector and source responses, \mathbf{D} and \mathbf{S} can be simply expressed by binary vectors using equations (11) and (12). Vector $\boldsymbol{\gamma}$ is a binary vector representing the blending operator for the p th individual solution in the g th generation. Parameter l_g equals the required bit length to parametrize a given blending code per single source. In our case, a binary string for a given source with a length of l_g accounts for the amplitude term and the phase term given by equation (6). Since the spectral properties of the dispersed source array (DSA) sources are

predetermined, the information on the source type is included in the frequency-dependent amplitude term.

Let n_g and n_p represent the numbers of generations and populations. The population in the g th generation, \mathbf{c}_g , consists of a set of parameter vectors, having n_p individuals as

$$\mathbf{c}_g = [\mathbf{c}_{g,1}, \mathbf{c}_{g,2}, \dots, \mathbf{c}_{g,n_p}]^T. \quad (13)$$

The first step is to randomly generate the initial population, \mathbf{c}_0 , across a given problem space:

$$\mathbf{c}_0 = \begin{bmatrix} \mathbf{c}_{0,1} \\ \mathbf{c}_{0,2} \\ \vdots \\ \mathbf{c}_{0,n_p} \end{bmatrix} = \begin{bmatrix} [\mathbf{d}_{0,1}, \mathbf{s}_{0,1}, \boldsymbol{\gamma}_{0,1}]^T \\ [\mathbf{d}_{0,2}, \mathbf{s}_{0,2}, \boldsymbol{\gamma}_{0,2}]^T \\ \vdots \\ [\mathbf{d}_{0,n_p}, \mathbf{s}_{0,n_p}, \boldsymbol{\gamma}_{0,n_p}]^T \end{bmatrix}. \quad (14)$$

Each parameter vector in the initial population, $\mathbf{c}_{0,p}$, then goes to the forward process (going to $\mathbf{P}'_{0,p}$ from \mathbf{R} and \mathbf{W} along with $\mathbf{D}_{0,p}$, $\mathbf{S}_{0,p}$ and $\Gamma_{0,p}$). The inverse process subsequently follows (going to $\langle \mathbf{R}_{0,p} \rangle$ and $\langle \mathbf{W}_{0,p} \rangle$ from $\mathbf{P}'_{0,p}$). Using equation (8), a set of objective functions for the initial population, \mathbf{j}_0 can be obtained by

$$\mathbf{j}_0 = \begin{bmatrix} [J_{R0,1}, J_{W0,1}]^T \\ [J_{R0,2}, J_{W0,2}]^T \\ \vdots \\ s [J_{R0,n_p}, J_{W0,n_p}]^T \end{bmatrix} = \begin{bmatrix} \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{0,1}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{0,1}\|_2^2 \right]^T \\ \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{0,2}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{0,2}\|_2^2 \right]^T \\ \vdots \\ \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{0,n_p}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{0,n_p}\|_2^2 \right]^T \end{bmatrix}. \quad (15)$$

Using \mathbf{c}_0 and \mathbf{j}_0 , we train the CNN. As shown in Fig. 2, the neural network architecture in this study consists of layers that are commonly used in CNN architectures: three pairs of 2D convolutional layers (LeCun *et al.* 1998) and ReLU (rectified linear unit) layers (Hahnloser *et al.* 2000), followed by a fully connected and softmax layer (Bishop 2006). In a convolutional layer, we apply zero padding and a stride of one along the height and width dimensions. The channel size and kernel size of each layer are given in Fig. 2. In this study, we select square shape kernels. Instead of down-sampling feature maps within the CNN via a pooling layer, we first arrange

Algorithm 1. Survey-design workflow**Input:** \mathbf{R} and \mathbf{W} **Output:** $\mathbf{c}_{0:n_g}$ ($= [\mathbf{c}_0, \dots, \mathbf{c}_{n_g}]^T$) and $\mathbf{j}_{0:n_g}$ ($= [\mathbf{j}_0, \dots, \mathbf{j}_{n_g}]^T$)

- 1: Set n_g and n_p ($= n_c + n_m$)
- 2: Generate initial population, \mathbf{c}_0 ($= [\mathbf{c}_{0,1}, \dots, \mathbf{c}_{0,n_p}]^T$)
- 3: **for** $g = 0 \rightarrow n_g$ **do**
- 4: **if** $g = 0$ **then**
- 5: $\mathbf{c}_{0:g-1} = \emptyset$ and $\mathbf{j}_{0:g-1} = \emptyset$
- 6: **else**
- 7: Compute selection probability of each solution in \mathbf{c}_{g-1}
- 8: **for** $p = 1 \rightarrow n_c/2$ **do**
- 9: **while** stopping criterion not met **do**
- 10: Select two parental solutions from $\mathbf{c}_{0:g-1}$ using roulette wheel selection
- 11: Perform crossover operation to generate $\mathbf{c}_{g,2p-1}$ and $\mathbf{c}_{g,2p}$
- 12: Evaluate $\mathbf{c}_{g,2p-1}$ and $\mathbf{c}_{g,2p}$ with the CNN trained from $g - 1^{\text{th}}$ generation
- 13: **end while**
- 14: **end for**
- 15: **for** $p = n_c + 1 \rightarrow n_p$ **do**
- 16: **while** stopping criterion not met **do**
- 17: Select one parental solution from $\mathbf{c}_{0:g-1}$ using roulette wheel selection
- 18: Perform mutation operation to generate $\mathbf{c}_{g,p}$
- 19: Evaluate $\mathbf{c}_{g,p}$ with the CNN trained from $g - 1^{\text{th}}$ generation
- 20: **end while**
- 21: **end for**
- 22: **end if**
- 23: **for** $p = 1 \rightarrow n_p$ **do**
- 24: Generate $\mathbf{D}_{g,p}$, $\mathbf{S}_{g,p}$ and $\mathbf{\Gamma}_{g,p}$ from $\mathbf{c}_{g,p}$ ($= [\mathbf{d}_{g,p}, \mathbf{s}_{g,p}, \gamma_{g,p}]^T$)
- 25: Generate $\mathbf{P}'_{g,p}$ from \mathbf{R} and \mathbf{W} along with $\mathbf{D}_{g,p}$, $\mathbf{S}_{g,p}$ and $\mathbf{\Gamma}_{g,p}$ via FWMod
- 26: Estimate $\langle \mathbf{R} \rangle_{g,p}$ and $\langle \mathbf{W} \rangle_{g,p}$ from $\mathbf{P}'_{g,p}$ via JMI
- 27: Compute reflectivity and velocity errors using equation 8
- 28: **end for**
- 29: Update $\mathbf{c}_{0:g}$ ($= [\mathbf{c}_{0:g-1}, \mathbf{c}_g]^T$) and $\mathbf{j}_{0:g}$ ($= [\mathbf{j}_{0:g-1}, \mathbf{j}_g]^T$)
- 30: Apply fast non-dominated sort to $\mathbf{c}_{0:g}$ based on $\mathbf{j}_{0:g}$
- 31: Compute crowding distances of each non-domination rank
- 32: Train the CNN using $\mathbf{c}_{0:g}$ and $\mathbf{j}_{0:g}$
- 33: Sort $\mathbf{c}_{0:g}$ by the elitism based primarily on non-domination rank and secondary on crowding distance to derive the fittest population, \mathbf{c}_g
- 34: **end for**

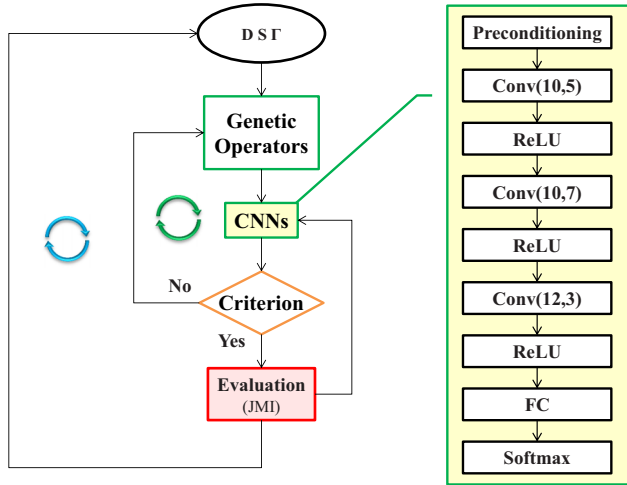


Figure 2 The scheme to update survey parameters. Genetic operators update \mathbf{D} , \mathbf{S} and $\mathbf{\Gamma}$ simultaneously. Newly generated parameters then go to the CNN. Survey parameters are regenerated till the classification result meets the criterion (the expected residue between \mathbf{R} and $\langle \mathbf{R} \rangle$ as well as between \mathbf{W} and $\langle \mathbf{W} \rangle$). Only selected solutions go to the actual evaluation that involves JMI. Designs having smaller objective-function values form a new generation. Additionally, the evaluation results are used to train the CNN applied to the classification step in the next iteration. The diagram at the right depicts the applied network architecture. Conv means a convolutional layer with two numbers indicating the channel size and the kernel size, respectively. ReLU indicates a rectified linear unit layer and FC a fully connected layer.

