Geophysical Journal International

Geophys. J. Int. (2018) **214,** 1683–1700
Advance Access publication 2018 June 02
GJI Marine geosciences and applied geophysics

doi: 10.1093/gji/ggy221

Decimetric-resolution stochastic inversion of shallow marine seismic reflection data: dedicated strategy and application to a geohazard case study

Giuseppe Provenzano,[1] Mark E. Vardy[2] and Timothy J. Henstock[1]

[1]*Ocean and Earth Science, National Oceanography Centre, Southampton, University of Southampton, European Way, Southampton* SO14 3ZH*, UK. E-mail: gp1n13@soton.ac.uk*
[2]*Marine Geosciences Group, National Oceanography Centre, Southampton, European Way, Southampton* SO14 3ZH*, UK*

Accepted 2018 June 1. Received 2018 May 23; in original form 2018 February 12

SUMMARY

Characterization of the top 10–50 m of the subseabed is key for landslide hazard assessment, offshore structure engineering design and underground gas-storage monitoring. In this paper, we present a methodology for the stochastic inversion of ultra-high-frequency (UHF, 0.2–4.0 kHz) pre-stack seismic reflection waveforms, designed to obtain a decimetric-resolution remote elastic characterization of the shallow sediments with minimal pre-processing and little *a priori* information. We use a genetic algorithm in which the space of possible solutions is sampled by explicitly decoupling the short and long wavelengths of the *P*-wave velocity model. This approach, combined with an objective function robust to cycle skipping, outperforms a conventional model parametrization when the ground-truth is offset from the centre of the search domain. The robust *P*-wave velocity model is used to precondition the width of the search range of the multiparameter elastic inversion, thereby improving the efficiency in high-dimensional parametrizations. Multiple independent runs provide a set of independent results from which the reproducibility of the solution can be estimated. In a real data set acquired in Finneidfjord, Norway, we also demonstrate the sensitivity of UHF seismic inversion to shallow subseabed anomalies that play a role in submarine slope stability. Thus, the methodology has the potential to become an important practical tool for marine ground model building in spatially heterogeneous areas, reducing the reliance on expensive and time-consuming coring campaigns for geohazard mitigation in marine areas.

Key words: Geomechanics; Inverse theory; Acoustic properties; Controlled source seismology.

1 INTRODUCTION

A quantitative high-resolution physical model of the top 50 m of the subseabed is of key importance for a wide range of geohazard and offshore engineering applications: identification of potential shallow landsliding (Vanneste *et al.* 2013); monitoring of gas storage sites (Cevatoglu *et al.* 2015); and assessment of offshore structures stability (Vardy *et al.* 2012). Currently, engineering-scale sediment characterization relies heavily on direct sampling of the seabed and *in situ* measurements (e.g. Stoker *et al.* 2009). This is expensive and time-consuming for large areas, as well as being liable to alter the sediment properties during the coring process (Clare *et al.* 2017; Monrigal *et al.* 2017; Vardy *et al.* 2017).

Variations in lithology and pore-pressure conditions produce mechanical layering in the shallow subsurface, that can be conducive to changes in the stability conditions of submarine slopes (Vardy *et al.* 2012; Vanneste *et al.* 2013). Corresponding anomalies in compressibility, shear resistance and density, have recognizable footprints on the amplitude and phase of the multi-offset reflected seismic waveforms (Ostrander 1984; Ruthenford & Williams 1989). Therefore, seismic reflection data have the potential to be used as a remote sensing tool for shallow geohazard estimation (Mallick & Dutta 2002; Vardy *et al.* 2017). As opposed to reservoir-scale seismic exploration, ultra-high-frequency (UHF, 0.2–4.0 kHz) multi-channel marine seismic reflection data are historically under-used in industry for offshore engineering design, being most often limited to a semi-quantitative interpretation of the reflection amplitudes and facies geometries. Recent advances, however, have shown the potential use of UHF data as a quantitative tool (Vardy *et al.* 2017), from acoustic quality factor estimation (Pinson *et al.* 2008) and acoustic impedance inversion (Vardy 2015), to elastic pre-stack inversion of the full waveform (Provenzano *et al.* 2016, 2017).

Full waveform inversion (FWI) is a process by which the initial state of knowledge about a given parametrization of the propagation medium (model) is iteratively improved, by maximizing the fitness between the observed seismograms and the data computed using a forward modelling operator that approximates the wave equation (Mora 1980; Tarantola 1984; Virieux & Operto 2009). The oscillating nature of the data makes the high-frequencies of the inversion strongly non-linear with respect to the long-wavelengths of the *P*-wave model (cycle skipping; Tarantola 1984; Fichtner 2011), and non-linearity is enhanced by the interdependency among the different elastic parameters (Operto *et al.* 2013; Gholami *et al.* 2013a,b). A deterministic inversion thus requires an accurate starting model and an equally accurate estimation of the Hessian matrix (Virieux & Operto 2009; Fichtner 2011; Operto *et al.* 2013; Dagnino *et al.* 2014).

Alternatively, the inversion can be tackled with a stochastic approach, which samples the model space with a density proportional to the *Posterior Probability Density function (PPD)* (Sambridge & Mosegaard 2002; Sen & Stoffa 2013). In addition to being less reliant upon the accuracy of the starting solution, stochastic seismic inversion, in principle, allows lots of potentially useful information on the performance of the inversion to be extracted from the PPD (Sambridge 1999; Tarantola 2005). This includes, but it is not limited to, the multiparameter solution uncertainties and crosstalk resulting from the noise content, the limited offset and bandwidth of the data, and the inherent interdependency of coupled parameters. However, most often deterministic FWI is the only feasible approach in 3-D environments because of the high number of model evaluations required in high-dimensionality spaces and the computing cost of the seismogram modelling.

Stochastic optimizers, such as genetic algorithm (GA; Goldberg 1989), simulated annealing (Rothman 1985) and particle swarm optimization (Kennedy & Eberhart 1995), implement analogies between numerical optimization and natural biological and physical phenomena. They represent a compromise between the extensive exploration of the model space, and the exploitation of the current state of information about the model (Sambridge & Mosegaard 2002). Although providing a biased posterior distribution of models (Sambridge 1999; Sen & Stoffa 1996; Sambridge & Mosegaard 2002; Aleardi & Mazzotti 2017), they are computationally more affordable than a pure Bayesian approach, and more robust against local minima entrapment than a deterministic algorithm (Tarantola 1984; Sambridge & Mosegaard 2002; Sen & Stoffa 2013).

These algorithms have been applied to FWI of both land and marine data, for reservoir characterization and shallow drilling hazard assessment, (Stoffa & Sen 1991; Sen & Stoffa 1992; Mallick & Dutta 2002; Mallick & Adhikari 2015; Sajeva *et al.* 2016; Aleardi & Mazzotti 2017), especially when a 1-D parametrization of the problem is sensible, and hence the number of unknowns lower. Stochastic optimizers would be especially beneficial to the inversion of seismic data for shallow geohazard purposes, because of the difficulties of obtaining a reliable starting elastic model from a typical UHF reflection data set, as well as the capability of providing solution error bounds for geological interpretation and engineering design (Morgan *et al.* 2014; Vardy 2015; Vardy *et al.* 2015).

In this paper, we present a submetric resolution stochastic seismic inversion methodology based upon a GA, custom-built for limited offset, limited bandwidth seismic reflection data. The performance of the stochastic algorithm is tested on both synthetic and real data. In summary, this paper aims at demostrating that:

(1) Our proposed strategy is effective at reducing the dependency from the search range design compared to a genetic algorithm with a conventional parametrization.

(2) A two-stage multi-parameter inversion with data-driven preconditioning improves the performance of the inversion, and the interpretability of the estimated elastic model.

(3) Elastic layering, corresponding to key-features for slope instability, can be successfully identified and quantitatively described; in an area subject to shallow landsliding (Finneidfjord, northern Norway). A sediment bed corresponding to the glide plane of multiple-landslide events is accurately located and its changes in pore-fluid saturation quantified.

This underpins the application of UHF seismic reflection data as a remote sensing tool for ground-model building and geohazard estimation, reducing the need for expensive and time-consuming coring campaigns.

2 DATA AND METHODOLOGY OUTLINE

2.1 Example estimation of a starting model from UHF data

Shallow marine seismic reflection data typically suffer from the limited offset and lack of low frequencies of the source-acquisition system, which severely limits the sensitivity to the broad-band multiparameter space (Mora 1980; Jannane 1989). Furthermore, most shallow marine data sets are single-component and do not contain post-critical reflections, which limits the possibility of uncoupling the effect of the different elastic parameters (Kormendi & Dietrich 1991), as well as the sensitivity to the long wavelength distribution of shear properties and density (Operto *et al.* 2013; Provenzano *et al.* 2017). As a result, the solution of the inverse problem is highly dependent upon the starting model (Virieux & Operto 2009; Operto *et al.* 2013).

In high-resolution reflection seismic data, due to the limited aperture, a starting model is typically obtained from limited-offset moveout velocity analysis (e.g. Aleardi *et al.* 2016), from which an interval *P*-wave velocity profile is derived (Dix 1955). The uncertainty associated with this process can be high, and bias the seismic inversion towards unreliable solutions.

As an example, we simulate a synthetic UHF seismic reflection data set (Fig. 1a), acquired in a shallow marine context with a 1-D isotropic elastic shallow subsurface (Fig. 1b); the model is a 7-m thick stack of 30 homogeneous layers, representing a site with a strong *P*-wave velocity gradient and properties reasonable for near-surface sediments (Hamilton 1970). We suppose no data-independent information is available in the site about the *P*-wave velocity trend, and we want to derive interval velocities from reflection moveout semblance analysis.

The source wavelet is the signature of a Boomer electro-acoustic plate, whose effective bandwidth spans from 0.2 to 2.5 kHz (Verbeek & McGee 1995); the simulated streamer comprises 60 channels with 1-m spacing and minimum offset of 13 m. A realistic level of band-limited random noise has been added to the data; namely, the signal-to-noise ratio, computed with respect to the seafloor reflection amplitude, is equal to 50. Note in Fig. 1(b) how the strong *P*-wave velocity gradient determines a compression of the time delay among reflections at the longest offsets, which makes it difficult to accurately use moveout-based velocity estimation tools. The broad moveout semblance maxima (Fig. 2) produce uncertainties in the

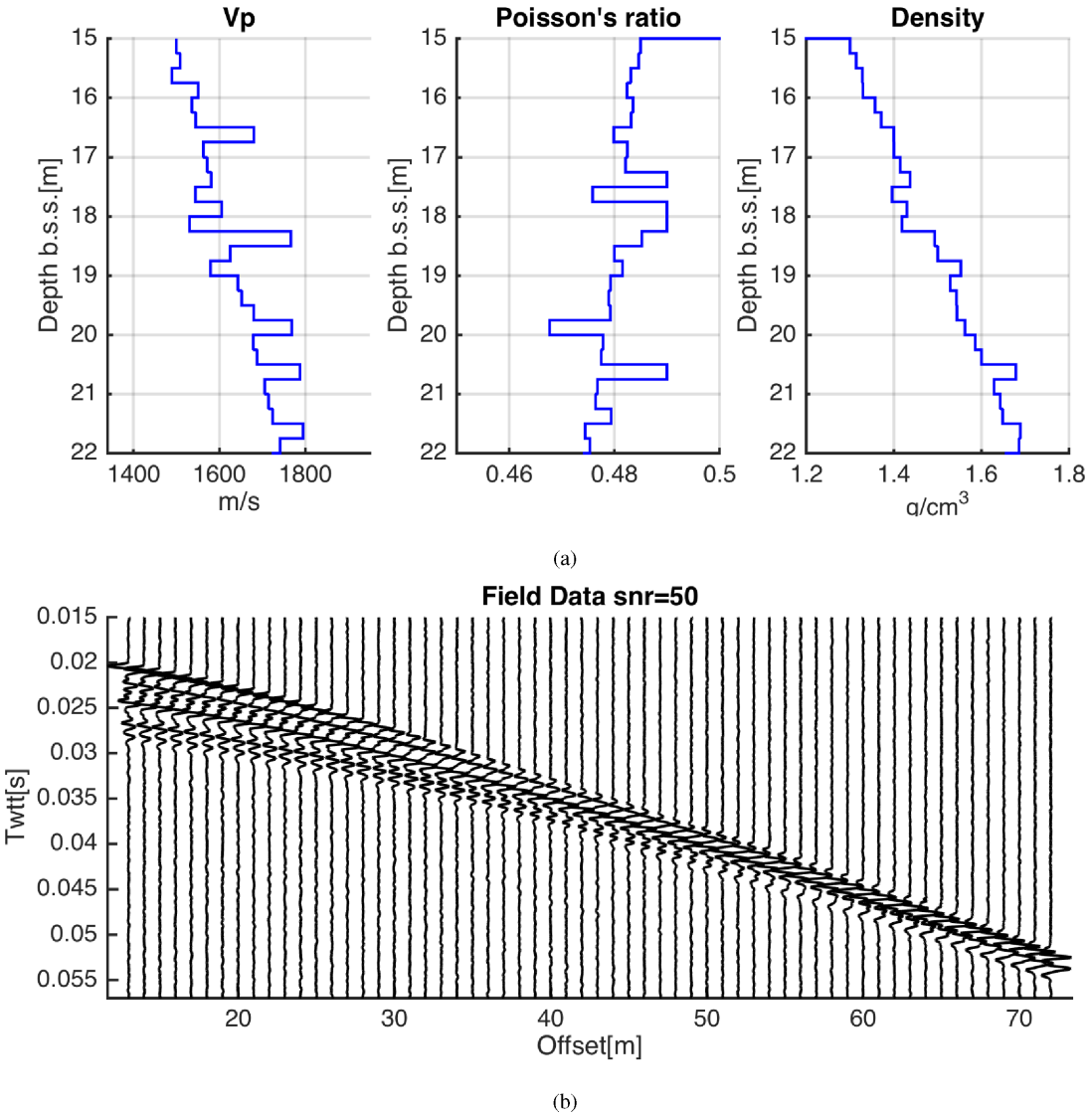

Figure 1. 1-D elastic model and synthetic data. (a) Horizontally layered elastic model, parametrized as *P*-wave velocity (V_p), Poisson's ratio (ν) and density (ρ); depth relative to the sea surface. (b) Common-shot multichannel reflection seismic data simulating the acquisition in a shallow marine environment. The seismogram is computed using the *Oases* software (Schmidt & Jensen 1985; Schmidt & Tango 1986), and contaminated with band-limited random noise; signal-to-noise ratio is equal to 50, computed with respect to the strongest reflection

