

1 Efficient sequential geoacoustic inversion: algorithm and  
2 application

3 Mandolesi E<sup>1</sup>, Dettmer J<sup>1,2</sup>, Dosso S<sup>1</sup>, Holland C<sup>3</sup>, and Martinelli S<sup>3</sup>

4 <sup>1</sup>School of Earth and Ocean Sciences, University of Victoria Victoria BC Canada

5 <sup>2</sup>Department of Geosciences, University of Calgary Calgary BC Canada

6 <sup>3</sup>Applied Research Laboratory, Pennsylvania State University State College PA, USA

7 October 6, 2017

8 **Abstract**

9

10 **1 Introduction**

11 Trans-dimensional (trans-D) Bayesian inversion has received much attention in recent years as tool to  
12 estimate physical parameters that describe the environment as well as their uncertainties. Bayesian  
13 inference is a technique derived from Bayes' theorem and generated a greater and greater interest with  
14 the development of computer technology. Bayesian inference has many possible uses in geophysics explo-  
15 ration, solving, at least from a theoretical perspective, the problem of model selection (i.e. parametrisa-  
16 tion). However, with respect to standard linearized inversion methods, Bayesian inference results to be  
17 particularly expensive in terms of computation time, especially for complex inverse problems which aim  
18 to account for many datasets at once. One way to account for large data volumes is to consider Bayesian  
19 particle filters. For example, Yardim et al. [1] proposed three different types of filters to track spatial  
20 and temporal geoacoustic environmental parameters, and more recently Dettmer et al. [2] proposed a  
21 particle filter based on annealed importance sampling (AIS) to estimates range-dependent seabed param-  
22 eters from reflection-coefficient data collected along a track. Notwithstanding improvements in computer  
23 power, enhancements in the efficiency of available methods are of great interest to a broad community  
24 of scientists who aim to handle larger and larger datasets with the best possible accuracy and speed. In  
25 this paper a novel sequential Bayesian method to invert a large number of sequential reflection-coefficient  
26 data sets from surveys in shallow water is presented.

27 The algorithm presented here is thus not truly a particle filter, but it uses concepts and techniques  
 28 from particle filters. On the basis of the algorithm developed here, inversions of both simulated and  
 29 measured large datasets are carried out. Simulation analysis show that the proposed sampler is effective  
 30 and greatly reduces computation time to estimate full uncertainty distributions along the track. Without  
 31 such efficiency increment, it will not be possible to invert a large dataset as the one collected as part of  
 32 the Clutter09 experiment and presented here.

## 33 **2 Experiment and Modeling**

34 Ocean acoustic data were collected along a 14-km transect in May, 2009, on the outer shelf of the Malta  
 35 Plateau, south of Sicily in the Mediterranean Sea, as part of the Clutter 09 experiment (Fig. 1). The  
 36 bathymetry varied smoothly between  $\sim$ 130–140 m along the track. The geometry for the experiment is  
 37 illustrated in Fig. 2. The AUV travelled at a speed of 2 kn (1.03 m/s) and a height of about 12 m above  
 38 the seabed, towing an acoustic source—an Ultra Electronics 2-100 MPS cylindrical projector mounted  
 39 in a spheroidal-shaped tow body—and a horizontal array of 32 hydrophones. The distance between the  
 40 AUV and the source was 2.6 m and the first hydrophone was 10.38 m from the source; the hydrophones  
 were spaced at 1.05 m with a total array length of 32.55 m.

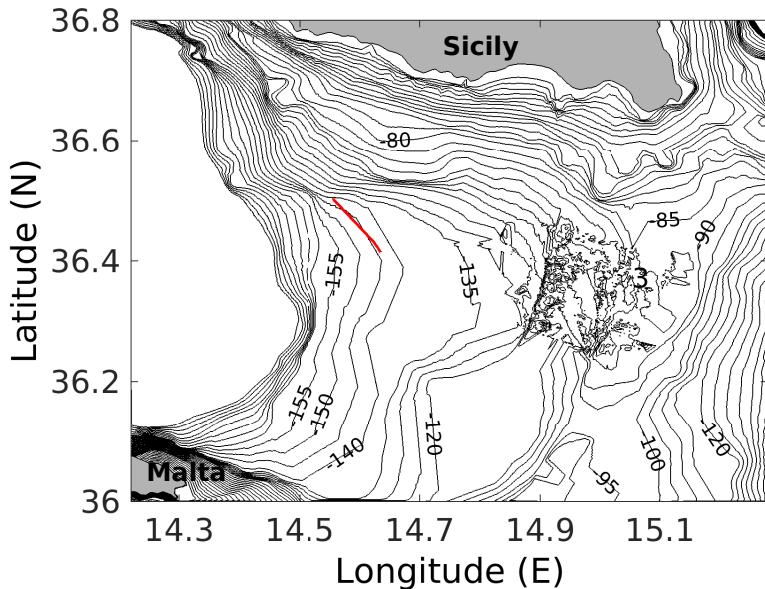


Figure 1: Bathymetric map of the survey area. Grey color depicts land masses. The red line shows the AUV trajectory.