blending and sampling operators into a 3D matrix, \mathbf{Y} , its size being $c_w \times c_h \times N$, where c_w and c_h are the width and the height of \mathbf{Y} . They equal the numbers of columns and rows in \mathbf{X} , respectively. N equates to the number of DSA source types. For the p th individual in the initial population, the n^{th} feature map of \mathbf{Y} is given by

$$\mathbf{Y}_{0,p}^{(n)} = \lambda_n |\mathbf{L}[\mathbf{D}_{0,p}^T \mathbf{D}_{0,p} (\mathbf{k} \boldsymbol{\tau}_{0,p}^T) \mathbf{S}_{n,0,p} \mathbf{S}_{n,0,p}^T]|, \quad (16)$$

with

$$\mathbf{k} \in \{1\}^{c_h}, \quad \boldsymbol{\tau}_{0,p} = [\tau_{1,0,p}, \dots, \tau_{c_h,0,p}]^T. \quad (17)$$

Matrix \mathbf{S}_n contains the spatial locations of the n th source type of the DSA acquisition, λ_n is a user-defined weight parameter and vector $\boldsymbol{\tau}$ contains information on the activation time applied to each source. Matrix \mathbf{L} is the forward transform operator to the wavenumber domain. Since we assume delta-functioned detector and source responses, \mathbf{Y} becomes frequency independent, leading to a reduction in the data size. Although a deeper and more complex network presumably enables us to directly input \mathbf{D} , \mathbf{S} and $\mathbf{\Gamma}$ into the CNN, for now this mathematical operation helps to achieve reasonable

classification results with an affordable computation time. After the initial iteration, the CNN is embedded into our GA. As mentioned, its task is to exclude individual solutions that are likely to not satisfy the predetermined threshold criterion for proceeding to the next generation.

To minimize the vectorized objective function in equation (8), we utilize non-dominated sorting and crowding distance approaches (Deb *et al.* 2002). Algorithm 2 illustrates the procedure to assess the individual solution. It starts with finding the first-rank solutions. This involves a calculation of two entities: (1) domination count $q_{g,p}$ (the number of solutions dominating a given solution, $\mathbf{c}_{g,p}$) and (2) a set of individuals, $\mathbf{c}_{g,p}^0$, being dominated by $\mathbf{c}_{g,p}$. For all solutions with rank one, their domination counts become zero. For each member dominated by these solutions, its domination count is reduced by one. Any members having a zero domination count obtain rank two. This procedure is repeated until the ranks of all solutions are identified.

We then derive the selection probability of each solution according to its rank as

$$G(\mathbf{c}_{g,p}) = r_{g,p} / \sum_{j=1}^{n_p} r_{g,j}, \quad (18)$$

with

$$r_{g,p} = \exp(-\beta \text{Rank}_{g,p} / \max_{j \in n_p} \text{Rank}_{g,j}), \quad (19)$$

where β is a scalar controlling the diversity of the selection. The new population is formed by elitism replacement based primarily on the domination rank. Additionally, to discriminate solutions with the same rank, we analyse their crowding distances. An infinite crowding distance is given to two solutions within the t th rank having the maximum and the minimum value of a given objective function (in our case either J_R or J_W). Crowding distance values for other solutions are then calculated by the sum of individual distance values corresponding to each objective function as

$$d_{g,j} = \left| \frac{J_R(\mathbf{c}_{g,j+1}) - J_R(\mathbf{c}_{g,j-1})}{\max_{j \in n_t} J_{R,g,j} - \min_{j \in n_t} J_{R,g,j}} \right| + \left| \frac{J_W(\mathbf{c}_{g,j+1}) - J_W(\mathbf{c}_{g,j-1})}{\max_{j \in n_t} J_{W,g,j} - \min_{j \in n_t} J_{W,g,j}} \right|, \quad (20)$$

with

$$j \in [2, \dots, n_t - 1], \quad \mathbf{c}_{g,j} \in \mathbf{f}_t, \quad (21)$$

where \mathbf{f}_t is a set of solutions belonging to the t th rank and n_t is the number of solutions in \mathbf{f}_t . This ensures the diversity of

Algorithm 2. Fast non-dominated sort**Input:** $\mathbf{c}_{0:g} (= [\mathbf{c}_0, \dots, \mathbf{c}_g]^T)$ and $\mathbf{j}_{0:g} (= [\mathbf{j}_0, \dots, \mathbf{j}_g]^T)$ **Output:** $\mathbf{f}_{1:t} (= [\mathbf{f}_1, \dots, \mathbf{f}_t]^T)$

```

1: for each  $\mathbf{c}_{k,h} \in \mathbf{c}_{0:g}$  do
2:    $\mathbf{q}_{k,h}^0 = \emptyset$  and  $q_{k,h} = 0$ 
3:   for each  $\mathbf{c}_{i,j} (\neq \mathbf{c}_{k,h}) \in \mathbf{c}_{0:g}$  do
4:     if  $\mathbf{c}_{k,h} \prec \mathbf{c}_{i,j}$  then
5:        $\mathbf{q}_{k,h}^0 = \mathbf{q}_{k,h}^0 \cup \mathbf{c}_{i,j}$ 
6:     else
7:        $q_{k,h} = q_{k,h} + 1$ 
8:     end if
9:   end for
10:  if  $q_{k,h} = 0$  then
11:     $\text{Rank}_{k,h} = 1$ 
12:     $\mathbf{f}_1 = \mathbf{f}_1 \cup \mathbf{c}_{k,h}$ 
13:  end if
14: end for
15:  $t = 1$ 
16: while  $\mathbf{f}_t \neq \emptyset$  do
17:   $\mathbf{q} = \emptyset$ 
18:  for each  $\mathbf{c}_{k,h} \in \mathbf{f}_t$  do
19:    for each  $\mathbf{c}_{i,j} \in \mathbf{q}_{k,h}^0$  do
20:       $q_{i,j} = q_{i,j} - 1$ 
21:      if  $q_{i,j} = 0$  then
22:         $\text{Rank}_{i,j} = t + 1$ 
23:         $\mathbf{q} = \mathbf{q} \cup \mathbf{c}_{i,j}$ 
24:      end if
25:    end for
26:  end for
27:   $t = t + 1$ 
28:   $\mathbf{f}_t = \mathbf{q}$ 
29: end while

```

the new generation by assigning a higher priority to a more isolated solution in the objective-function space.

According to the selection probabilities of the individuals, parental solutions that feed into crossover and mutation operations are selected. Let n_c and n_m be the number of parental solutions to apply crossover and mutation, respectively. The sum of n_c and n_m is then equal to n_p . Crossover mixes the portions of two parental solutions to generate new solutions. Mutation then adds a random refinement to a single parental solution. Our parameter sequence consists of different parameter vectors related to the blending and sampling operators, each

having a different length and constraint. Thus, we apply crossover and mutation to each binary vector separately. Newly generated solutions are evaluated using the CNN trained in the previous iteration to classify solutions according to their likelihood to pass the predetermined threshold based on the non-domination rank. Genetic operators repeatedly generate new \mathbf{d} , \mathbf{s} and $\boldsymbol{\gamma}$ until the CNN classifies all candidates in a given generation as likely to pass the threshold.

All the new candidates \mathbf{c}_g of generation g obtained in this way go to the JMI process to derive a set of objective-function vectors:

$$\mathbf{j}_g = \begin{bmatrix} [J_{R,g,1}, J_{W,g,1}]^T \\ [J_{R,g,2}, J_{W,g,2}]^T \\ \vdots \\ [J_{R,g,n_p}, J_{W,g,n_p}]^T \end{bmatrix} = \begin{bmatrix} \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{g,1}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{g,1}\|_2^2 \right]^T \\ \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{g,2}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{g,2}\|_2^2 \right]^T \\ \vdots \\ \left[\sum_{\omega} \|\Delta \hat{\mathbf{R}}_{g,n_p}\|_2^2, \sum_{\omega} \|\Delta \hat{\mathbf{W}}_{g,n_p}\|_2^2 \right]^T \end{bmatrix}. \quad (22)$$

We combine \mathbf{c}_g with the solutions obtained from all previous generations, $\mathbf{c}_{0:g-1}$, and \mathbf{j}_g with the objective functions from all previous generations, $\mathbf{j}_{0:g-1}$, as

$$\mathbf{c}_{0:g} = [\mathbf{c}_{0:g-1}, \mathbf{c}_g]^T = [\mathbf{c}_0, \dots, \mathbf{c}_g]^T \quad (23)$$

and

$$\mathbf{j}_{0:g} = [\mathbf{j}_{0:g-1}, \mathbf{j}_g]^T = [\mathbf{j}_0, \dots, \mathbf{j}_g]^T. \quad (24)$$

Since the new individuals do not necessarily attain better results, the n_p best solutions in $\mathbf{c}_{0:g}$ are selected on the basis of elitism. They are subsequently replaced to \mathbf{c}_g which is then used for the next generation. In this way, the best parental solutions are carried into the new generation.