velocity picking that translate into a broad range of possible interval *P*-wave velocity profiles; such variability is in general beyond a cycle-skipping safe domain (Virieux & Operto 2009; Fichtner 2011; Aleardi *et al.* 2016), and suggests that moveout velocity analysis is not a robust tool to obtain a starting *P*-wave model for local-search FWI. In the absence of reliable Cone Penetrometer Tests (CPT) and/or core-logs, the starting *S*-wave velocity and density profiles would also be derived from this *P*-wave model, further jeopardizing the chance to converge to the global minimum of the multiparameter objective function.

The latter is in fact a typical situation in engineering-scale seismic exploration and clarifies the potential benefits of a seismic inversion strategy that reduces the dependency upon the initial state of information. In the next sections, we demonstrate that, although the inherent sensitivity limitation of this kind of data cannot be overcome, our stochastic approach can still provide a robust solution when a wrong velocity trend is used to build the starting model search range.

2.2 Proposed genetic algorithm-based strategy

Stochastic optimizers like the GA (Goldberg 1989) represent a compromise between the systematic exploration of the model space and the exploitation of the current state of information about the physical model (Sambridge & Mosegaard 2002); although more robust to cycle-skipping than deterministic FWI (Sajeva *et al.* 2016), they can similarly suffer from local-minima entrapment, depending on the model space dimensionality and the nature of the objective function (Sajeva *et al.* 2017).

In this paper, we introduce relevant changes to a classic evolutionary algorithm framework, in order to attenuate the bias inherited from ill-characterized *a priori* distributions, and reduce the effective size of the model space. The proposed strategy is based on the genetic algorithm developed by Vardy (2015), adapted to the multiparameter FWI of multichannel data (Fig. 3). The subsurface is assumed to be horizontally stratified (Fig. 1), and each isotropic elastic layer is parametrized in terms of *P*-impedance (*Z*), Poisson's

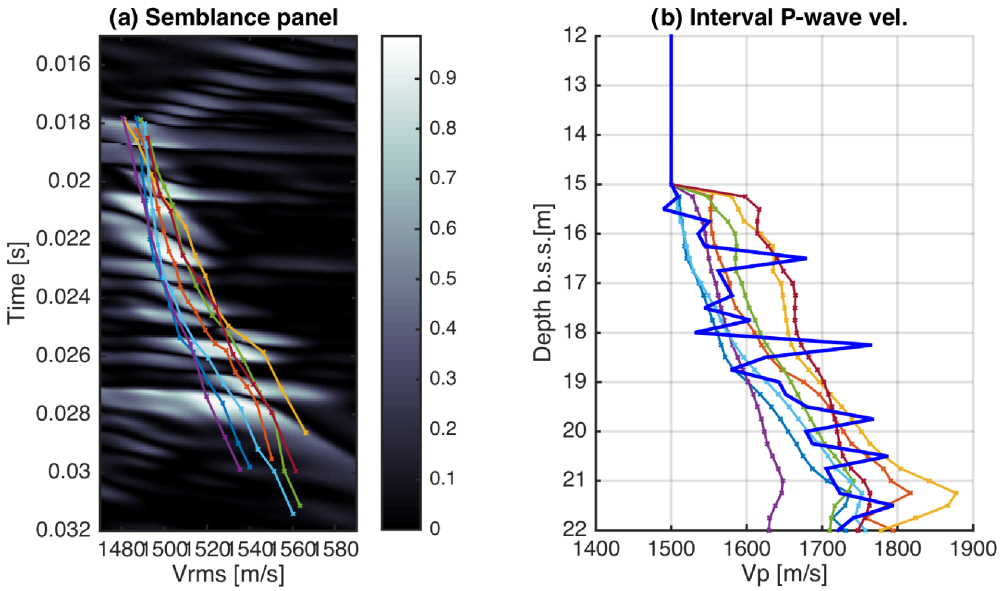

Figure 2. Deriving low-frequency *P*-wave velocity model from NMO analysis: (a) Stacking velocity semblance panel with possible V_{rms} pickings; (b) interval *P*-wave velocity models (dotted-lines) derived using Dix (1955) equation from the V_{rms}, and true V_p model (solid blue).

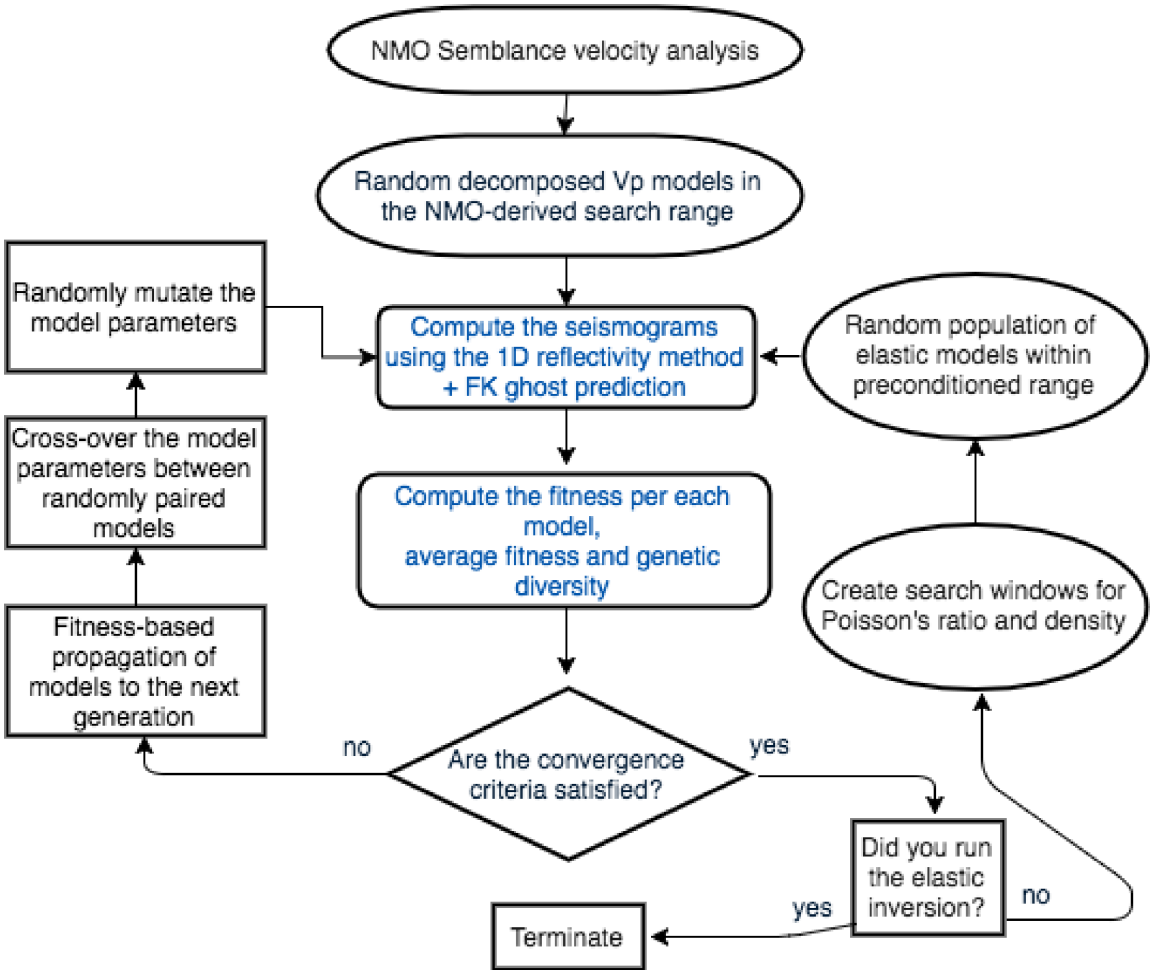

Figure 3. Inversion workflow. The flow chart summarises the proposed two-stage inversion strategy for the reconstruction of the multi-parameter elastic model. The steps for the computation of the models performance for each generation are highlighted in blue. The forward model is parallelized using openMP.

ratio (ν) and density (ρ) (Debski & Tarantola 1995; Igel *et al.* 1996; Provenzano *et al.* 2017). The inversion workflow can be divided into three stages:

(1) In the first stage, the goal is to obtain a high-fidelity *P*-wave velocity model, robust to inaccurate *a priori* information. An initial population of *P*-wave profiles is generated within a search-window built around the semblance-derived interval velocities. Unlike a classic GA, which performs a purely random exploration of the model space, here each individual of the initial population results from the superposition of a random long-wavelength component and a high-frequency perturbation, defined according to the frequency band of the data, and the minimum expected velocity. For each model, the multi-channel seismogram is computed using a homogeneous-layer method (Schmidt & Jensen 1985; Schmidt & Tango 1986) and the value of the objective function is calculated. The following generation is then populated using a stochastic remainder criterium (Vardy 2015); all models with fitness better than average are propagated to the next generation and the remaining individuals are randomly selected among the entire population. The survivors are crossed-over and mutated according to user-defined crossover and mutation probabilities. The long- and short-wavelength components of the model are completely decoupled, with the crossover and mutation elements of the GA operating on each independently.

(2) After convergence is attained, the data are inverted for the elastic properties (Z, ν, ρ) . The genetic algorithm operates on the multi-parameter elastic space, preconditioned by the robust *P*-wave velocity model obtained in the previous stage. Shear wave velocity and density trends, which cannot be constrained independently by single-component marine seismic data alone (Operto *et al.* 2013), are derived from the long wavelengths of the *P*-wave velocity profile, by using appropriate rock-physics relationships (Hamilton 1970; Richardson & Briggs 1993; Mavko *et al.* 2009). By contrast, the high-frequency reflective component, of the elastic model is estimated from the offset-dependent seismic waveforms. The width of the search range is modulated by the local *P*-wave reflectivity; the higher the *P*-impedance contrast, the wider the range. The rationale is to bias the inversion towards regions where anomalies in the elastic properties are most probable: Poisson's ratio and density are expected to change where *P*-wave impedance changes (Hamilton 1970; Mavko *et al.* 2009), that is, at interfaces between sediment types. Thereby the effective size of the model space is selectively reduced, thus improving the efficiency in high-dimensional parametrizations (Tarantola 2005; Sajeva *et al.* 2016).

(3) Multiple independent runs are performed, each starting from an independent random population of models (Sen & Stoffa 1992; Vardy 2015). This produces an ensemble of statistically independent solutions, from which the model parameters error boundaries can be computed, without the bias of genetic drift and model interdependency (e.g. Vardy 2015). The process is iterated until the estimate of the median model is stable, as in Stoffa & Sen (1991). The aim of this process is to obtain a measure of the reproducibility of the solution rather than an importance sampling of the PPD, which is not achievable using a GA without further exploration of the model space (Stoffa & Sen 1991; Sambridge 1999; Aleardi & Mazzotti 2017).

In the next section, we provide experimental proof of the efficacy of the methodology on both synthetic and real data: the decoupled approach, is effective at retrieving the true *P*-wave velocity model even when it lies at the edges of the range, while the structural preconditioning improves the efficiency of the elastic inversion by reducing the size of the model space. Both synthetic data (Fig. 1) and a real case study are presented. We will hereafter refer to our proposed strategy as *Decomposed Genetic Algorithm* (DGA). In analogy with Bayesian methods, the search domain will be referred to as *a priori* probability density function (PDF).