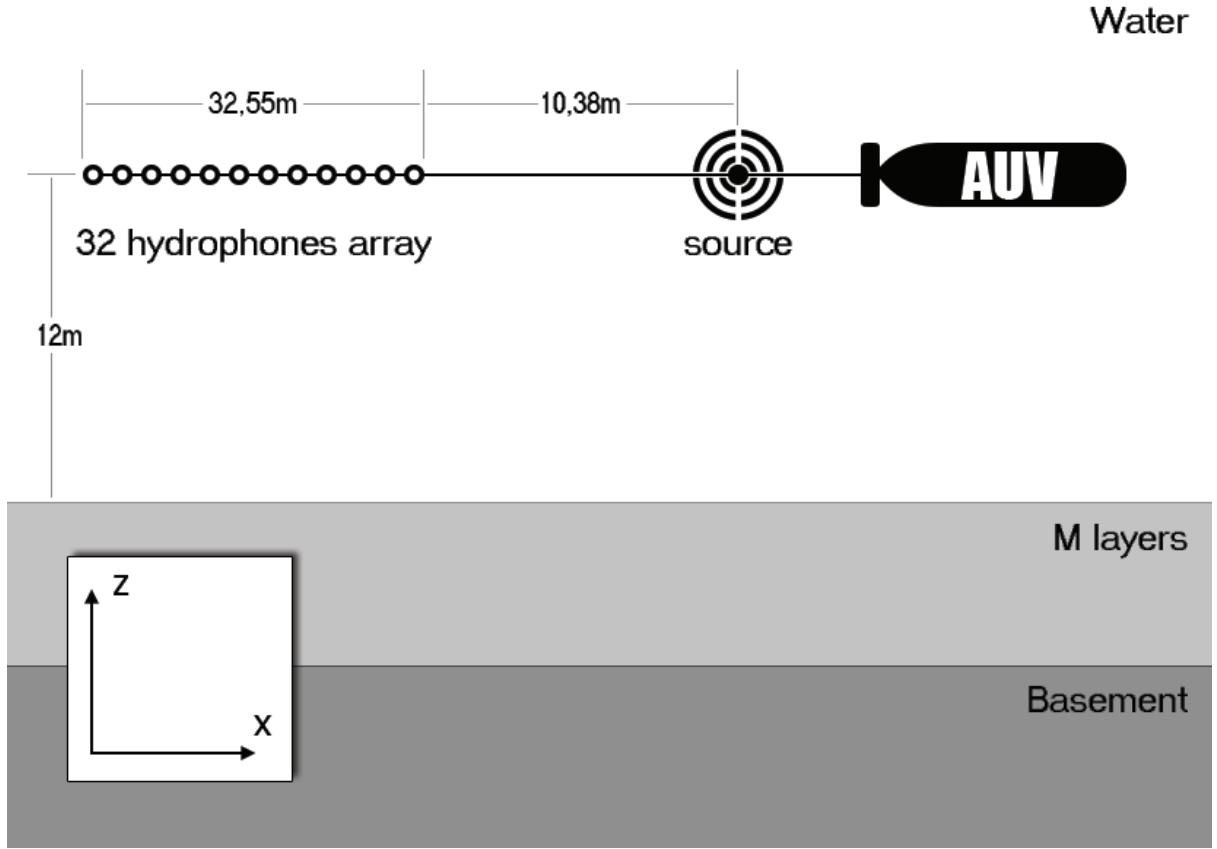


Figure 2: Schematic diagram of the experiment.[18]