Finally, we train the CNN using $\mathbf{c}_{0:g}$ and $\mathbf{j}_{0:g}$. We apply a fivefold cross-validation to evaluate the predictive classifier. The procedure randomly splits all the samples into five subsets with equal size. One of the five subsets is used as a testing set, whereas the remaining four subsets are assembled to form a training set. This process is repeated five times by alternately using every subset for testing and the remaining subsets for training. This interchange enables all the samples to be used for training as well as for testing. The mean of prediction

errors along with the variance among the five trials indicates the predictive performances of the classifiers. The best of the five models is applied to the next iteration. At each generation, the threshold criterion applied in the classification is updated, starting from a mild value and moving to a harsh one. To improve the stability of the CNN, we modify the number of epochs and the strength of the L2-regularization term applied to every weight in the network through the course of the iterations. The former parameter is increased while the latter is decreased with the number of iterations. Controlling them is considered as an effective and simple means to prevent overfitting (Murphy 2012). It is worth noting that the update of survey parameters is based primarily on the actual performance of JMI, while the CNN classifier has a supporting role in the proposed scheme.

The iterative procedure stops when the objective function is sufficiently small, or when the number of generations reaches the pre-defined number of generations n_g . The potential risk of our approach is that we may end up in some local minimum. Fortunately, since we are dealing with survey design, this is still acceptable as long as the resulting design is efficient and leads to the required image quality. It should be noted that various constraints on the blending and sampling operators can be imposed within the genetic operators in order to prevent the generation of any undesired solutions, such as operationally infeasible designs. It is also noteworthy that the proposed optimization scheme is not a mandated choice while our survey-design workflow can accommodate different methods. Both metaheuristics and neural networks can be

flexibly modified and designed, enabling us to adapt them for a problem-specific task. Therefore, a user can freely select a preferred framework for a given survey-design problem.

NUMERICAL EXAMPLES

We numerically simulated four acquisition scenarios: one representing standard blended acquisition and the other three incorporating the dispersed source array (DSA) concept. Table 1 summarizes the spatial-sampling schemes used in this study. In the standard scenario, detectors and sources are regularly deployed, while in the other three the geometries are irregular. Additionally, the DSA scenarios use fewer sources than the standard one. Figure 3 and Table 2 compare the source properties of the standard and DSA scenarios. The standard scenario uses a spatially uniform source signature with a wide frequency range. On the contrary, the DSA scenarios employ three source types, each having a dedicated narrow frequency bandwidth and a spatial-sampling scheme according to its frequency range. This illustrates that the DSA scenarios emit significantly less energy in both space and frequency. Note

Table 1 Detector and source sampling in standard and DSA scenarios

	Standard	DSA
Detector interval	40 m at regular	Irregular
The number of detectors	50	50
Source interval	20 m at regular	Irregular
The number of sources	100	70 (see Table 2)

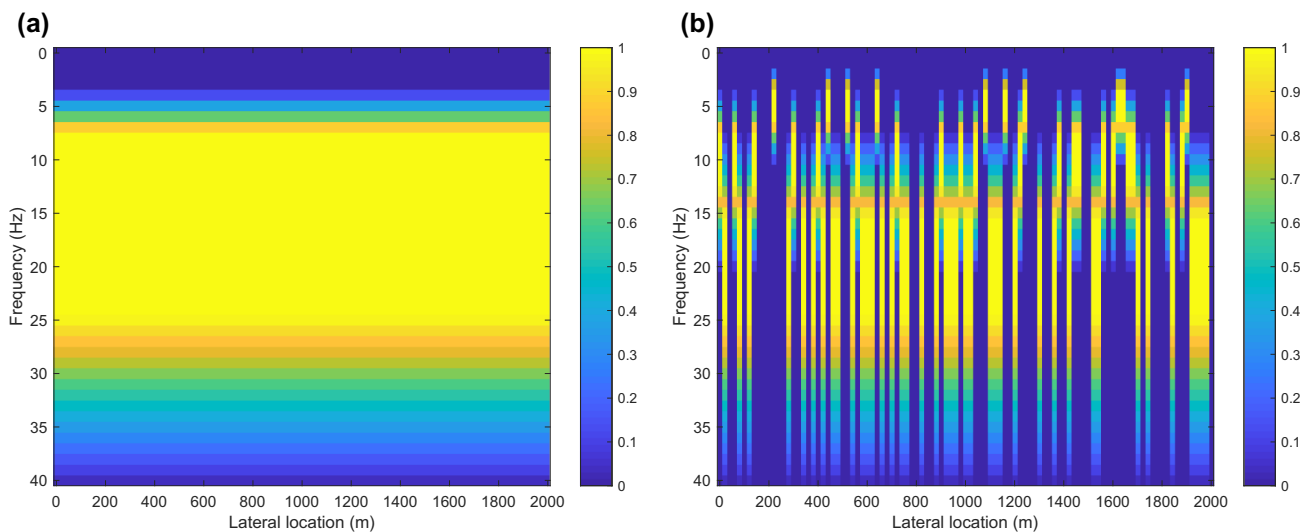


Figure 3 Source responses in the frequency-space domain. (a) Standard scenario using spatially uniform source signatures with a wide bandwidth. (b) DSA scenario using three types of source units, each having a specific sampling requirement and a dedicated narrow bandwidth.

Table 2 Source properties in standard and DSA scenarios

	Frequency range (Hz)	The number of sources
Standard source	4-8-24-40	100
DSA source type 1	2-4-6-10	10
DSA source type 2	4-8-12-20	20
DSA source type 3	8-16-24-40	40

Four corner frequencies, (1) low cut, (2) low pass, (3) high pass and (4) high cut, describe the frequency range of each source type. Compared to the standard scenario, the DSA scenario employs fewer sources, each having a limited bandwidth.

that we applied dedicated low-frequency sources, called DSA source type 1, whose frequency range is not well covered by the source used in the standard scenario (Fig. 3 and Table 2). The DSA concept allows a lower frequency source to be more coarsely sampled and a higher frequency source to be more densely sampled. This effectively prevents both oversampling of the lower frequencies and undersampling of the higher frequencies (Berkhout 2012a; Caporal *et al.* 2018).

Figure 4 shows a blended shot record from the standard scenario and one from a DSA scenario that exemplify our blending and spatial-sampling schemes. Unlike the standard scenario shown in Fig. 4(a), sources with different frequency responses are blended in the DSA scenario shown in Fig. 4(b). Our synthetic data contain both primaries and internal mul-

tiples. These data are directly fed into the joint migration inversion (JMI) process to obtain reflectivity and velocity estimates, that is, without deblending and reconstruction of missing traces.

For comparison purposes, we provide three DSA scenarios, two of which are obtained from blending and sampling operators created by 800 realizations of uniformly distributed, random variables. We derived the probability density function (PDF) by kernel density estimation using the 800 realizations. We show one result, called ‘P50’, having the mode value in the estimated PDF, which we assume to be representative for the situation where we rely on a single random realization to embed irregularity into blending and sampling operators. In addition, we provide the best result, ‘P1’, which is assumed to represent the outcome of a Monte Carlo approach. The previously mentioned three scenarios (standard, P50 and P1) are compared with the result from our optimized design.

Figure 5 shows the acquisition configurations applied to the standard and three DSA scenarios, respectively. In our examples, the detectors and sources are placed at the surface ($z_d = z_s = 0$ m). The detectors are stationary, making the maximum offset in the data 2000 m. All scenarios use the same number of detectors yet they are deployed differently. In each record of this study, two active sources are blended with a 1000 m distance separation having different activation

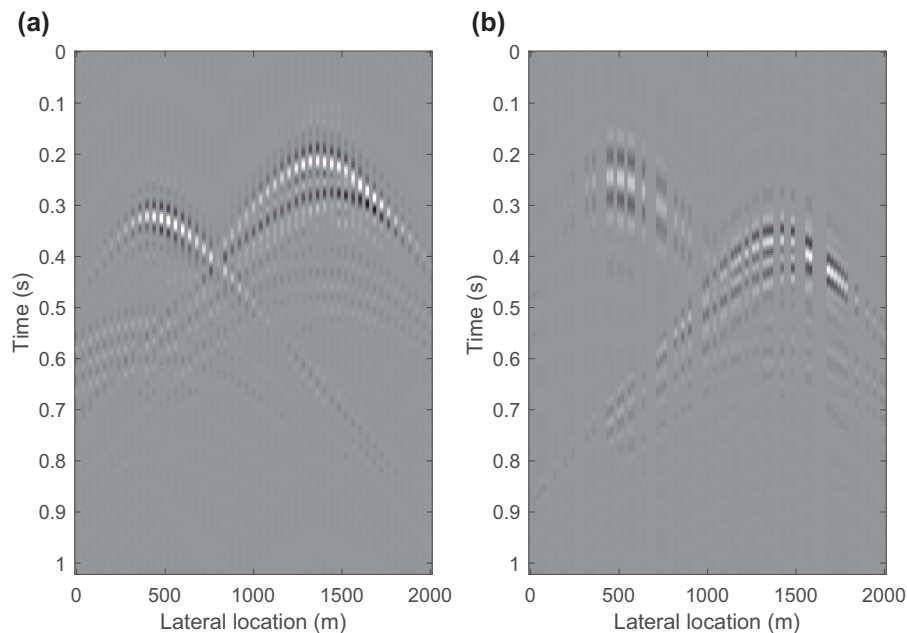


Figure 4 Blended shot record. (a) A blended shot record from the standard case where two sources employ the same signature with a regular detector interval. (b) A blended shot record from the DSA case where two sources employ different bandwidths and activation times with an irregular detector interval.