3 IMPLEMENTATION AND RESULTS

3.1 Robust *P*-wave velocity estimation with constant density and shear properties

Here we compare the proposed DGA to a GA inversion with a traditional parametrization, for the reconstruction of an accurate *P*-wave velocity model starting from an ill-constructed *a priori* search domain. The data and the true elastic model are shown in Fig. 1. The search window is built around the interval *P*-wave velocity model derived from semblance velocity analysis (Dix 1955), in purple in Fig. 2, which represents an underestimate of the velocity gradient of about 5 per cent. Hence, the true model is significantly offset from the centre of the uniform *a priori* PDF.

The model is parametrized as a stack of 30 layers with a thickness of 25 cm each, down to a depth of 7-m below the seafloor. Only *P*-wave velocity is inverted for, whereas Poisson's ratio and density are kept fixed at this stage to a uniform profile ($\nu = 0.48$ and $\rho = 1.5\,\mathrm{g\,cm^{-3}}$). A population of 160 individuals sampled from the *a priori* PDF evolve through 150 generations according to an objective function $o(\mathbf{m})$ based on the L1 distance between the true data $\mathbf{r}(t)$ and the synthetic data $\mathbf{s}(t)$, computed for the model $\mathbf{m}$ (Menke 1989):

$$o(\mathbf{m}) = \sum_{j=1}^{M}\sum_{i=1}^{N} |r_{i,j} - s_{i,j}(\mathbf{m})| \tag{1}$$

,where M is the number of traces and N the number of time samples per trace. In this experiment, we compare the results obtained by computing the L1 residuals on (i) the real waveform and (ii) its instantaneous phase modified according to Jimenez-Tejero *et al.* (2015). This is the inverse tangent of the ratio between the absolute value of the Hilbert transform $\mathbf{h}(t)$ of the trace and the real waveform $\mathbf{w}(t)$:

$$\phi(t) = \tan^{-1}\frac{|\mathbf{h}(t)|}{\mathbf{w}(t)}. \tag{2}$$

The objective function computed on the the modified instantaneous phase (hereafter referred to as MIP) is independent from the waveform amplitude, and has proven in deterministic FWI to be beneficial for the inversion of data lacking low frequencies and acquired with short streamers (Jimenez-Tejero *et al.* 2015; Provenzano *et al.* 2017).

The DGA explores the model space by explicitly decomposing each random subsurface profile into a slowly-varying component, in the order of the dominant propagated wavelength ($\geq\lambda_{dom}$), and an independent high-frequency perturbation, in the order of the tuning thickness for the maximum frequency contained in the data ($\simeq\lambda_{min}/4$). This is achieved by generating two random series of chromosome values (i.e. *P*-wave velocity per layer), with correlation lengths, respectively, equal to λ_{dom} and $\lambda_{min}/4$, and superimposing them to create the broad-band subsurface model (Fig. 4). For the *P*-wave velocities and source bandwidth of this test, the correlation lengths of the short and long wavelength components are, respectively, 0.25 and 1.25 m (Table 1). The decomposed approach is taken in both the sampling of the starting random population from

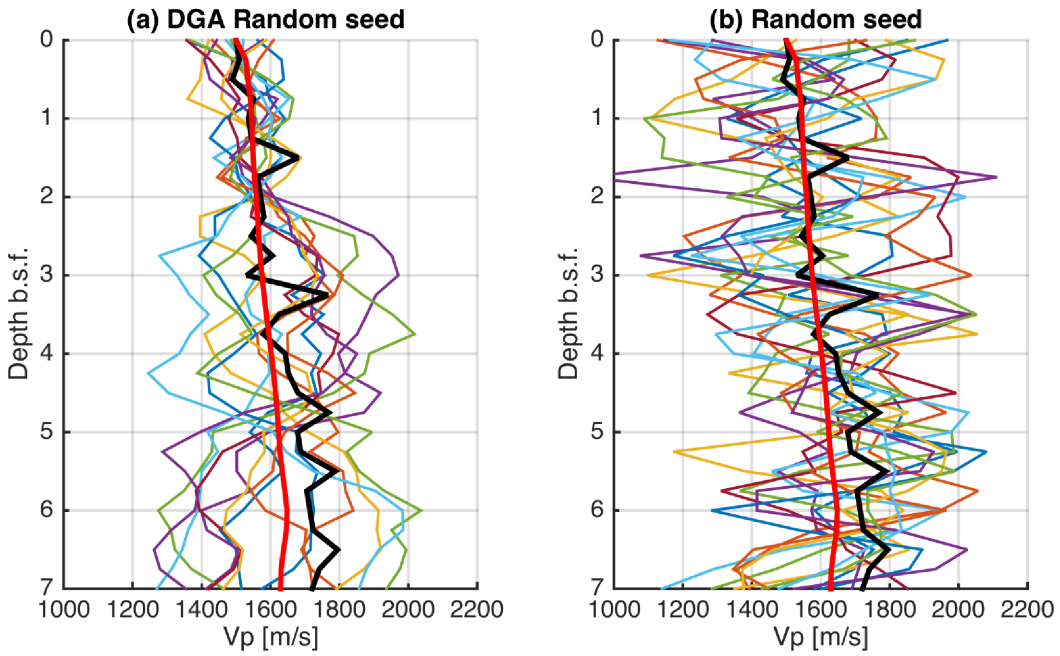

Figure 4. Random seeds. (a) Random models (thinner lines) generated by the decomposed algorithm within a search range designed around the semblance velocity model (thick red), and true V_p model (thick black). Correlation length for the high-frequency component is equal to the thickness of the individual layer (0.25 cm $\simeq \lambda_{min}/4$), and $\simeq$1 m for the long component. (b) The random models are instead generated by sampling randomly the same V_p search window.

Table 1. DGA inversion parameters. User-defined genetic algorithm parameters for the first-stage *P*-wave velocity inversion, and the following preconditioned simultaneous elastic inversion.

	One parameter (V_p)	Elastic (Z, ν, ρ)
L(m)	1.25	-
S(m)	0.25	0.25
Number of variables	30	90
Mutation prob.	0.2	0.001
Crossover prob.	0.4	0.6
Number of individuals	160	320
Number of generations	150	150
Number of runs	50	50

the search range, and in the mutation and crossover through the generations. High mutation rates and relatively low crossover probabilities are used to enhance the explorative nature of the algorithm, and prevent from earlier convergence to local minima (Sen & Stoffa 2013). The process thus accounts for the dual nature of the subsurface properties the seismic data are sensitive to, that is, kinematic and reflective (Mora 1980; Jannane 1989). By contrast, in a conventionally parametrized GA the reconstruction of the broadband earth model results from random mutation and crossover operating on each layer individually. Unlike the DGA, the model samples of the GA are high-variance profiles clustered around the centre of the search domain, and are in fact more likely to produce geologically unrealistic features (Fig. 4).

In Figs 5 and 6, we compare the median model and the 66 and 95 per cent solution confidence limits obtained from 50 independent inversion runs. The variation band of minimum and average L1 data misfit across the generations is also shown. The results show that the spectrally decomposed algorithm outperforms the conventional parametrization in attenuating the bias of the starting model: regardless the seismic attribute used, the true model is in fact included in the 95 per cent confidence limit of the DGA solution (grey shaded area in Figs 5 and 6). However, the median model of the DGA waveform-based inversion is offset from the true *P*-wave profile, and the reproducibility of the solution is poor (Fig. 5). The MIP-based misfit functional (Jimenez-Tejero *et al.* 2015) ensures more stable results (Fig. 6), proving a higher sensitivity to mid-to-long wavelength changes in the model, and lower liability to cycle skipping. In contrast, the waveform based objective function strongly suffers from local minima entrapment, because of the high non-linearity of the objective function with respect to the *P*-wave velocity trends (Virieux & Operto 2009; Operto *et al.* 2013), and the mid-wavelength sensitivity gap due to the limited bandwidth and short offset of the data (Jannane 1989). Therefore, the modified model exploration strategy is effective at attenuating the footprint of inaccurate *a priori* information, but only if combined with an appropriate objective function.

By decomposing the *P*-wave velocity model into independent long and short components, and using a misfit criterium robust to cycle skipping, the inaccuracy in the starting model has been compensated for. Also, the model heterogeneities have been correctly located. The next step is to use this result to precondition the high-resolution multi-parameter elastic inversion.

3.2 Structure-preconditioned elastic inversion

Here we exploit the information contained in the estimated *P*-wave velocity model, in order to guide the inversion for the elastic properties. The model is parametrized in terms of *P*-impedance (Z), Poisson's ratio (ν) and density (ρ) (e.g. Debski & Tarantola 1995; Igel *et al.* 1996; Provenzano *et al.* 2017). No independent constraint on the shear and density macro-model is posed by the single-component, limited offset data. So, in the absence of independent geological and geotechnical information, a sensible relationship with the *P*-wave velocity is needed to derive the elastic starting model. We make the assumption that such a relationship is available in our synthetic test, and is representative only of the long-wavelength component of the model. The high-wavenumber component, on

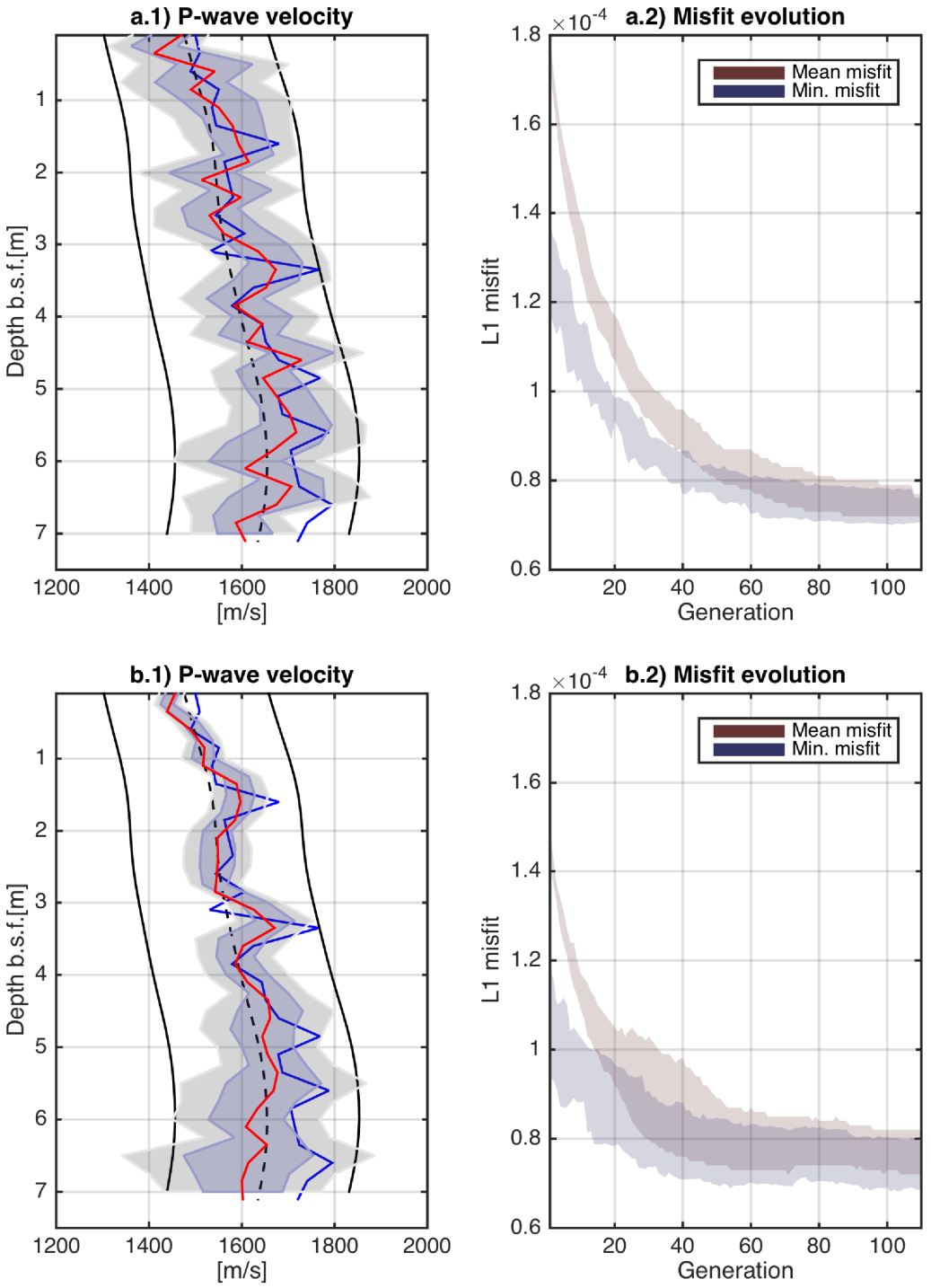

Figure 5. V_p inversion. Waveform L1 misfit functional. (a) Results obtained using a GA with a conventional strategy; (b) results obtained using our DGA. True model (blue), median model (red) and shaded areas for the 66 (blue) and 95 (grey) per cent confidence intervals of the solution are computed from an ensemble of 50 independent realizations. The variability range of the average and minimum misfit evolution over the generations is shown. The final minimum misfit is lower for the DGA inversion, and the true model is included in the confidence interval. However, the high variability of the solution reveals a remarkable liability to cycle skipping.

the other hand, varies independently for each parameter, simulating changes in lithology and/or fluid saturation (Igel *et al.* 1996; Mallick & Dutta 2002); such anomalies are the target of the elastic inversion, and are estimated entirely from the offset-dependent reflected wavefield.