41  
 42     The source generated a 1-s duration linear frequency modulated pulse that swept from 800 to 3500 Hz  
 43     (referred to as a “ping”) every 3 s, which was recorded by the receiver array. The recorded time series  
 44     were time-windowed to separate the direct and the bottom-reflected arrivals and Fourier transformed.  
 45     The ratio between the spectra of direct and the bottom-reflected arrivals, corrected for source directivity  
 46     and geometric spreading, defines the seabed reflection coefficient in the frequency domain [3]. Given  
 47     the source-receiver geometry, the reflection coefficients were computed for seabed grazing angle of 29°–  
 48     67°. Reflection coefficients were averaged in six 25-Hz wide frequency bands with band centers at 988,  
 49     1113, 1288, 1913, 2263, and 2513 Hz. To improve stability, the frequency-averaged reflection coefficients  
 50     were stacked (averaged) spatially over groups of 10 consecutive pings and assigned the source-receiver  
 51     range of the central ping, which results in an along-track seabed acoustic footprint of ∼ 20 m for the  
 52     averaged data. This procedure yielded a total of 1487 sequential data sets along the track. As part of  
 53     the stacking procedure, the standard deviation of the stack-averaged reflection coefficients was computed  
 54     for every frequency band and grazing angle. The data uncertainties estimated this way represents only  
 55     the internal consistency of the data within the stack, and a small number of highly-inconsistent pings  
 56     were removed from the data set. Moreover, to account for unknown error processes effecting the data  
 57     (both measurement and theory errors), the standard deviations were multiplied by scale factors of 1.2

58 for the low-frequency bands (988, 1113, and 1288 Hz) and 1.4 for the high-frequency bands (1913, 2263,  
 59 and 2513 Hz). These scale factors were determined by a trial-and-error procedure such that inversions  
 60 (described in Sec. 3) could generally fit the sequential data sets to statistically-acceptable levels.

61 The geometry of the experiment, in particular the distance of the source and array from the seabed,  
 62 meant that acoustic wave-front curvature was non-negligible, so, to invert these data, a procedure to  
 63 efficiently model spherical-wave reflection coefficients (rather than simpler plane-wave reflections coeffi-  
 64 cients) is applied. This procedure considers spherical waves as a superposition of plane waves incident  
 65 on the seabed and integrates over all plane-wave angles leading to the Sommerfeld integral for the  
 66 spherical-wave reflection coefficient,  $R_s$ ,

$$R_s(\theta, \omega) = G_\omega \int_0^\infty \frac{R_p(k_z, \omega) k_z}{k_r} e^{ik_r z_t} J_0(r k_z) dk_z, \quad (1)$$

67 where  $\theta$  is the seabed grazing angle,  $\omega$  is the angular frequency,  $R_p$  is the plane-wave reflection coefficient,  
 68  $k_r = k \cos \phi$  and  $k_z = k \sin \phi$  are the horizontal and vertical wavenumbers, respectively,  $k = \omega/c_w$  is the  
 69 wavenumber for water sound speed  $c_w$ ,  $z_t = 2H - (z - z_s)$ ,  $H$  is the water depth,  $z_s$  is the source depth,  
 70  $z$  is the receiver depth,  $J_0(\cdot)$  is the zeroth-order Bessel function of the first kind,  $r$  is the horizontal  
 71 separation between source and receiver,  $G_\omega = ir_t e^{-ikr_t}$ ,  $r_t = \sqrt{r^2 + z_t^2}$ , and  $i$  is the imaginary unit [4]  
 72 (see the measurement geometry in Fig. 3). The integrand in Eq. (1) depends on the seabed geoacoustic  
 73 parameters only through the plane-wave reflection coefficient  $R_p$ , which can be computed by standard  
 74 recursive algorithms [5]. The Bessel functions are known to be highly oscillatory, so the numerical  
 75 evaluation of  $R_s$  can be computationally expensive. In this work we use the efficient algorithm based on  
 76 Levin integration and the hybrid CPU/GPU parallel implementation developed by Quijano et al. [4] to  
 77 predict spherical reflection-coefficients. The modeled reflection-coefficient data are frequency-averaged  
 similarly to the measured data.

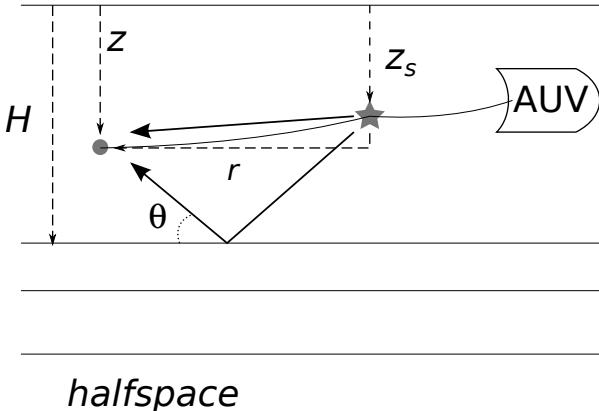


Figure 3: Spherical reflection-coefficients modeling scheme. The star represents the source, the grey dot represents one hydrophone in the array.  $H$  is the water depth and a set of seabed layers divides the sediment/water interface from the half-space that represents the basement.