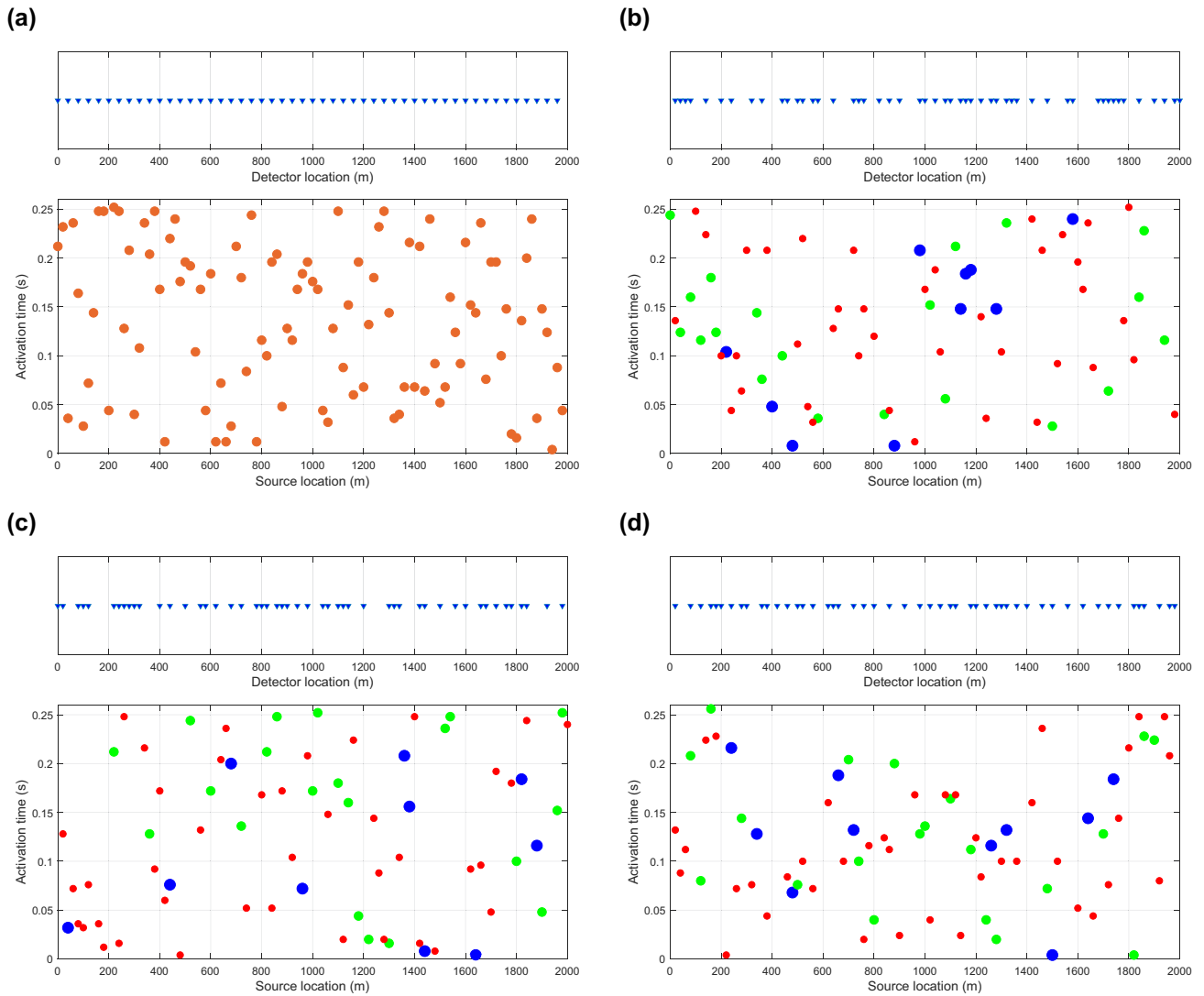


Figure 5 Acquisition scenarios used in this study: (a) standard, (b) P50, (c) P1 and (d) optimized designs. Each figure contains two plots: spatial sampling of detectors (top) and that of sources as well as activation times (bottom). In the standard scenario, all sources employ the same signature with irregular activation times and with a regular detector interval. In three DSA scenarios, blue, green and red circle markers with different marker sizes correspond to DSA source types 1, 2 and 3, respectively. Activation times of these DSA sources are irregular with an irregular detector interval.

times ranging from 0 to 256 ms. In the DSA scenarios, three types of source units are irregularly distributed according to the requirements defined in Table 2.

Figure 6(a,b) shows the true subsurface properties used in this study in terms of reflectivity and velocity, respectively. The model contains a lens-shaped high-velocity body above three weak horizontal reflectors. Figure 6(c,d) shows the initial reflectivity model, which involves no contrast, and the initial velocity model, which contains a gradient only. The estimation by JMI starts with these models having (almost) no

indication of true geological features. Figure 7(a,b) shows the JMI results from the standard acquisition design. Although some oblique lineaments are still recognizable below the lens body, both reflectivity and velocity models are estimated reasonably well. Figure 7(c,d) shows the JMI results from P50. This scenario apparently accentuates linear artefacts, leading to some jitter on the three horizontal reflectors. The lateral velocity variation, particularly beneath the high-velocity lens, adversely affects the kinematics of wave propagation. This explains the undesired structural undulations on the three

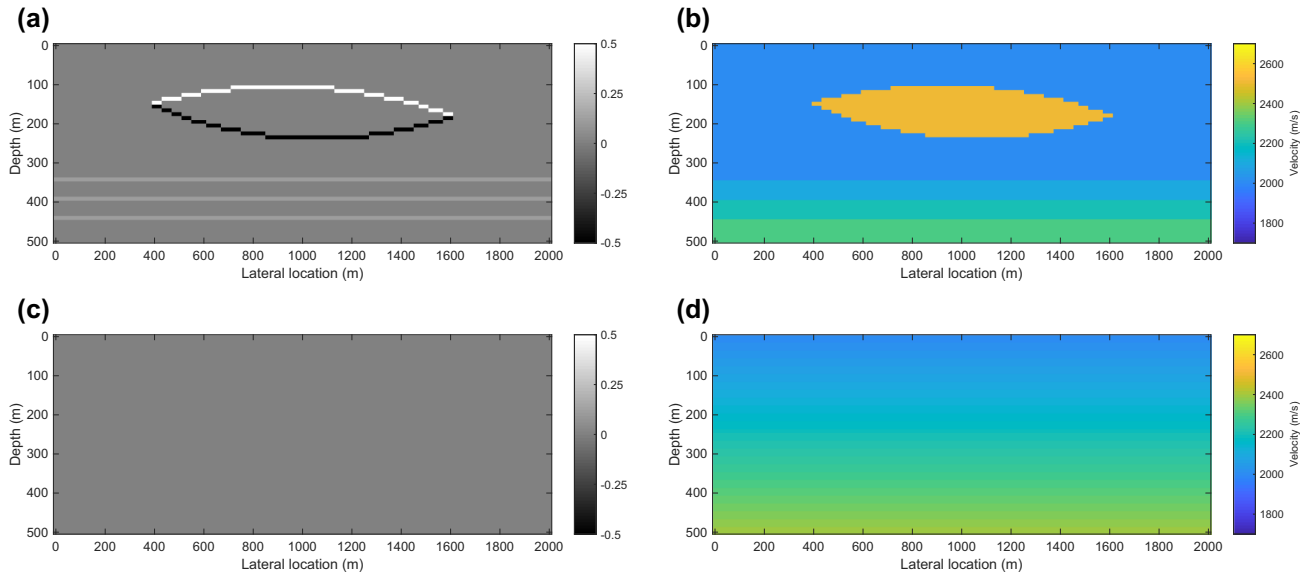


Figure 6 Subsurface model used in this study: (a) true reflectivity model, (b) true velocity model, (c) initial reflectivity model and (d) initial velocity model.

reflectors. Compared to P50, some improvements are observable in P1 (Fig. 7e,f). However, it still is hard to find a justifiable rationale for the applied DSA scheme compared to the standard one in terms of the JMI quality. The optimized design, however, attains notable enhancement in the JMI results (Fig. 7g,h). The lens-shaped body can be easily delineated in both reflectivity and velocity estimates. Reduction of artefacts improves the coherency of the reflectors. The optimized design also achieves a robust estimation of the velocity model, which enables all the reflectors to be recovered close to their actual locations. Even edge effects are well reduced in the optimized design compared to the other three scenarios.

In addition to the observations made from reflectivity and velocity estimates, Figure 8 quantitatively differentiates the overall data quality of each scenario. Figure 8(a) shows a cross-plot of J_R against J_W from 800 random realizations. The colours of the circle markers indicate the number of realizations. The blue, green, red and cyan squares represent the standard, P50, P1 and optimized design, respectively. The three dashed lines in Fig. 8(a) are constant probability density contours each of which represents the boundary of the area that contains a certain percentage (25%, 50% and 75% from the inner to the outer contour) of the estimated PDF. It shows a close to unimodal distribution indicating an increase in data points towards a single peak in the $J_R - J_W$ space. This implies that P50 obtained from the mode value in the PDF is expected to reasonably represent the anticipated data quality in the case where we use a single random realization to design blending

and sampling operators. If it is assumed that the PDF based on the outcome of our Monte Carlo optimization procedure is correct, we observe that the cumulative probability of the objective-function values from our optimized design turns out to be smaller than 10^{-13} . Statistically, this suggests that an enormous number of random realizations are needed to reach a result that is equivalent to our optimized design. On the other hand, our workflow is capable of obtaining it with 800 realizations. Figure 8(b) shows a cross-plot of J_R against J_W from our approach. This clearly demonstrates that the proposed workflow effectively and efficiently minimizes both J_R and J_W through the course of iterations. The optimized DSA acquisition scenario consequently leads to proper reflectivity and velocity estimates even compared to the standard design.