The high-dimensionality of the multi-parameter problem determines a rapid increase of the model space volume with the range width, hence of the number of model evaluations needed to identify and sample its high-fitness region (Tarantola 2005; Sajeva *et al.* 2016, 2017). In our specific test, each elastic layer is described by a triad of independent parameters (Z, ν and ρ), for a total of 90

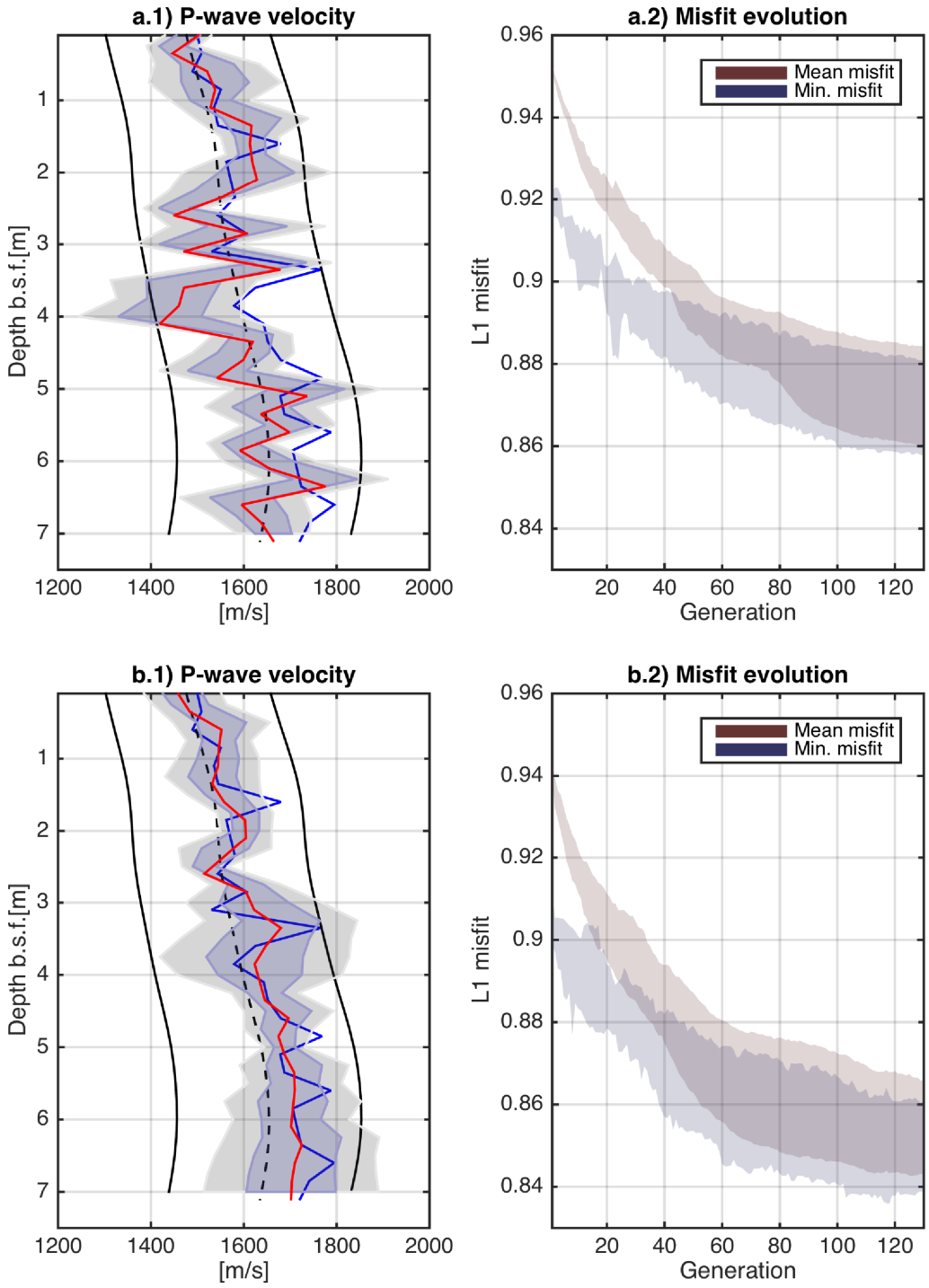

Figure 6. V_p inversion. Complex trace L1 misfit functional. (a) Results obtained using a GA with a conventional strategy; (b) results obtained using our DGA. True model (blue), median model (red) and shaded areas for the 66 (blue) and 95 (grey) per cent confidence intervals of the solution are computed from an ensemble of 50 independent realizations. The variability range of the average and minimum misfit evolution over the generations is shown. Note how the DGA solution (panel b) is less dependent from the central model (dashed line) of the *a priori* distribution, and that the final minimum misfit is lower than the traditional GA.

degrees of freedom. It can be therefore beneficial to reduce the volume of the model space by selectively narrowing the search domain. For this purpose, we further exploit the information contained in the *P*-wave velocity profile to precondition the elastic inversion. The exploration of the elastic model space is biased towards model heterogeneity locations identified by regions of contrasting *P*-wave velocity, from the results of the first inversion stage.

The preconditioned search range for either ν and ρ, for each depth, is compactly defined as:

$$R = F(V_{p\,sm}) \pm w_0 p \tag{3}$$

where

$$p = \frac{\left|\frac{\partial^2 V_\mathrm{p}}{\partial d^2}\right|}{\max\left(\left|\frac{\partial^2 V_\mathrm{p}}{\partial d^2}\right|\right)} \quad (4)$$

the search range R is thus obtained from the smooth P-wave velocity $V_{\mathrm{p\ sm}}$ using the rock-physics relationship F; the half-width w_0 can be modulated by the second derivative of V_p, computed with respect to the depth d, and normalized to one. In Fig. 7, we compare the state of information available from the moveout based interval acoustic velocities with the one from the FWI P-wave profile; the latter allows us to build a more accurate multiparameter search window, that reduces the risk of ruling out the true model from the *a priori* model distribution.

The inversion parameters are summarized in Table 1. Compared to the V_p-stage, we use a lower mutation rate, a higher crossover probability and a higher number of individuals per generation, to account for the higher number of unknowns (Stoffa & Sen 1991; Sen & Stoffa 1992). In Figs 8 and 9, we show the inversion results and the marginal confidence intervals for the elastic inversion, respectively, with and without range-width modulation. The confidence limits in both case are a high-fidelity representation of the elastic model, despite the low sensitivity of limited-offset marine data to shear properties variations, within the narrow Poisson's ratio range of shallow sediments (Mallick & Dutta 2002; Provenzano *et al.* 2017). However, the fully-preconditioned test, with range modulation, produces solution confidence regions with better defined shear and density heterogeneities, especially in the deeper parts of the model.

In Fig. 10, the synthetic seismogram for the final FWI-model is overlaid to the real one, and the offset-dependent L1 misfit is plotted for each DGA stage. It is worth pointing out that the V_p-model, although obtained via instantaneous phase inversion, is responsible for a significant reduction of the waveform misfit, especially at the far offsets. The final misfits for the elastic models obtained with the two preconditioning techniques are similar; despite the higher misfit in some of the farthest channels, the fully-preconditioned solution has a higher fitness value, thanks to the better match at the higher-energy shortest offsets.

With a view to reducing the size of the model space, the range-width preconditioning will be key in the inversion of the real case study, where the presence of a 50-cm thick gas-saturated layer requires the inversion to explore a wide range of elastic moduli values in a thick spatial parametrization.

3.3 Identification of shallow weak layers

3.3.1 Acquisition design and background information

Here we apply the elastic FWI to two UHF multi-channel seismic data sets acquired on a marine slope prone to shallow landsliding, in the Sørfjorden side-fjord near the town of Finneidfjord (Norway). The extensive suite of high frequency geophysical, geotechnical and geological data in the study area (e.g. Steiner *et al.* 2012) identifies a composite 50-cm thick clay-rich bed with low stiffness, low density and high overpressure ratio that lies at shallow depth within the background silty-clay sediments (Vanneste *et al.* 2012; Vardy *et al.* 2012; Vardy 2015). This layer has been recognized as the failure plane for multiple submarine landslides from the last few decades and is thus referred to as *event bed* (L'Heureux *et al.* 2012; Steiner *et al.* 2012; Vanneste *et al.* 2012, 2015; Vardy *et al.* 2012). The saturation state of the event bed changes across the basin, from a water-saturated zone in the north to a gas-bearing area to the south (Vardy *et al.* 2012; Morgan *et al.* 2014).

The seismic source is a Boomer wide-band electro-acoustic plate with an effective bandwidth spanning from 0.2 to 2.5 kHz. The far-field signature has been measured during acquisition, and used as source wavelet in the inversion. The acquisition system is a 60-channel streamer with 1.0 m group spacing and maximum offset equal to 72 metres. A single receiver group has a length of 1 m and is made up by seven elements, and both source and receiver group wavenumber filters have been included in the modelling

The inversion data sets at the two locations are supergathers obtained by stacking 15 adjacent common depth point (CDPs) gathers with 1 m spacing, in order to attenuate the random noise energy. Data pre-processing is limited to muting the direct arrival and bandpass frequency filtering in the desired modelling bandwidth (0.2–2.0 kHz). The receiver ghost reflections are a strong source of coherent noise in the data (Fig. 11), particularly at site A, where the seafloor ghost reflection has higher energy than the event bed reflection and the respective traveltimes intersect. We include the receiver ghosts in the forward model by integrating the 1-D solver with a custom-built frequency–wavenumber filter that accounts for the strong variation of the streamer depth and sea-surface reflection coefficient with offset (Provenzano *et al.* 2017).

Because of the availability of a reliable ground-truth, the challenging nature of the seismic data and the geohazard implications, this is an excellent case study to test the potential of UHF seismic inversion.

3.3.2 Elastic FWI and ground-truthing

We apply the inversion at two key locations, outside (site A) and inside (site B) the gas front (Fig. 11), with the aim to locate and characterise the event bed, and quantify the changes in partial gas saturation. P-wave velocity and bulk density measurements from a Multi Sensor Core Logger (MSCL) core proximal to site A (Fig. 11) provide the low-frequencies to build a reliable starting model; therefore the computational effort in this section will be devoted to the identification of geohazard-relevant features. The 1-D approximation is justified by the almost plane-parallel geometry of the shallow reflectors in the pre-stack depth-migrated (PSDM) image (Fig. 11). Preliminary analysis of the available reflection angle range suggests that, due to the short aperture of the data compared to the target depth, density and P-impedance cannot be resolved independently (Provenzano *et al.* 2017), therefore a full elastic (Z, ν, ρ) multi-parameter inversion would be heavily ill-posed. As shown in Igel *et al.* (1996) and Provenzano *et al.* (2016, 2017), a more appropriate parametrization in this case is (Z, ν).

Inversion parameters at site A are presented in Table 2. The relatively narrow range of expected values allowed for the simultaneous (Z, ν) inversion to converge to a satisfactory solution. The median model attains an excellent match with the MSCL acoustic impedance within the seismic resolution (Fig. 12), capturing the composite structure of the low-impedance anomaly between 3.5 and 4.0 m depth. The results suggest that the strongest heterogeneity at the event bed is the bottom interface, whereas the top one is probably a graded boundary at the seismic resolution. The Poisson's ratio model, on the other hand, does not contain important discontinuities, most probably because of the low-sensitivity of reflection data to changes in shear properties in the range of non-lithified sediments ($\geq$0.49; Hamilton 1970; Mallick & Dutta 2002; Provenzano *et al.*

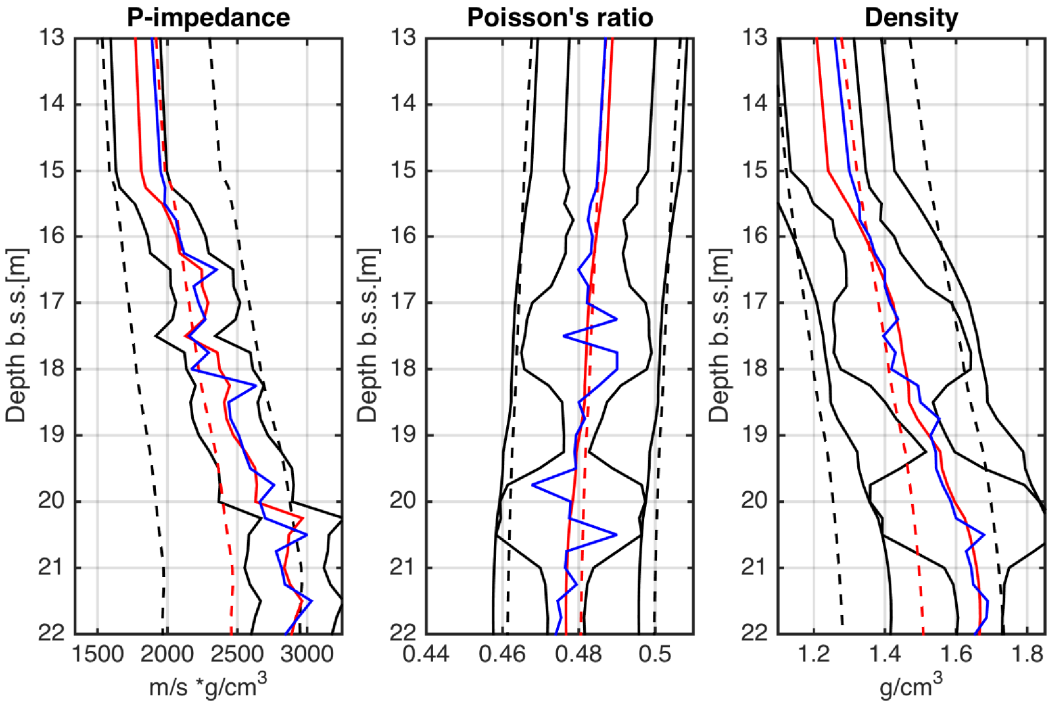

Figure 7. Advantage of using the *P*-wave velocity model to precondition the search window. Elastic model parametrized as (Z, ν, ρ). Note that the Z model obtained at the first stage of DGA (solid red), allows us to build accurate estimates of the shear and density trends (also in solid red). The derived ranges (solid black) contain the true model (blue). Using the semblance *P*-wave model with the same empirical relationship (dashed red), increases the chances of excluding the true model from the search window (dashed black). In this example, this is particularly evident for density.