<sup>79</sup> **3 Inverse theory and algorithms**

<sup>80</sup> This section summarizes the theory and algorithms used to invert large sequential reflectivity data sets.

<sup>81</sup> **3.1 Trans-D Bayesian inversion**

<sup>82</sup> Let  $\mathbf{m}$  be a vector of  $M$  model parameters (considered random variables) and  $\mathbf{d}$  be a vector of  $N$  data;

<sup>83</sup> Bayes' theorem states

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{m}) P(\mathbf{d}|\mathbf{m})}{P(\mathbf{d})} = \frac{P(\mathbf{m}) \mathcal{L}(\mathbf{m})}{\int P(\mathbf{m}') \mathcal{L}(\mathbf{m}') d\mathbf{m}'}. \quad (2)$$

<sup>84</sup> In Eq. (2),  $P(\mathbf{m})$  is the prior density representing probabilistic information on the parameters independent of the data.  $P(\mathbf{d}|\mathbf{m})$  is the conditional probability of  $\mathbf{d}$  given  $\mathbf{m}$ ; however, when  $\mathbf{d}$  represents (fixed) measured data, this term is interpreted as the likelihood of  $\mathbf{m}$ , written  $\mathcal{L}(\mathbf{m})$ .  $P(\mathbf{d})$  is a normalization term known as the Bayesian evidence. Finally,  $P(\mathbf{m}|\mathbf{d})$  is the PPD, representing the probability of a set of model parameters given the data and prior information.

<sup>89</sup> In Bayesian inversion, Markov chain Monte Carlo (MCMC) methods are commonly used to draw dependent random samples from the PPD while maintaining detailed balance (reversibility) for the <sup>90</sup> Markov chain according to

$$P(\mathbf{m}|\mathbf{d}) H(\mathbf{m}'|\mathbf{m}) = P(\mathbf{m}'|\mathbf{d}) H(\mathbf{m}|\mathbf{m}'), \quad (3)$$

<sup>92</sup> where  $H(\mathbf{m}'|\mathbf{m})$  is the probability of a random transition from the current model  $\mathbf{m}$  to a new model <sup>93</sup>  $\mathbf{m}'$ . Metropolis-Hastings (MH) sampling is based on treating  $H(\mathbf{m}'|\mathbf{m})$  as the product of a proposal <sup>94</sup> probability density  $Q(\mathbf{m}'|\mathbf{m})$  and an acceptance probability  $A(\mathbf{m}'|\mathbf{m})$ , with  $A(\mathbf{m}'|\mathbf{m})$  defined to satisfy <sup>95</sup> detailed balance as

$$A(\mathbf{m}'|\mathbf{m}) = \min \left[ 1, \frac{Q(\mathbf{m}|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m})} \frac{P(\mathbf{m}'|\mathbf{d})}{P(\mathbf{m}|\mathbf{d})} \right] = \min \left[ 1, \frac{Q(\mathbf{m}|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m})} \frac{P(\mathbf{m}') \mathcal{L}(\mathbf{m}')}{P(\mathbf{m}) \mathcal{L}(\mathbf{m})} \right]. \quad (4)$$

<sup>96</sup> In practice, a random number  $\zeta$  is drawn from a uniform distribution on  $[0, 1]$  such that if  $\zeta \leq A(\mathbf{m}'|\mathbf{m})$ , <sup>97</sup>  $\mathbf{m}'$  is accepted as the next model in the chain and, if not, another copy of  $\mathbf{m}$  is included in the chain.