As described previously, we performed a fivefold cross-validation at every 50 realizations to evaluate classification accuracies while training the convolutional neural network (CNN) (Fig. 9). A vertical error bar indicates the minimum and the maximum accuracies obtained from each validation. Blue and red circle markers represent the mean value from five cross-validations for training and testing, respectively. Although we adjusted the regularization strength applied on each weight in the network and the number of epochs to obtain a reliable model, due to the insufficient number of samples the testing results still exhibit some indication of overfitting at the early stage of the iterative procedure. Since we altered the threshold criterion and some parameters within the CNN as mentioned previously, the validation results among

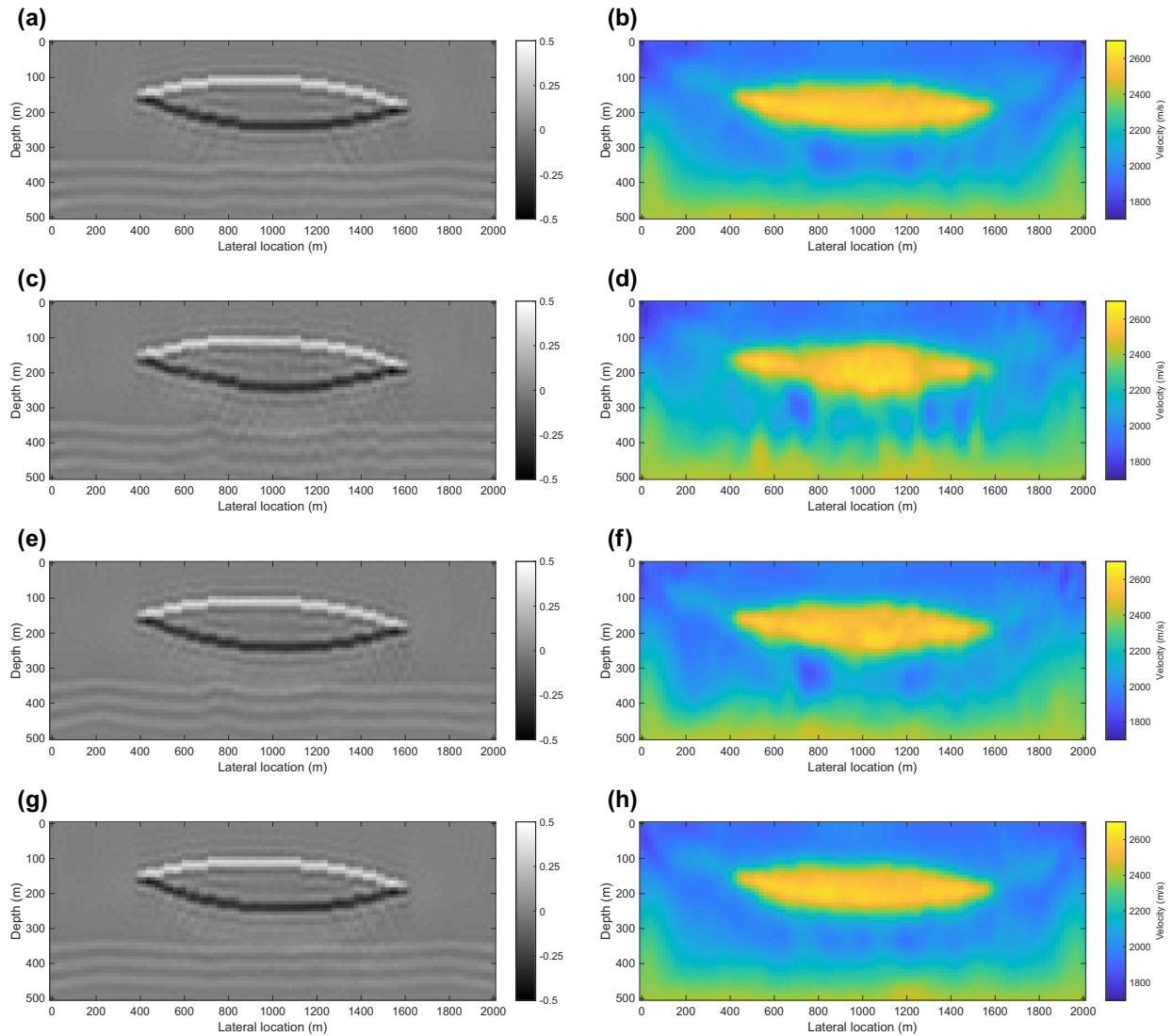


Figure 7 Estimated reflectivity and velocity models: (a and b) standard, (c and d) P50, (e and f) P1 and (g and h) optimized design, respectively. The notable enhancement in the performance of JMI due solely to the choice of survey parameters is easily recognizable.

different stages are not directly comparable. Nevertheless, the classification performance evidently improves through the course of iterations. Additionally, the difference in accuracies between training sets and testing sets becomes insignificant after a couple of hundred realizations, where the classification achieves accuracies well above 90% for both training and testing sets with a small discrepancy between the minimum and the maximum values. This indicates that our network architecture, along with the chosen hyperparameters, manages the bias-variance trade-off reasonably well and successfully relates the survey parameters to the resultant JMI quality.

DISCUSSION

Insufficiencies in quality and quantity of available subsurface information potentially cause uncertainty in \mathbf{R} and \mathbf{W} used to optimize survey parameters in the proposed workflow although its degree varies with the situations. In this respect, we investigated the effect of this uncertainty on the joint migration inversion (JMI) results. Figure 10(a) depicts velocity perturbations applied to the true model where we randomly modified the velocity of each layer with a value in the range of ± 100 m/s. These velocity perturbations enable each

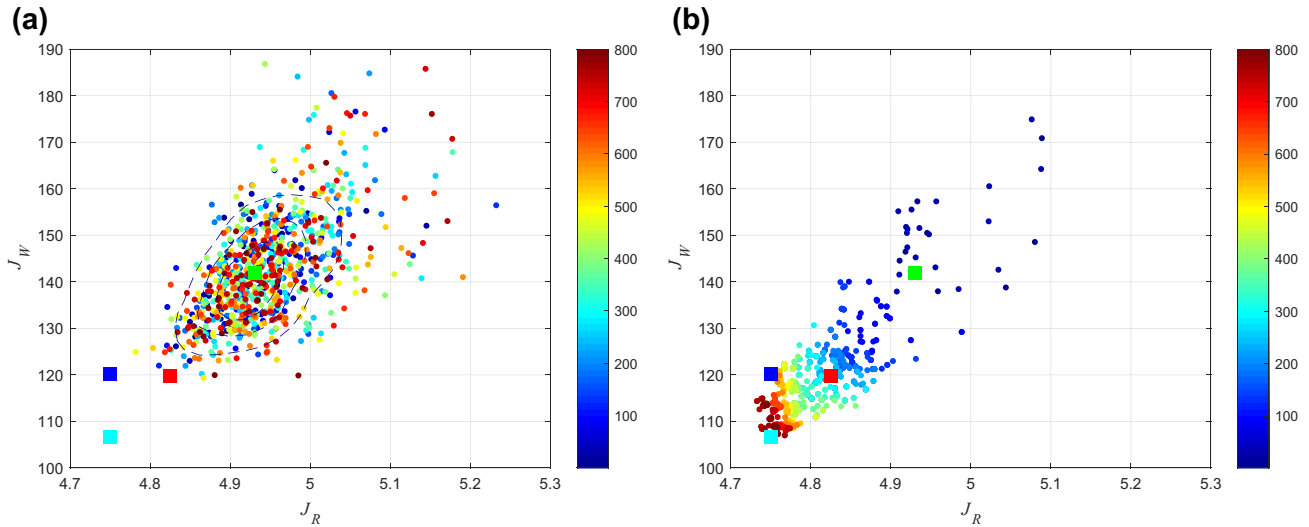


Figure 8 The progress of the optimization from (a) random realization and (b) proposed approach. Plot colours indicate the realization number. The blue, green, red and cyan squares are the results from the standard, P50, P1 and optimized design, respectively. In (a), the three dashed lines are constant probability density contours from kernel density estimation each of which represents the boundary of the area that contains a certain percentage of data points (25%, 50% and 75% from the inner to the outer contour).