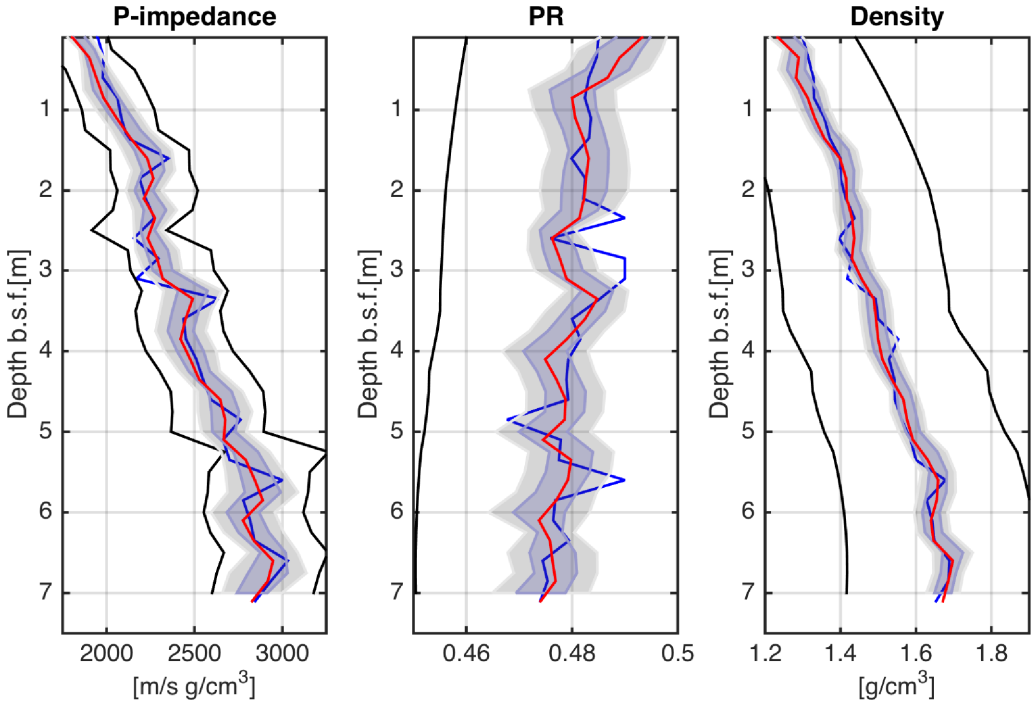

Figure 8. Elastic multiparameter inversion with long-wavelength preconditioning. Median model (red) and 66 (blue) and 95 (grey) per cent confidence intervals obtained from 50 independent inversion runs; true model in blue and range boundaries in solid black. In this test, the preconditioning on Poisson's ratio and density is only in the long-wavelength of the *P*-wave velocity model obtained in the previous stage.

2017). In Fig. 13, we compare the real seismogram to the computed one for the best-fit model; note that the small amplitude negative polarity reflection associated with the event bed is correctly represented in the inversion, despite the strong receiver ghost, which overlaps the up-going reflection within a significant range of offset. The small reduction of normalized offset-dependent misfit between the starting low-frequency model and the FWI-solution, is explained

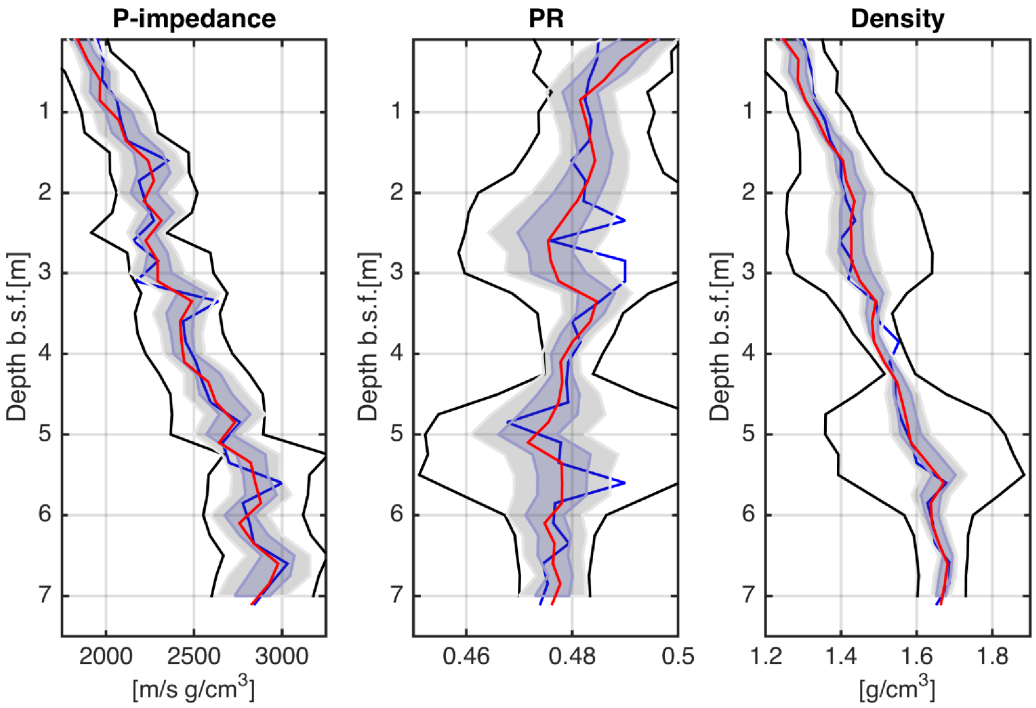

Figure 9. Elastic multiparameter inversion with full preconditioning. Median model (red) and 66 (blue) and 95 (grey) per cent confidence intervals obtained from 50 independent inversion runs; true model in blue and range boundaries in solid black. In this test, full preconditioning is used, and the width of the search range is modulated by the local *P*-wave velocity model heterogeneities obtained in the previous stage.

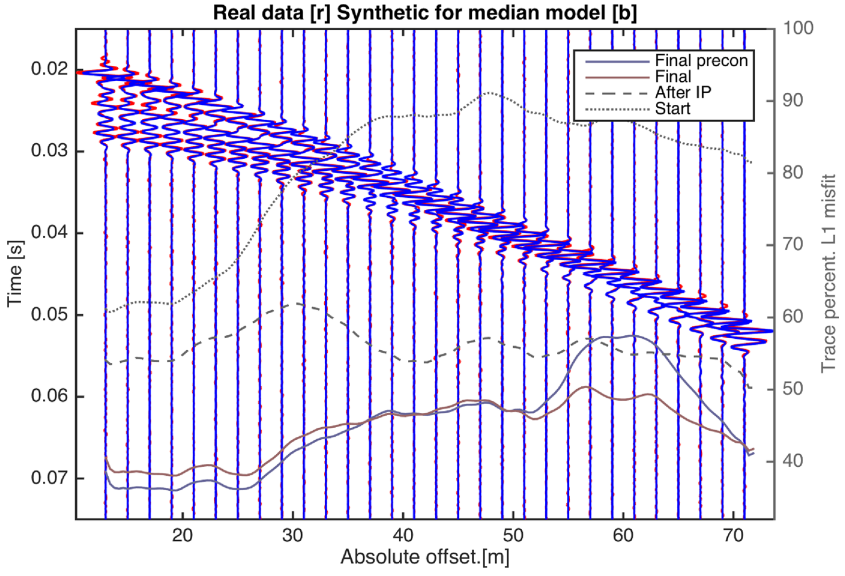

Figure 10. Two-stage elastic inversion offset-dependent misfit. Real seismogram (red), synthetic seismogram (blue) for the median model after elastic FWI. Overlaid, offset-dependent trace-normalized L1 misfit for the starting NMO model (dash-dot grey), the final V_p model (dash grey), and the two median models, with and without range-width preconditioning (respectively, blue and red).

by the low energy of the subsurface reflections compared to the primary and ghost seafloor reflections.

The inversion at site B is performed using the parameters summarized in Table 3. Because we want account for the possible presence of free gas, broader (Z, ν) search windows are required to allow for the inversion to converge close to the true elastic model; due to the gas saturation, *P*-impedance values close-to or lower-than the water column acoustic impedance are expected, and Poisson's ratio can span a broad range of values (Anderson & Hampton 1980a; Ostrander 1984; Tóth *et al.* 2014). Attempts to invert simultaneously Z and ν failed to converge to a meaningful solution, thus the two-stage preconditioned strategy is applied to reduce the size of the model space. The first stage of *P*-impedance inversion places the event bed between 7.5 and 8.0 m below the seafloor (b.s.f.); this model is then smoothed using a moving average filter with a correlation length of 50 cm ($\simeq\lambda/2$, Fig. 14), and used to precondition the range width of the elastic inversion according to eq. (3.2). The final model shows

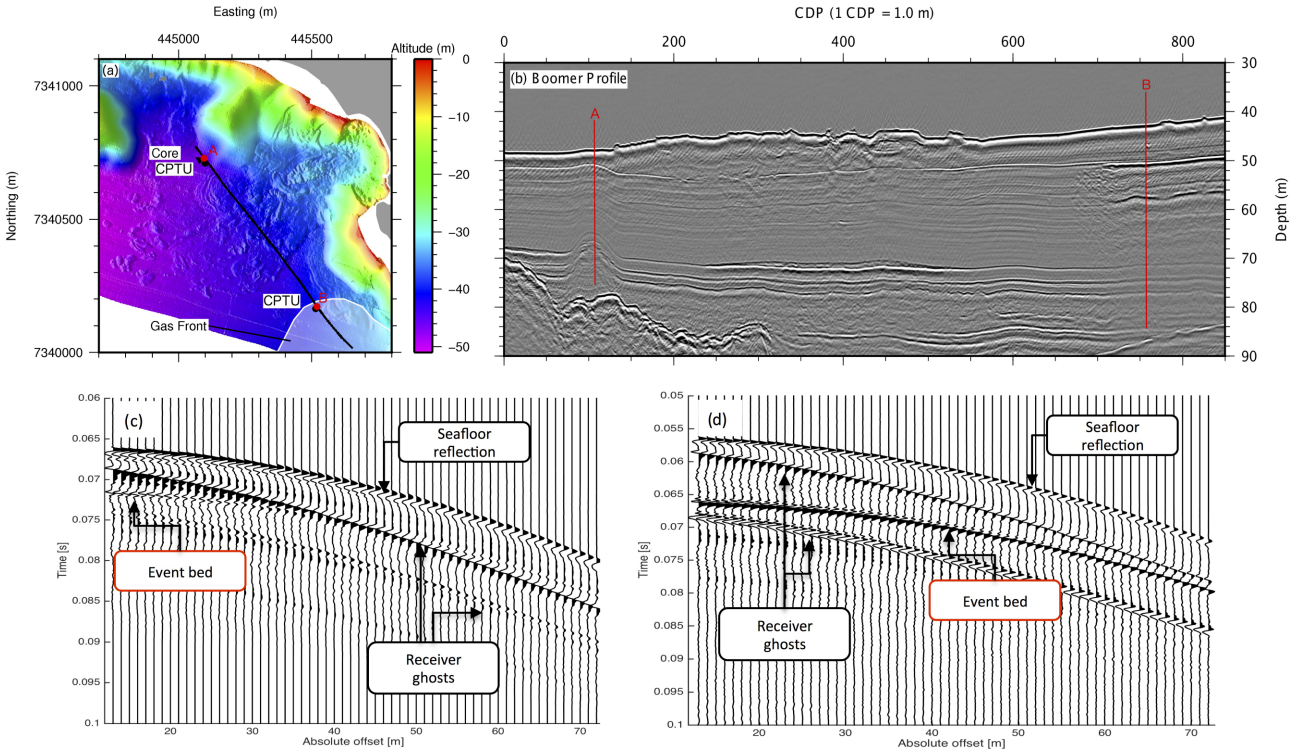

Figure 11. (a) Bathymetric image of the study area, with location of the two CDP supergathers (site A and B). The seabed morphology clearly shows indicators of multiple marine landslides deposits. The gas front area is shaded. (b) Migrated Boomer seismic line crossing the gas front. (c and d) Seismic gather, respectively, of sites A and B. Note how the strong receiver ghost reflections cross the event bed reflection in site A.

Table 2. Site A inversion parameters. User-defined genetic algorithm parameters for the simultaneous elastic inversion.

	One parameter (V_p)	Elastic (Z, ν)
S(m)	-	0.15
Number of variables	-	120
Mutation prob.	-	0.001
Crossover prob.	-	0.6
Number of individuals	-	420
Number of generations	-	100
Number of runs	-	50

a drop of Z at 7.4 m b.s.f., with a 95 per cent confidence interval reaching values under 600 m s^{-1} · g cm^{-3}, correlated to a decrease in ν. The confidence intervals for ν are broad, and the solution for Z shows a probably non-physical long wavelength harmonic trend, as a consequence of the limited sensitivity of the short-offset reflection data in a relatively deep-water environment (Mallick & Dutta 2002; Operto *et al.* 2013). Nevertheless, the FWI univocally identifies a shallow horizon with reduced bulk modulus, consistent with the presence of free gas in the pore space. The solution models after each stage account for a significant reduction of the offset dependent misfit in Fig. 15 compared to the initial state of information, and the final computed seismogram shows an excellent match with the real one.

3.3.3 *Can we quantify the free-gas content?*

As an example application, we use the FWI results to estimate a partial gas-saturation distribution in the sediment column. The presence of small amounts of free gas in the pore space is expected to produce a strong increase in the bulk sediment compressibility, having, on the other hand, little effect on its shear modulus (Mavko *et al.* 2009). Therefore, a drop of P-wave velocity associated with a decrease of Poisson's ratio is expected (Aki & Richards 2002). Also, gas-bubble resonance produce frequency-dependent changes in attenuation and P-wave velocity, which are significant at frequencies higher than the characteristic resonance frequency of the dominant bubble size (Anderson & Hampton 1980a; Tóth *et al.* 2015).

Under the assumption that bulk mechanical effects dominate in our frequency band (0.2–2.5 kHz; Riedel & Theilen 2001), we apply the geo-acoustic model by Anderson & Hampton (1980b) to estimate the partial gas saturation (Tóth *et al.* 2014). The model accounts for the increase of compressibility due to the presence of gas in the pore space, with respect to a water-saturated sediment at a given hydrostatic pressure. The model predicts the three-phase medium bulk modulus (K_m), and the shear modulus (μ), from which the (Z, ν) couple at each depth can be derived assuming an elastic isotropic mechanics (Mavko *et al.* 2009). The partial gas-saturation value is obtained by iteratively minimising the difference between the FWI-model and the predicted one.

Fig. 16 shows the results obtained at site B by fitting the FWI results within the confidence intervals, and compares the FWI-solution with the elastic properties predicted by the geo-acoustic model. A free-gas anomaly is placed at 7.5 m below the seafloor, with a partial gas saturation confidence area in the order of 0.03 per cent; this in good agreement with previous results obtained independently at the same site employing different high-resolution seismic sources with dissimilar inversion techniques (Vanneste *et al.* 2013; Morgan *et al.* 2014). The same procedure applied to site A resulted in a gas saturation one order of magnitude lower (Fig. 17).

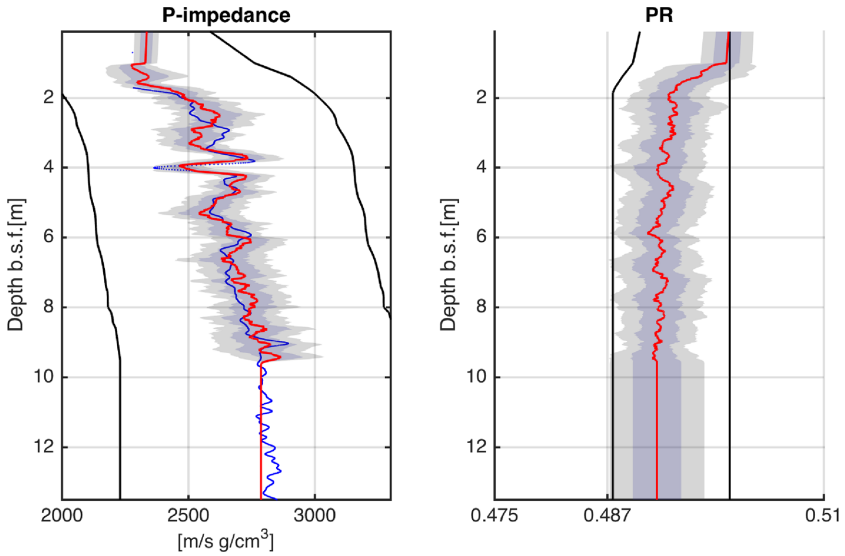

Figure 12. Site A elastic inversion results. Median (solid red) and 66–95 per cent confidence intervals (respectively blue and grey) obtained from 50 independent inversion runs; MSCL Impedance measured *in situ* (blue); search boundaries (solid black lines).

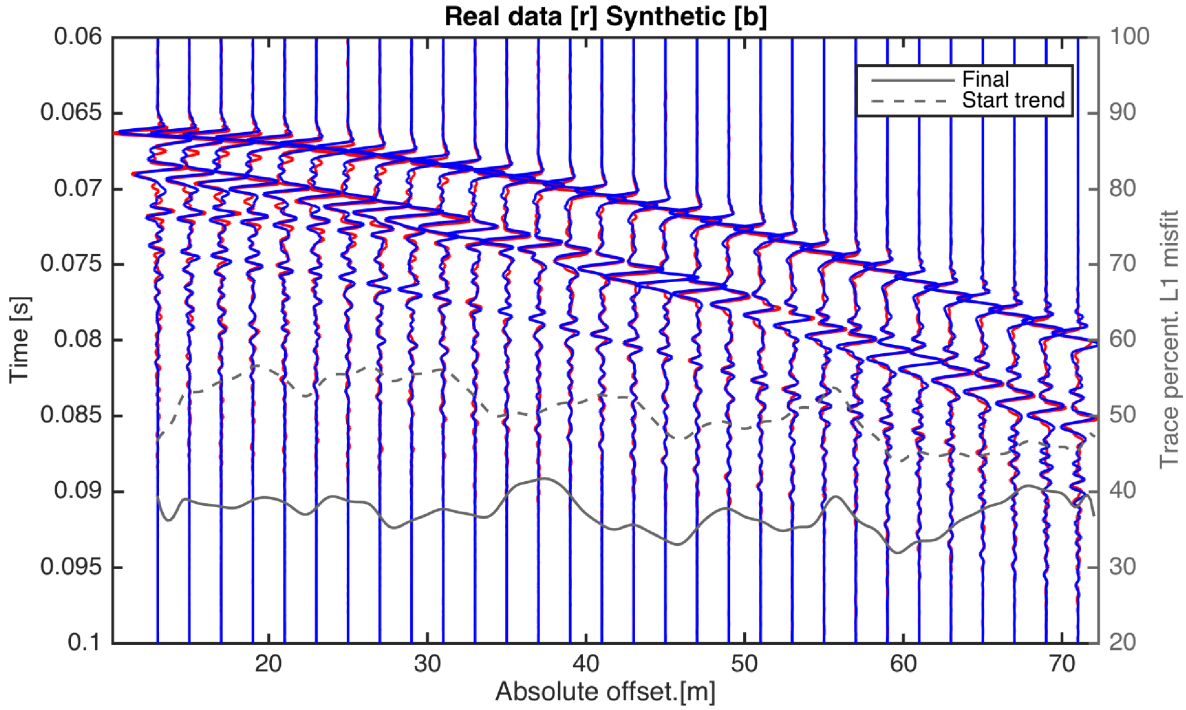

Figure 13. Site A elastic inversion offset-dependent misfit. Real seismogram (red), synthetic seismogram (blue) for the median model after elastic FWI. Overlaid, offset-dependent trace-normalized L1 misfit for the starting low-frequency model (dash-dot grey) and the final FWI model (solid grey).

Table 3. Site B inversion parameters. User-defined genetic algorithm parameters for the first-stage *P*-wave velocity inversion, and the following preconditioned simultaneous elastic inversion.

	One parameter (V_p)	Elastic (Z, ν)
S(m)	0.2	0.2
Number of variables	35	70
Mutation prob.	0.01	0.001
Crossover prob.	0.6	0.6
Number of individuals	320	500
Number of generations	120	150
Number of runs	10	50

3.4 Discussion

The real case study demonstrates stable results, in excellent agreement with the geotechnical ground-truth at the seismic resolution. This makes a strong case for the use of seismic data as a remote characterization tool for decimetric-scale features relevant to geohazard. In general, even when the final solution confidence intervals are broad, as in site B, the stochastic inversion succeeds at improving the *a priori* state of information about the subsurface, in a purely data-driven way. In Site A, the solution for *P*-impedance reproduces the MSCL V_p data, almost perfectly within the seismic resolution,

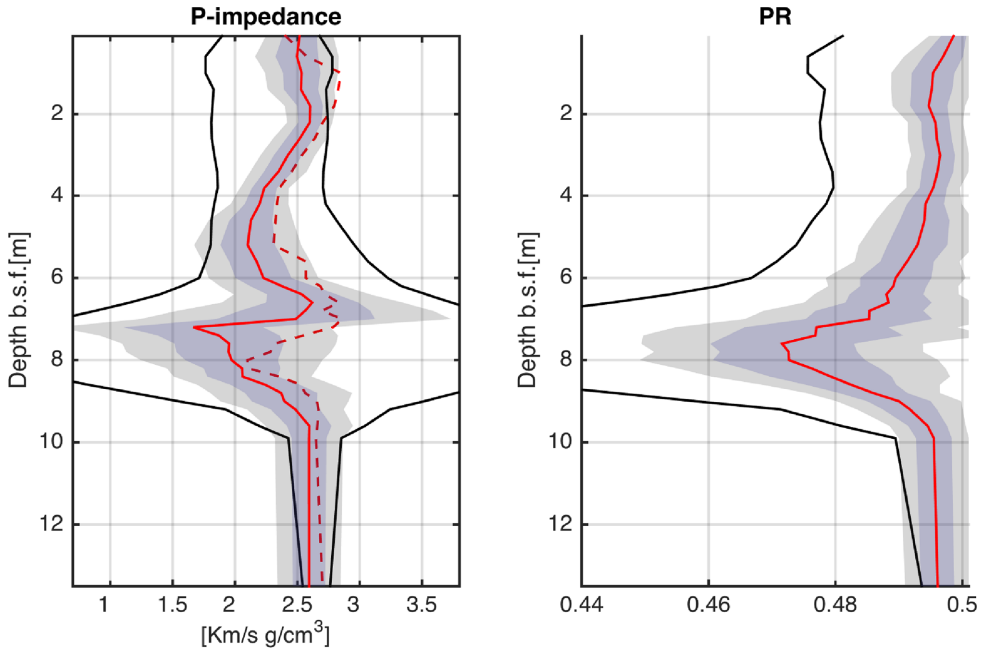

Figure 14. Site B elastic inversion results. Median model (solid red) and 66–95 per cent confidence intervals (respectively, blue and grey) obtained from 50 independent inversion runs; search boundaries (solid black lines); the red dashed line is the smoothed *P*-impedance model from the single-parameter inversion, whose local rate of change is used to precondition the width of the search range for the elastic inversion.

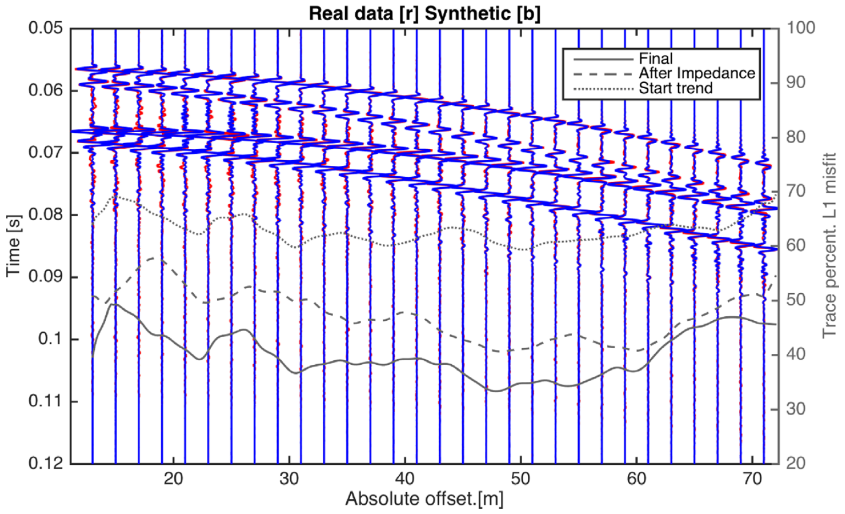

Figure 15. Site B elastic inversion offset-dependent misfit. Real seismogram (red), synthetic seismogram (blue) for the final model after elastic FWI. Overlaid, offset-dependent trace-normalized L1 misfit for the starting low-frequency model (dash-dot grey), the final V_p model (dash grey), and the final FWI model (solid grey).

constituting a virtual *in situ* elastic log. In site B, our structure-preconditioned strategy is fundamental to obtain convergence in the simultaneous multi-parameter inversion. The (Z, ν) model univocally identifies a shallow horizon showing a typical signature of gas saturation, as expected from the available independent data.