subsurface model to employ different \mathbf{R} and \mathbf{W} . For each arbitrary derived subsurface model, we performed full-wavefield modelling (FWMoD) to simulate four different \mathbf{P} ' using the four survey designs discussed earlier (standard, P50, P1 and optimized). We then applied JMI to these four datasets to obtain different reflectivity and velocity estimates. As shown in Fig. 10(b), this procedure was applied to all the subsurface models. Figure 11 shows a comparison among the JMI results from the four different acquisition scenarios for each subsurface model. The vertical axes of Fig. 11(a,b) indicates relative differences in J_R and J_W , respectively, with respect to the corresponding objective-function values from the optimized design. A negative value indicates that the optimized design attains the smaller misfit. The figure clearly shows that our optimized design achieves proper reflectivity and velocity estimates for all the models. Although certain variations can be observed among different subsurface models, the relative relationship among the four designs in terms of JMI quality still holds. For the reflectivity estimation, the standard and optimized designs attain comparable results. Next, P1 produces a somewhat less optimal outcome, followed by P50. For the velocity estimation, a distinct improvement can be seen in the optimized design. It clearly outperforms the standard design as well as the P1 design, the two of which provide almost comparable results. Then, the P50 result is suboptimal for all the subsurface models. Hence, this suggests that blending and sampling operators derived from our approach

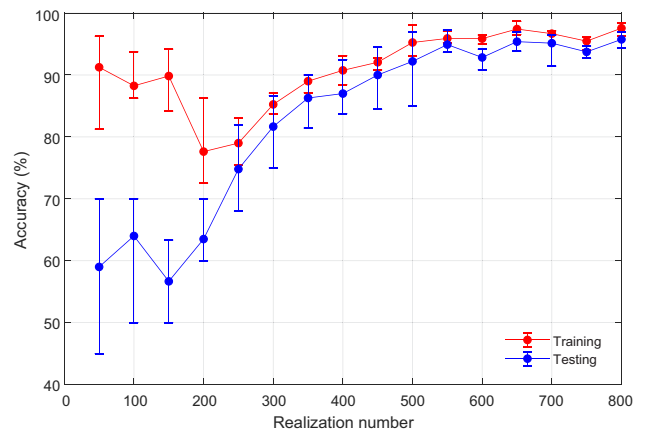


Figure 9 Classification accuracy of the predictive model at every 50 realizations. The red and blue markers represent mean accuracies at each fivefold cross-validation from training and testing sets, respectively. Error bars indicate the minimum and maximum accuracies obtained from each validation.

can still provide optimum JMI results even when prior subsurface properties used for the survey-design scheme are not precisely known.

One of the unique properties of JMI is its ability to explain and utilize multiple reflections. They can provide subsurface illumination at different angles from the same shot and may carry information on areas that primaries cannot reach (Berkhout 2012b; Verschuur and Berkhout 2015).

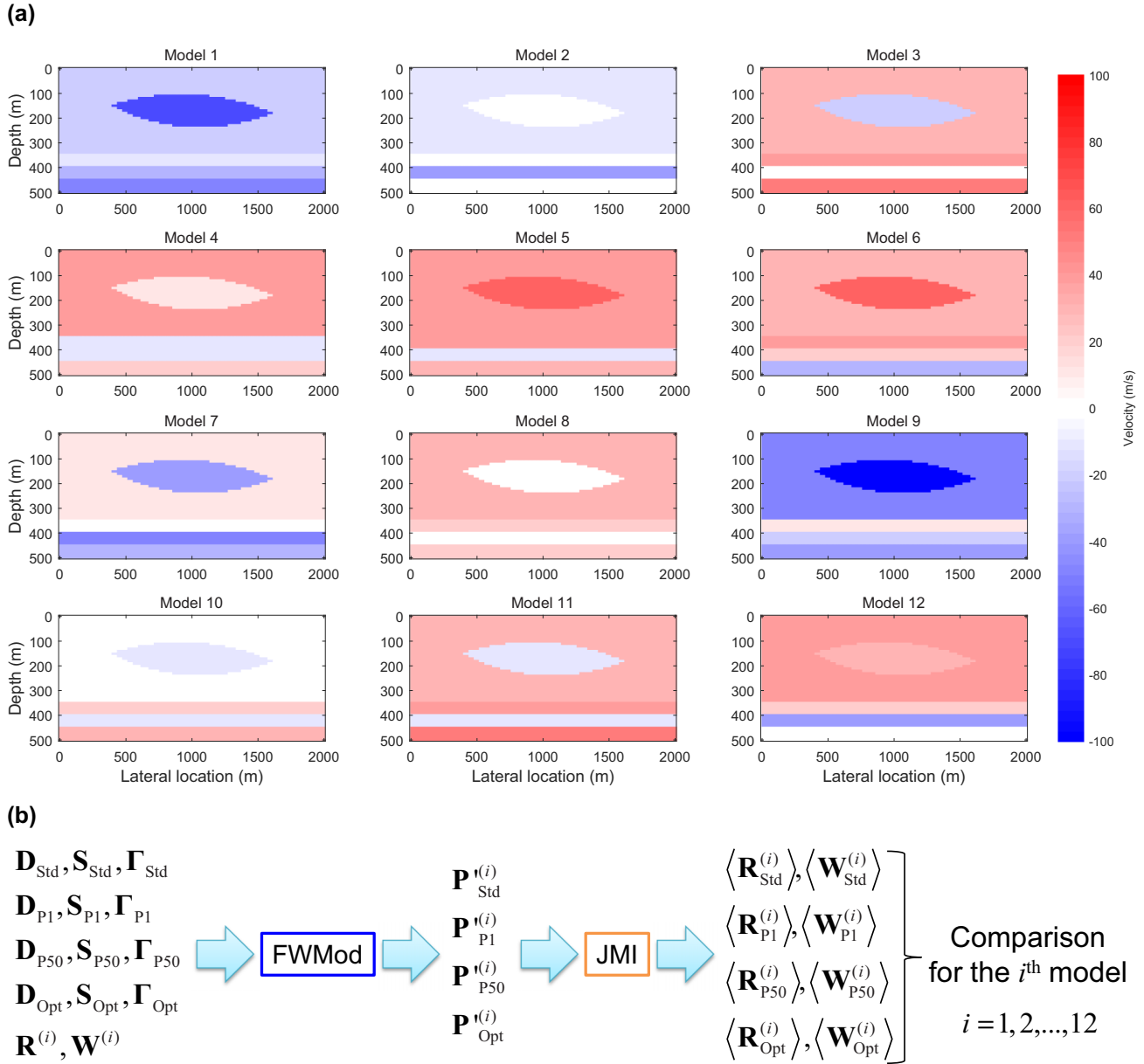


Figure 10 Analysis on the effect of uncertainty in prior subsurface information. (a) Various velocity perturbations applied to the subsurface model. (b) Procedure of the analysis for the i th subsurface model. For subscripts in matrices, ‘Std’, ‘P1’, ‘P50’ and ‘Opt’ refer to the applied acquisition scenario.

Internal multiples are also capable of illuminating the subsurface in downward and upward directions (Davydenko and Verschuur 2013). By treating multiples as signal, they can also assist primaries in the subsurface illumination, provided that proper processing algorithms are implemented. For instance, Berkhout and Verschuur (2016) illustrated a contribution of multiples by comparing JMI results with and without the use of multiples. Seismic surveys are often designed to

sample primary reflections. On the contrary, less emphasis is given to multiples as they are to be attenuated in the subsequent processing. Exploiting features of multiples potentially allows us to ease the sampling requirements in data acquisition (Verschuur and Berkhout 2015). Alternatively, it can enhance the subsurface illumination with the same acquisition footprint (Kumar *et al.* 2016). The proposed workflow, aimed at improving the JMI results, then enables us to derive

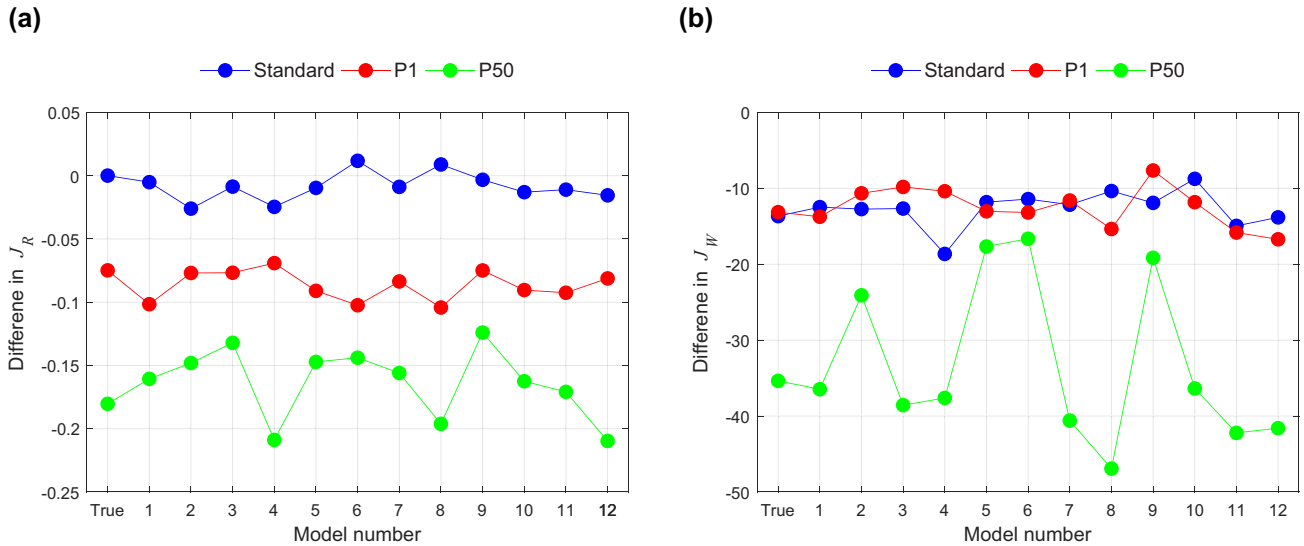


Figure 11 JMI results for different designs from different subsurface models: (a) reflectivity estimate and (b) velocity estimate. The plots in blue, red and green show relative differences between JMI results with respect to the optimized design for the standard, P1 and P50 design, respectively.

acquisition scenarios that can account for the effect of both primary and multiple reflections in the area of interest.