In the real data example, the availability of a *P*-wave velocity profile, measured on a core proximal to the inversion site, allowed us to design an accurate search range for the stochastic optimiser. When this is not the case, low-resolution *P*-wave models obtained from the reflection kinematics (e.g. from NMO analysis), can be used for the purpose. However, especially in sites with a strong shallow velocity gradient, inaccuracy in the semblance-derived starting model could bias the genetic algorithm away from the true broadband model. On the other hand, a uniform search range, between the maximum and the minimum expected velocities, significantly increases the number of samples required for an effective exploration of the model space (Tarantola 2005), therefore reducing the chances of convergence within a feasible computing time (*curse of dimensionality*, e.g.; Sajeva *et al.* 2016, 2017). Thus, we suggest that reflection moveout analysis can be used as a quick method to narrow the range of possible earth models, and that this should be combined with an appropriate seismic inversion strategy. In the synthetic example, the proposed decoupled parametrization of the *P*-wave velocity model, combined with a misfit criterium based on the instantaneous phase (Jimenez-Tejero *et al.* 2015), proved to be able to compensate for

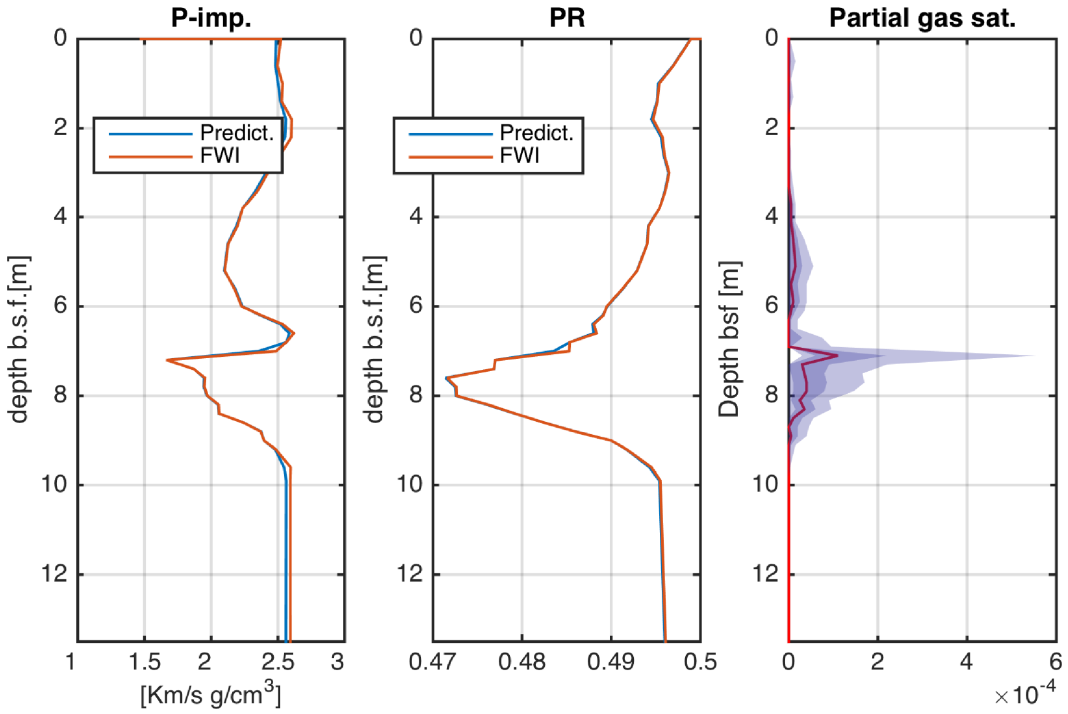

Figure 16. Site B. Partial gas saturation from Z and ν. P-impedance and Poisson's ratio predicted (blue) using Anderson & Hampton (1980b) model are compared with the FWI-solution (red). In the right panel, log-scale partial gas saturation estimated by fitting the median (red), and the 66 per cent (blue) and 95 per cent (grey) confidence limits.

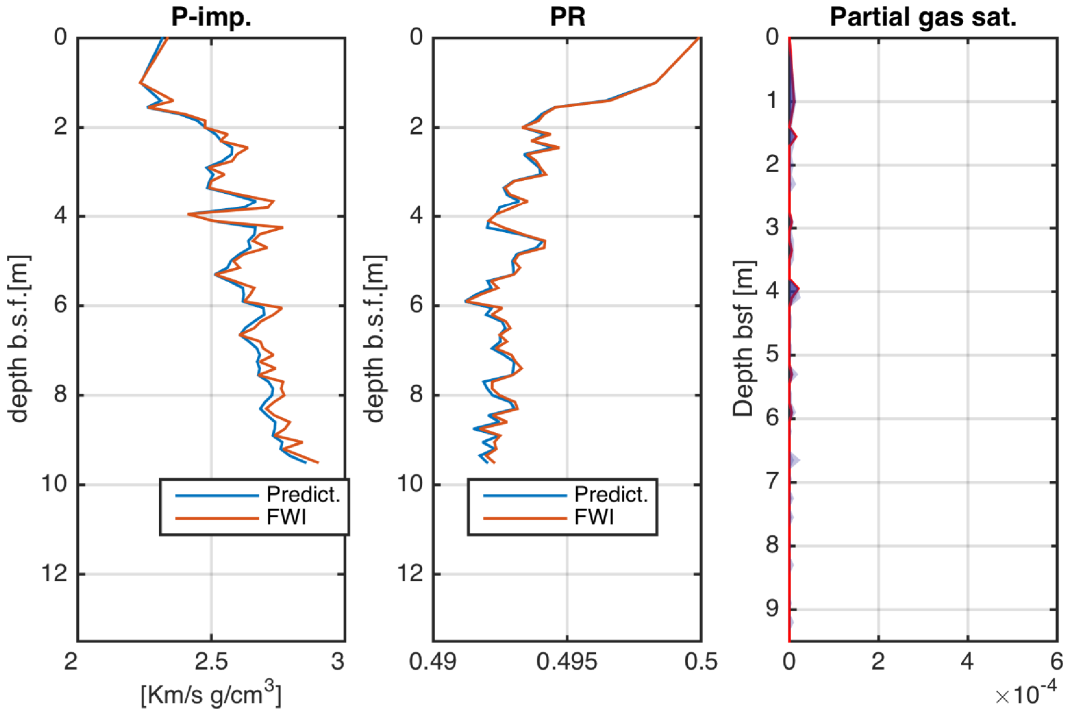

Figure 17. Site A. Partial gas saturation from Z and ν. P-impedance and Poisson's ratio predicted (blue) using Anderson & Hampton (1980b) model are compared with the FWI-solution (red). In the right panel, log-scale partial gas saturation estimated by fitting the median (red), and the 66 per cent (blue) and 95 per cent (grey) confidence limits. The same horizontal scale as in Fig. 16 is used keep consistency with the plot for site B.

ill-constructed search ranges, attenuating the bias of the starting model on the solution.

In the preconditioned elastic stage, a robust relationship among the long wavelength of the elastic parameters is only needed to build the search window, and can reflect the expected sediment type (Hamilton 1970; Richardson & Briggs 1993), or previous information about the recent geological history of the study area (e.g. Vardy *et al.* 2017). Detectable converted S-wave and post-critical

reflections would probably allow us to constrain independently the shear properties long wavelengths, but this is often not the case in short-offset marine data (Kormendi & Dietrich 1991; Igel *et al.* 1996; Vardy *et al.* 2017).

The elastic stage of reflection FWI aims at estimating the high-frequency elastic structures responsible for the offset-dependent seismic reflectivity, which are indicators of local anomalies in sediment fabric, lithology, or partial fluid saturation. The constraint posed by reflection data on the elastic parameters is highly hierarchical (Tarantola 1986; Igel *et al.* 1996), namely higher for *P*-wave impedance, whereas independent density anomalies can only be detected if wide reflection angles are available (>40°; Provenzano *et al.* 2017). Sensitivity to differential changes of *P* wave and *S* wave in non-lithified, water-saturated media is also low, because they span a narrow Poisson's ratio range, with little footprint on the AVO response (Mallick & Dutta 2002; Provenzano *et al.* 2017). This explains, in the synthetic example, the lower accuracy of the reconstructed shear and density profiles compared to the *P*-wave impedance result, and, in the real data, the absence of significant Poisson's ratio interfaces correlated to the event bed. By contrast, small amounts of gas in the pore space correspond to greater Poisson's ratio anomalies, as in site B (Ostrander 1984; Mallick & Dutta 2002).

The aforementioned sensitivity issues in the multi-parameter elastic space can in principle be quantitatively estimated from the posterior model ensemble as solution non-uniqueness. Several techniques have been developed in the literature in order to make the evolutionary optimiser a proxy to Bayesian statistics (e.g. Sen & Stoffa 1992, 1996; Sambridge 1999; Aleardi & Mazzotti 2017). However, it is beyond the scope of this paper to address the issue of an effective importance sampling of the model space using a genetic algorithm. The approach of this work is to perform multiple inversion runs, starting from independent random model populations (Sen & Stoffa 1992; Vardy 2015), in order to estimate the reproducibility of the solution without the bias of genetic drift and model inter-dependency (e.g. Vardy 2015). In the real case study elastic inversion, 50 independent runs were spread across 6 nodes; each run took 16 hr on a 16-core 2.6 GHz node, for a total computing cost of about 5 days. The synthetic inversion example, despite the two-stage approach, had a similar computing cost, because of the narrower bandwidth, the lower number of layers and the fewer unknowns.

In the real case study, the partial gas saturation boundaries estimated from the elastic model confidence intervals are consistent with previous independent geophysical studies (e.g. Vanneste *et al.* 2013; Morgan *et al.* 2014). While the gas saturations at this site have not been measured *in situ* (through pressure coring, or similar), the consistency amongst different geophysical methods is promising, especially because of the potential destabilising effect of gas-pockets within near-seafloor beds (e.g. Vanneste *et al.* 2013). However, similarly to the cited works, the compressible-gas model used (Anderson & Hampton 1980b) does not account for gas bubble resonance. As shown by Tóth *et al.* (2015), the validity of this assumption is highly dependent on the dominant gas bubble size, which exerts an important influence on apparent *P*-wave velocity dispersion and attenuation (Anderson & Hampton 1980a). Such frequency-dependent effects are negligible only at frequencies lower than the gas bubble characteristic frequency (Wilkens & Richardson 1998). Since the bubble size is not known to the accuracy required to rule out *P*-wave velocity dispersion at the upper end of the spectrum of our data ($\simeq$2.0−3.0 kHz), this approach could have underestimated the true gas saturation value, and should therefore be considered a lower bound.

4 CONCLUSIONS

We proposed a dedicated strategy for the pre-stack waveform inversion of ultra-high-frequency marine reflection data, based on a genetic algorithm with a carefully constructed model space, and demonstrated the potential of the method as a tool for the remote characterization of decimetric thickness layers.

In summary:

(1) We have demonstrated that a genetic algorithm optimization is not inherently robust against inaccurate starting models, and requires appropriate model parametrization and a careful choice of the objective function. A spectrally decoupled exploration of the model space, combined with a complex-trace based objective function has been shown to increase robustness against inaccurate *a priori* distribution, for example, derived from inaccurate pickings in reflection moveout analysis.

(2) We have also shown that this robust *P*-wave velocity model can be used to precondition the multi-parameter elastic inversion, in order to reduce the size of the model space in high-dimensionality parametrizations. This allows us to obtain a complex elastic model starting from little *a priori* information.

(3) A real case study has confirmed the potential of stochastic seismic inversion as a tool for the remote characterization of decimetric-scale structures useful for shallow geohazard assessment. A sedimentary bed correlated to the failure depth of multiple landslides in the study area has been identified. Signatures of changes in its partial gas saturation have been detected in the elastic model. Within the seismic resolution and sensitivity, an excellent match with the ground truth has been obtained.

The computing cost of this approach depends on the time used by a single forward model, and the number of forward models necessary for a satisfactory exploration of the model space. These factors still prevent stochastic optimization from being applied to 2-D/3-D elastic FWI, in which the number of independent parameters can be in the order of thousands. Nevertheless, a set of 1-D elastic inversions can be performed at selected locations, and used to inform the reconstruction of the lateral heterogeneity of the subsurface. Furthermore, such one-dimensional models can be interpolated to obtain a data-driven, accurate starting model for deterministic 2-D/3-D elastic FWI. The proposed strategy for engineering-scale FWI is thus a precious practical tool to complement information from bathymetry, sub-bottom profilers, cores and CPTUs for shallow geohazard assessment, reducing the need of expensive and time consuming geotechnical sampling campaigns in areas prone to shallow marine landsliding.

ACKNOWLEDGEMENTS

The work presented here was funded through the SEABED consortium with additional funding through the International Centre for Geohazards (ICG). The authors would like to thank Antonis Zervos, Tim Minshull and Hector Marin Moreno for the fruitful discussions regarding inversion and geomechanics. We also thank Rene-Edouard Plessix, Mrinal Sen and Andrea Tognarelli for the useful comments, that significantly improved the quality of the manuscript. The field data preconditioning was performed using a combination of Landmark's ProMAX software, MatLab and Seismic Unix. The

forward modelling was performed using the MIT OASES algorithm. All other processing used custom-written algorithms. The authors acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work.

REFERENCES

Aki, K. & Richards, P., 2002. *Quantitative Seismology,* University Science Books.

Aleardi, M. & Mazzotti, A., 2017. 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm-Gibbs sampler approach, *Geophys. Prospect.,* **65,** 64–85.

Aleardi, M., Tognarelli, A. & Mazzotti, A., 2016. Characterisation of shallow marine sediments using high-resolution velocity analysis and genetic-algorithm-driven 1D elastic full-waveform inversion, *Near Surf. Geophys.,* **14,** 449–460.

Anderson, A. & Hampton, L., 1980a. Acoustics of gas-bearing sediments I. Background, *J. acoust. Soc. Am.,* **67,** 1865–1889.

Anderson, A. & Hampton, L., 1980b. Acoustics of gas-bearing sediments II. Measurements and models, *J. acoust. Soc. Am.,* **67,** 1890–1903.

Cevatoglu, M., Bull, J., Vardy, M., Gernon, T., Wright, I. & Long, D., 2015. Gas migration pathways, controlling mechanisms and changes in sediment acoustic properties observed in a controlled sub-seabed CO2 release experiment, *Int. J. Greenhouse Gas Control,* **38,** 26–43.

Clare, M.A., Vardy, M.E., Cartigny, M., Talling, P., Himsworth, M., Dix, J., Harris, J. & Whitehouse, R., 2017. Direct monitoring of active geohazards: emerging geophysical tools for deep-water assessments, *Near Surf. Geophys.,* **15,** 427–444.

Dagnino, D., Salláres, V. & Ranero, C.R., 2014. Scale and parameter-adaptive model-based gradient pre-conditioner for elastic full-waveform inversion, *Geophys. J. Int.,* **198,** 1130–1142.

Debski, W. & Tarantola, A., 1995. Information on elastic parameters obtained from the amplitudes of reflected waves, *Geophysics,* **60,** 1426–1436.

Dix, C., 1955. Seismic velocities from surface measurements, *Geophysics,* **20,** 68–86.

Fichtner, A., 2011. *Full Seismic Waveform Modelling and Inversion,* Springer.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A. & Virieux, J., 2013a. Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 2: Synthetic and real data case studies from Valhall, *Geophysics,* **78,** 107–124.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J., 2013b. Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis, *Geophysics,* **78,** 81–105.

Goldberg, D., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning,* Addison Wesley Publishing Company.

Hamilton, E., 1970. Sound velocity and related properties of marine sediments, North Pacific, *J. geophys. Res.,* **75,** 4423–446.

Igel, H., Djikpesse, H. & Tarantola, A., 1996. Waveform inversion of marine reflection seismograms for P-impedance and Poisson's ratio, *Geophys. J. Int.,* **124,** 363–371.

Jannane, M., 1989. Short Note. Wavelength of earth structures that can be resolved from seismic reflection data, *Geophysics,* **54**(7), 906–910.

Jimenez-Tejero, C., Dagnino, D., Salláres, V. & Ranero, C., 2015. Comparative study of objective functions to overcome noise and bandwidth limitations in full waveform inversion, *Geophys. J. Int.,* **203,** 632–645.

Kennedy, J. & Eberhart, R., 1995. Particle swarm optimization, in *Proceedings of IEEE International Conference Neural Networks,* Vol. **4,** pp. 1942–1948, IEEE, Perth, WA, Australia.

Kormendi, F. & Dietrich, M., 1991. Nonlinear waveform inversion of plane-wave seismograms in stratified elastic media, *Geophysics,* **56**(5), 664–674.

L'Heureux, J.-S. *et al.*, 2012. Identification of weak layers and their role for the stability of slopes at Finneidfjord, northern Norway, in: *Submarine Mass Movements and their Consequences. Advances in Natural and Technological Hazards Research,* Vol. **31,** pp. 321–330, Springer, Dordrecht.

Vardy , M.E., Henstock, T.J., Clare, M.A., Forsberg, C.F. & Provenzano, G., 2017. State-of-the-art remote characterization of shallow marine sediments: the road to a fully integrated solution , *Near Surf. Geophys.,* **15,** 387–402.

Mallick, S. & Adhikari, S., 2015. Amplitude variation with offset and pre-stack waveform inversion: a direct comparison using a real data example from the Rock Springs Uplift, Wyoming, USA, *Geophysics,* **80**(2), B45–B59.

Mallick, S. & Dutta, N., 2002. Shallow water flow prediction using prestack waveform inversion of conventional 3D seismic data and rock modeling, *The Leading Edge,* **21,** 675–680.

Mavko, G., Mukerji, T. & Dvorkin, J., 2009. *The Rock Physics Handbook,* Cambridge Univ. Press.

Menke, W., 1989. *Geophysical Data Analysis and Inverse Problems,* Academic Press INC.

Monrigal, O., Jong, I.D. & Duarte, H., 2017. An ultra-high-resolution 3D marine seismic system for detailed site investigation, *Near Surf. Geophys.,* **15,** 335–345.

Mora, P., 1980. Inversion = migration + tomography, *Geophysics,* **54**(12), 1575–1586.

Morgan, E., Vanneste, M. & Vardy, M., 2014. Characterization of the slope-destabilizing effects of gas-charged sediment via seismic surveys, *Offshore Technology Conference,* Houston, TX, USA.

Operto, S., Prieux, Y., Ribodetti, A., Brossier, R., Metivier, L. & Virieux, J., 2013, **32**. A guided tour of multiparameter full waveform inversion with multicomponent data, *Leading Edge,* pp. 936–947.

Ostrander, W., 1984. Plane-wave reflection coefficients for gas sands at non-normal angles of incidence, *Geophysics,* , **49,** (10), 1637–1648.

Pinson, L., Henstock, T., Dix, J. & Bull, J., 2008. Estimating quality factor and mean grain size of sediment from high-resolution seismic data, *Geophysics,* **73,** G19–G28.

Provenzano, G., Vardy, M.E. & Henstock, T., 2016. Pre-stack waveform inversion of VHF marine seismic reflection data. A case study in Norway, in *Near Surface Geoscience 2016 – Second Applied Shallow Marine Geophysics Conference.,* Barcelona, Spain, 4-8 September.

Provenzano, G., Vardy, M.E. & Henstock, T.J., 2017. Pre-stack full waveform inversion of ultra-high-frequency marine seismic reflection data, *Geophys. J. Int.,* **209,** 1593–1611.

Richardson, M. & Briggs, K., 1993. On the use of acoustic impedance values to determine sediment properties, in *Acoustic classification and mapping of the seabed,* **15,** part 2, Proceedings of the Institute of Acoustics, 15–24.

Riedel, M. & Theilen, F., 2001. AVO investigations of shallow marine sediments, *Geophys. Prospect.,* **49,** 198–212.

Rothman, D., 1985. Nonlinear inversion, statistical mechanics, and residual statics esiimation, *Geophysics,* **50,** 2784–2796.

Ruthenford, S. & Williams, R., 1989. Amplitude-versus-offset variations in gas sands, *Geophysics,* **54**(6), 680–688.

Sajeva, A., Aleardi, M., Stucchi, E. & Mazzotti, A., 2016. Estimation of acoustic macro models using a genetic full-waveform inversion: Applications to the marmousi model, *Geophysics,* **81,** 173–184.

Sajeva, A., Aleardi, M., Galluzzi, B., Stucchi, E., Spadavecchia, E. & Mazzotti, A., 2017. Comparing the performances of four stochastic optimisation methods using analytic objective functions, 1D elastic full-waveform inversion, and residual static computation, *Geophys. Prospect.,* **65,** 322–346.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm II. Appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Sambridge, M. & Mosegaard, K., 2002. Monte carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 1–29.

Schmidt, H. & Jensen, F., 1985. A full wave solution for propagation in multilayered viscoelastic media with application to Gaussian beam reflection at fluid-solid interfaces., *J. acoust. Soc. Am.,* **77,** 813–825

Schmidt, H. & Tango, G., 1986. Efficient global matrix approach to the computation of synthetic seismograms., *Geophys. J. Res. Astron. Soc*.

Sen, M. & Stoffa, P., 1992. Rapid sampling of model space using genetic algorithms: examples from seismic waveform inversion, *Geophys. J. Int.,* **108,** 281–292.

Sen, M. & Stoffa, P., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, *Geophys. Prospect.,* **44,** 313–350.

Sen, M. & Stoffa, P., 2013. *Global Optimization Methods in Geophysical Inversion,* Cambridge Univ. Press.

Steiner, A., L'Heureux, J.-S., Kopf, A., Vanneste, M., Longva, O. & Lange, M., 2012. *An In-Situ Free-Fall Piezocone Penetrometer for Characterizing Soft and Sensitive Clays at Finneidfjord (Northern Norway), in Submarine Mass Movements and Their Consequences,* pp. 99–109, Yamada, Y. , *et al.*, Springer.

Stoffa, P. & Sen, M., 1991. Nonlinear multiparameter optimization using genetic algorithms: inversion of plane wave seismograms, *Geophysics,* **56,** 1794–1810.

Stoker, M., Bradwell, T., Howe, J., Wilkinson, I. & McIntyre, K., 2009. Lateglacial ice-cap dynamics in NW Scotland: evidence from the fjords of the Summer Isles region., *Quat. Sci. Rev.,* **28,** 3161–3184.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49**(8), 1259–1266.

Tarantola, A., 1986. A strategy for nonlinear inversion of seismic reflection data, *Geophysics,* **51**(10), 1893–1903.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* Society for Industrial and Applied Mathematics.

Tóth, Z., Spiess, V., Mogollón, J.M. & Jensen, J., 2014. Estimanting the free gas content in Baltic Sea sediments using compressional wave velocity from marine seismic data, *J. geophys. Res.,* **119,** 1–17.

Tóth, Z., Spiess, V. & Keil, H., 2015. Frequency dependence in seismoacoustic imaging of shallow free gas due to gas bubble resonance, *J, geophys. Res.,* **120,** 8056–8072.

Vanneste, M., *et al.*, 2012. Assessing offshore geohazards: a multidisciplinary reasearch intiative to understand shallow landslides and their dynamics in coastal and deepwater environments, Norway, *Adv. Nat. Technol. Hazard Res.,* **31,** 29–41.

Vanneste, M. *et al.*, 2013. Finneidfjord, a field laboratory for integrated submarine slope stability assessments and characterization of landslide-prone sediments: a review, in *Offshore Technology Conference held in Houston, Texas, USA - OTC 13OTC-P-686-OTC*.

Vanneste, M. *et al.*, 2015. Integration of very-high-resolution seismic and CPTU data from a coastal area affected by shallow landsliding—the Finneidfjord natural laboratory, in *Proceedings of the Third International Symposium on Frontiers in Offshore Geotechnics,* OTC-TC-P-686.

Vardy, M., 2015. Deriving shallow-water sediment properties using post-stack acoustic impedance inversion, *Near Surf. Geophys.,* **13,** 143–154.

Vardy, M., L'Heureux, J.-S., Vanneste, M., Longva, O., Steiner, A., Forsber, C. &Brendryen J., 2012. Multidisciplinary investigation of a shallow near-shore landslide, Finneidfjord, Norway, *Near Surf. Geophys.,* **10,** 267–277.

Vardy, M., Vanneste, M., Henstock, T.J., Morgan, E. & Pinson, L., 2015. Can high-resolution marine geophysical data be inverted for soil properties? *Proc. Instit. Acoust,* **37**(1), 149–156.

Vardy, M., Vanneste, M., Henstock, T.J., Clare, M.A., Forsberg, C. & Provenzano, G., 2017. State-of-the-art remote characterization of shallow marine sediments: the road to a fully integrated solution, *Near Surf. Geophys.,* **15,** 387–402.

Verbeek, N. & McGee, T.M., 1995. Characteristics of high-resolution marine reflection profiling sources, *J. appl. Geophys.,* **33,** 251–269.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics., *Geophysics,* **74**(6).

Wilkens, R. & Richardson, M., 1998. The influence of gas bubbles on sediment acoustic properties: in situ, laboratory, and theoretical results from Eckernförde Bay, Baltic Sea, *Cont. Shelf Res.,* **18,** 1859–1892.