In this study, we utilize the standard implementation of JMI based on Staal and Verschuur (2013) and Staal (2015), which disregards certain subsurface characteristics as mentioned previously. In this respect, some recent studies have been carried out towards exploring phenomena which are not explained in this standard framework such as mode conversion (Berkhout 2014b), angle-dependent reflectivity (Qu, Sun and Verschuur 2018), horizontal wave propagation (Masaya and Verschuur 2017) and anisotropy (Alshuhail and Verschuur 2019). The research has been also extended to other perspectives such as applications to 3D data (El-Marhoul and Verschuur 2016), time-lapse data (Qu and Verschuur 2017) and integration of JMI and full-waveform inversion (FWI) (Sun *et al.* 2018). Furthermore, the presented numerical examples were designed fundamentally to demonstrate the proposed survey-design scheme. To this end, a simple 2D subsurface model was chosen which does not reflect the geological complexity encountered in actual fields. Hence, our future work is aligned with the recent developments in JMI coupled with the use of more complex and realistic 3D subsurface scenarios. To examine the effect of acquisition design on the JMI result, our study employed the idealized situation comprising FWMod as a forward-modelling tool and JMI as a property estimation tool. It is worth noting that the proposed approach is not necessarily limited to this implementation as the contents within the overall scheme can be freely and

flexibly altered depending on the needs of users. A possible option would be to replace them with different methodologies, for example, different forward-modelling schemes such as a finite-difference modelling as well as different velocity estimation and/or imaging algorithms such as FWI and least-squares migration (LSM). Our approach would then be able to design a seismic survey that can meet the quality requirements of the applied process(es).

This study demonstrates that optimally designed survey parameters lead to JMI results that are better than the results obtained with the standard scenario. Alternatively, the same results could have been obtained in a more efficient and cost-effective manner. Furthermore, the implementation of dispersed source array (DSA) is of help in contributing to a health, safety and environment (HSE) perspective (Caporal *et al.* 2018). The emission of acoustic energy may incur a potential environmental risk, particularly in marine surveys. Sound pressure level (SPL) and sound exposure level (SEL) are of primary concern to determine the effects of an acoustic source on the marine environments, in particular on marine mammals. Airgun clusters that generate an impulsive signal are widely used in the industry. As a broad frequency range of acoustic energy is instantaneously generated, the technique inevitably accentuates the peak pressure. Furthermore, these conventional marine sources inevitably emit high-frequency components, for example, above 100 Hz, which are normally discarded in seismic imaging yet significantly overlap with the hearing ranges of odontocetes and pinnipeds

(Southall *et al.* 2007). As illustrated in Fig. 3, the use of dedicated narrowband sources decreases both peak amplitude and total energy for each shot. It is also capable of preventing a seismic survey from the emission of unnecessary frequency components, yet of acquiring the information needed to characterize the subsurface. Hence, the technique contributes to reducing both SPL and SEL. The optimized DSA scenario uses reduced source locations without adversely affecting the data quality. These elements lower the total energy accumulated over the survey duration, known as cumulative SEL. Our optimization scheme is therefore capable of designing an efficient, cost-effective and environmentally favourable seismic survey that also enhances the performance of JMI.

CONCLUSION

The survey-design workflow introduced in this study simultaneously optimizes parameters that determine source blending and the spatial sampling of detectors and sources. The numerical examples in this study demonstrate that the joint migration inversion (JMI) results can vary with the design of survey parameters. Optimally designed parameters lead to the enhancement of both reflectivity and velocity models estimated directly from blended and irregularly sampled data. The proposed approach integrates a genetic algorithm (GA) and a convolutional neural network (CNN) to make the optimization of survey designs feasible within an affordable computation time. Genetic operators coupled with non-dominated sorting are capable of minimizing errors in reflectivity and velocity models by simultaneously updating the survey parameters. The workflow can effectively deal with complex blending and spatial-sampling schemes that employ the dispersed source array (DSA) concept. The neural network architecture applied to this study successfully relates the performance of JMI to the choice of survey parameters. The classification through a CNN helps us to minimize the computation of suboptimal designs while keeping the optimal ones into the iterative scheme. The proposed approach provides survey parameters to enhance the results of JMI that have been obtained by directly processing blended and irregularly sampled data without the need for deblending or data reconstruction. The resultant acquisition scenarios allow us to optimally estimate subsurface properties at an affordable cost and with a low environmental footprint.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation and gratitude to all the sponsors of DELPHI

consortium and INPEX Corporation for their continuous support.

ORCID

Shotaro Nakayama 

<https://orcid.org/0000-0003-4989-2778>

REFERENCES

- Abma, R., Zhang, Q., Arogunmati, A. and Beaudoin, G. 2012. An overview of BPs marine independent simultaneous source field trials. 82nd SEG annual meeting, Las Vegas, NV, USA, Expanded Abstracts, 1–5.
- Ahmadizar F., Soltanian K., AkhlaghianTab F. and Tsoulos I. 2015. Artificial neural network development by means of a novel combination of grammatical evolution and genetic algorithm. *Engineering Applications of Artificial Intelligence* **39**, 1–13.
- Aldawood A., Hoteit I. and Alkhalifah T. 2014. The possibilities of compressed-sensing-based Kirchhoff prestack migration. *Geophysics* **79**, S113–S120.
- Alshuhail A.A. and Verschuur D.J. 2019. Robust estimation of vertical symmetry axis models via joint migration inversion: including multiples in anisotropic parameter estimation. *Geophysics* **84**, C57–C74.
- Bak, S.H., Rask, N. and Risi, S. 2016. Towards adaptive evolutionary architecture. *Proceedings of the 5th International Conference on Evolutionary and Biologically Inspired Music, Sound, Art and Design, EvoMUSART 2016*, Porto, Portugal, pp. 47–62.
- Beasley, C.J., Ronald, E.C. and Jiang, Z. 1998. A new look at simultaneous sources. 68th SEG annual meeting, New Orleans, LA, USA, Expanded Abstracts, 133–135.
- Ben-Hadj-Ali H., Operto S. and Virieux J. 2011. An efficient frequency-domain full waveform inversion method using simultaneous encoded sources. *Geophysics* **76**, R109–R124.
- Berkhout, A.J. 1982. Seismic Migration, Imaging of Acoustic Energy by Wave Field Extrapolation, Part A: Theoretical Aspects. Elsevier.
- Berkhout A.J. 2008. Changing the mindset in seismic data acquisition. *The Leading Edge* **27**, 924–938.
- Berkhout A.J. 2012a. Blended acquisition with dispersed source arrays. *Geophysics* **77**, A19–A23.
- Berkhout A.J. 2012b. Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging. *Geophysics* **77**, S43–S50.
- Berkhout A.J. 2014a. Review paper: an outlook on the future of seismic imaging, part I: forward and reverse modelling. *Geophysical Prospecting* **62**, 911–930.
- Berkhout A.J. 2014b. Review paper: an outlook on the future of seismic imaging, part III: joint migration inversion. *Geophysical Prospecting* **62**, 950–971.
- Berkhout A.J., Blacquièrre G. and Verschuur D.J.E. 2012. Multiscattering illumination in blended acquisition. *Geophysics* **77**, P23–P31.
- Berkhout A.J. and Verschuur D.J. 2016. Enriched seismic imaging by using multiple scattering. *The Leading Edge* **35**, 128–133.

- Bishop, C. 2006. *Pattern Recognition and Machine Learning*. Springer.
- Boonyasiriwat, C. and Schuster, G.T. 2010. 3D multisource full-waveform inversion using dynamic random phase encoding. 80th SEG annual meeting, Denver, CO, USA, Expanded Abstracts, 1044–1049.
- Bouska J. 2010. Distance separated simultaneous sweeping, for fast, clean, vibroseis acquisition. *Geophysical Prospecting* 58, 123–153.
- Caporal M., Blacqui re G. and Davydenko M. 2018. Broadband imaging via direct inversion of blended dispersed source array data. *Geophysical Prospecting* 66, 942–953.
- Chen Y., Yuan J., Zu S., Qu S. and Gan S. 2015. Seismic imaging of simultaneous-source data using constrained least-squares reverse time migration. *Journal of Applied Geophysics* 114, 32–35.
- Davies, E., Tew, P., Glowacki, D.R., Smith, J. and Mitchell, T. 2016. Evolving atomic aesthetics and dynamics. *Proceedings of the 5th International Conference on Evolutionary and Biologically Inspired Music, Sound, Art and Design, EvoMUSART 2016*, Porto, Portugal, 17–30.
- Davydenko, M. and Verschuur, D.J. 2013. Full wavefield migration, using internal multiples for undershooting. 83rd SEG annual meeting, Houston, TX, USA, Expanded Abstracts, 3741–3745.
- Deb K., Pratap A., Agarwal S. and Meyarivan T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 182–197.
- El-Marhoul, B. and Verschuur, D.J. 2016. High-resolution, integrated 3D joint migration inversion of surface and VSP data. 86th SEG annual meeting, Dallas, TX, USA, Expanded Abstracts, 672–676.
- Florez, K.A., Mantilla, J.G. and Ramirez, A.B. 2016. Full waveform inversion (FWI) in time for seismic data acquired using a blended geometry. *2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, Bucaramanga, Colombia, 1–5.
- Hahnloser R.H., Sarpeshkar R., Mahowald M.A., Douglas R.J. and Seung H.S. 2000. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405, 947–951.
- Hennenfent G. and Herrmann F.J. 2008. Simply denoise: wavefield reconstruction via jittered undersampling. *Geophysics* 73, V19–V28.
- Herrmann F.J. 2010. Randomized sampling and sparsity: getting more information from fewer samples. *Geophysics* 75, WB173–WB187.
- Hinton G., Deng L., Yu D., Dahl G., Rahman Mohamed A., Jaitly N. et al. 2012. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine* 29, 82–97.
- Huang Y. and Schuster G.T. 2012. Multisource least-squares migration of marine streamer and land data with frequency-division encoding. *Geophysical Prospecting* 60, 663–680.
- Ibaraki S., Tomita I. and Sugimoto K. 2015. Aerodynamic design optimization of centrifugal compressor impeller based on genetic algorithm and artificial neural network. *Mitsubishi Heavy Industries Technical Review* 52, 77–82.
- Jeong W., Pyun S., Son W. and Min D.-J. 2013. A numerical study of simultaneous-source full waveform inversion with l1-norm. *Geophysical Journal International* 194, 1727–1737.
- Kontakis, A. and Verschuur, D.J. 2015. Combined focal and coherency-based deblending strategy. 85th SEG annual meeting, New Orleans, Louisiana, USA, Expanded Abstracts, 4667–4672.
- Kramer, O. 2016. *Machine learning for evolution strategies*. Springer.
- Kramer, O. 2017. *Genetic algorithm essentials*. Springer.
- Krebs J.R., Anderson J.E., Hinkley D., Neelamani R., Lee S., Baumbstein A. et al. 2009. Fast full-wavefield seismic inversion using encoded sources. *Geophysics* 74, WCC177–WCC188.
- Kumar A., Blacqui re G., Pedersen M.W.P. and Goertz A. 2016. Full-wavefield marine survey design using all multiples. *Geophysics* 81, P1–P12.
- LeCun Y., Bottou L., Bengio Y. and Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86, 2278–2324.
- Leung M.K.K., Xiong H.Y., Lee L.J. and Frey B.J. 2014. Deep learning of the tissue-regulated splicing code. *Bioinformatics* 30, i121–i129.
- Lin, T.T.Y. and Herrmann, F.J. 2009. Designing simultaneous acquisitions with compressive sensing. *71st EAGE Conference and Exhibition*, Amsterdam, The Netherlands, Extended Abstracts, S006.
- Mahdad A., Doulgeris P. and Blacqui re G. 2011. Separation of blended data by iterative estimation and subtraction of blending interference noise. *Geophysics* 76, Q9–Q17.
- Marcelin J.-L. 2004. A metamodel using neural networks and genetic algorithms for an integrated optimal design of mechanisms. *The International Journal of Advanced Manufacturing Technology* 24, 708–714.
- Masaya, S. and Verschuur, D.J. 2017. Seismic modeling refracted and reflected waves including internal multiples by one-way propagators. 87th SEG annual meeting, Houston, TX, USA, Expanded Abstracts, 4138–4142.
- Monteagudo, A. and Santos, J. 2015. Evolutionary optimization of cancer treatments in a cancer stem cell context. *Genetic and Evolutionary Computation Conference 2015, GECCO 2015*, Madrid, Spain, July 11–15, pp. 233–240.
- Moore, I., Dragoset, B., Ommundsen, T., Wilson, D., Ward, C. and Eke, D. 2008. Simultaneous source separation using dithered sources. 78th SEG annual meeting, Las Vegas, NV, USA, Expanded Abstracts, 2806–2810.
- Mosher, C.C., Kaplan, S.T. and Janiszewski, F.D. 2012. Non-uniform optimal sampling for seismic survey design. *74th EAGE Conference and Exhibition*, Copenhagen, Denmark, Extended Abstracts, X034.
- Murphy, K.P. 2012. *Machine Learning: A Probabilistic Perspective*. MIT Press.
- Nakayama S., Blacqui re G., Ishiyama T. and Ishikawa S. 2019. Blended-acquisition design of irregular geometries towards faster, cheaper, safer and better seismic surveying. *Geophysical Prospecting* 67, 1498–1521.
- Nakayama S., Mercado G., Benson M., Belaid K. and Garden M. 2015. Field-wide implementation of time and distance separated source techniques on a 3D OBC survey offshore Abu Dhabi, UAE. *First Break* 33, 47–53.
- Nemeth T., Wu C. and Schuster G.T. 1999. Least-squares migration of incomplete reflection data. *Geophysics* 64, 208–221.

- Perez-Liebana, D., Dieskau, J., Hunermund, M., Mostaghim, S. and Lucas, S.M. 2015. Open loop search for general video game playing. *Genetic and Evolutionary Computation Conference*, Madrid, Spain, pp. 337–344.
- Qu, S., Sun, Y. and Verschuur, D.J. 2018. Mitigating amplitude versus ray-parameter effect in joint migration inversion using a zero-lag cross-correlation objective function of redatumed wavefields. 88th SEG annual meeting, Anaheim, CA, USA, Expanded Abstracts, 1133–1137.
- Qu, S. and Verschuur, D.J. 2017. Simultaneous joint migration inversion for semi-continuous time-lapse seismic data. 87th SEG annual meeting, Houston, TX, USA, Expanded Abstracts, 5808–5813.
- Sakaguchi D., Tun M., Numakura R. and Wang B. 2018. Global optimization of recirculation flow type casing treatment in centrifugal compressors of turbochargers. *Journal of Mechanical Engineering Science* 232, 4461–4471.
- Sato T. and Fujita M. 2016. An investigation of design optimization of electric devices supported by machine learning. *IEICE Technical Report* 116, 289–294.
- Schmidhuber J. 2015. Deep learning in neural networks. An overview. *Neural Networks* 61, 85–117.
- Scirea, M., Togelius, J., Eklund, P. and Risi, S. 2016. Metacompose: a compositional evolutionary music composer. *Proceedings of the 5th International Conference on Evolutionary and Biologically Inspired Music, Sound, Art and Design, EvoMUSART 2016*, Porto, Portugal, 202–217.
- Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene, C.R.J. *et al.* 2007. Marine mammal noise exposure criteria: initial scientific recommendations. *Aquatic Mammals* 33, no. 4, 411–522.
- Sreekanth J. and Datta B. 2011. Comparative evaluation of genetic programming and neural network as potential surrogate models for coastal aquifer management. *Water Resources Management* 25, 3201–3218.
- Staal, X.R. 2015. Combined imaging and velocity estimation by joint migration inversion. PhD thesis, Delft University of Technology, The Netherlands.
- Staal, X.R. and Verschuur, D.J. 2013. Joint migration inversion, imaging including all multiples with automatic velocity update. 75th EAGE annual meeting, London, UK, Extended Abstracts, 10–13.
- Sun, Y., Kim, Y.S., Qu, S., Verschuur, D.J., Almomin, A. and van Borselen, R. 2018. Angle-dependent full wavefield migration based upon full waveform inversion and joint migration inversion. 88th SEG annual meeting, Anaheim, CA, USA, Expanded Abstracts, 4357–4361.
- Tang, Y. and Biondi, B. 2009. Least-squares migration/inversion of blended data. 79th SEG annual meeting, Houston, TX, USA, Expanded Abstracts, 2859–2863.
- Tarantola A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49, 1259–1266.
- Tenne, Y. and Goh, C.-K. 2010. *Computational intelligence in expensive optimization problems*. Springer.
- Verschuur D.J. and Berkhouit A.J. 2015. From removing to using multiples in closed-loop imaging. *The Leading Edge* 34, 744–759.
- Verschuur D.J., Staal X.R. and Berkhouit A.J. 2016. Joint migration inversion: simultaneous determination of velocity fields and depth images using all orders of scattering. *The Leading Edge* 35, 1037–1046.
- Verschuur D.J.E. and Berkhouit A.J. 2011. Seismic migration of blended shot records with surface-related multiple scattering. *Geophysics* 76, A7–A13.
- Wang, C., Yingst, D., Brittan, J., Farmer, P. and Leveille, J. 2014. Fast multi-parameter anisotropic full waveform inversion with irregular shot sampling. 84th SEG annual meeting, Denver, CO, USA, Expanded Abstracts, 1147–1151.
- Xie, L. and Yuille, A. 2017. Genetic CNN. *IEEE International Conference on Computer Vision*, pp. 1379–1388.