<sup>98</sup> In many practical inverse problems an appropriate parameterization consistent with the information <sup>99</sup> content of the data is not known *a priori*; for example, in the geoacoustic inversion considered here the <sup>100</sup> number of seabed-layer interfaces is unknown. In this case, the number of interfaces  $k$  can be included <sup>101</sup> hierarchically as a hyper-parameter to be estimated in the inversion. To do so, let  $\mathcal{K}$  be a countable set <sup>102</sup> of possible parameterizations for the problem and  $k$  be an index specifying parameter choices (e.g.,  $k$  <sup>103</sup> can indicate the number of seabed interfaces). In this case, Bayes' theorem can be written [6]

$$P(k, \mathbf{m}_k | \mathbf{d}) = \frac{P(k) P(\mathbf{m}_k | k) \mathcal{L}(k, \mathbf{m}_k)}{\sum_{k' \in \mathcal{K}} \int_{\mathcal{M}_{k'}} P(k') P(\mathbf{m}'_{k'} | k') \mathcal{L}(k', \mathbf{m}'_{k'}) d\mathbf{m}'_{k'}}, \quad (5)$$

104 where  $P(k) P(\mathbf{m}_k|k)$  and  $\mathcal{L}(k, \mathbf{m}_k)$  represents the prior and likelihood, respectively, for the parameter  
 105 set  $\mathbf{m}_k$  in parameter space  $\mathcal{M}_k$  associated with index  $k$ .

106 To apply MCMC sampling to the trans-D parameter space, moves from one parameterization to  
 107 another must be taken included in a probabilistic manner. This paper applies the birth/death reversible-  
 108 jump Markov-chain Monte Carlo (rjMCMC) algorithm in which the MH acceptance criterion, Eq. (4),  
 109 is generalized to the Metropolis-Hastings-Green (MHG) criterion [6]

$$A(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k) = \min \left[ 1, \frac{Q(k, \mathbf{m}_k|k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k)} \frac{P(k')}{P(k)} \frac{P(\mathbf{m}'_{k'}|k')}{P(\mathbf{m}_k|k)} \frac{\mathcal{L}(k', \mathbf{m}'_{k'})}{\mathcal{L}(k, \mathbf{m}_k)} |\mathbf{J}| \right]. \quad (6)$$

110 In Eq. (6),  $|\mathbf{J}|$  is the determinant of the Jacobian matrix for the transformation from state  $(k, \mathbf{m}_k)$  to  
 111  $(k', \mathbf{m}'_{k'})$ , which can be shown to be unity for the rjMCMC algorithm described below [6, 7].

112 In the geoacoustic inversion considered here, the seabed is modeled as a stack of  $k$  sediment layers,  
 113 with the  $i$ th layer ( $i = 1, \dots, k$ ) represented by interface depth  $z_i$ , sound speed  $c_i$ , density  $\rho_i$ , and  
 114 attenuation  $\alpha_i$ , overlying a basement half-space with parameters  $c_{k+1}$ ,  $\rho_{k+1}$ , and  $\alpha_{k+1}$ . The interface  
 115 depths form a partition model and are treated independently from the geoacoustic parameters, which are  
 116 required to have uniform, depth-independent prior densities. The rjMCMC algorithm consists of three  
 117 type of moves which are accepted/rejected according to Eq. (6): perturbations, which randomly change  
 118 parameter values of the existing parameterization; births, which create a new layer with an interface at a  
 119 random depth; and deaths, which delete a layer at random. To achieve  $|\mathbf{J}| = 1$  in Eq. (6), for birth moves  
 120 the geoacoustic parameter values for a new layer must at most depend on the existing parameter values  
 121 at that depth, and for death moves the parameter values for the resulting larger layer must depend only  
 122 the parameters of the portion of the original layer that is retained.

### 123 3.2 Likelihood function

124 In a Bayesian framework, it is crucial to specify the likelihood function, which is based on assumptions  
 125 on the data error statistics that must account for both measurements errors, due to factors such as  
 126 environmental and instrument noise, and theory errors, due to model approximations. For Gaussian-  
 127 distributed errors (consistent with the central limit theorem), the likelihood can be written

$$\mathcal{L}(\mathbf{m}) = \frac{1}{(2\pi)^{N/2} |\mathbf{C}_d|^{1/2}} \exp \left[ -\frac{1}{2} \mathbf{r}^\top(\mathbf{m}) \mathbf{C}_d^{-1} \mathbf{r}(\mathbf{m}) \right], \quad (7)$$

128 where  $\mathbf{r}(\mathbf{m}) = \mathbf{d} - \mathbf{d}(\mathbf{m})$  are the data residuals, i.e., the difference between the measured data and  
 129 data predicted for parameters  $\mathbf{m}$ , and  $\mathbf{C}_d$  is the data error covariance matrix. Under the assumption  
 130 of independent errors,  $\mathbf{C}_d$  is a diagonal matrix. Let be  $\sigma_i^2$  be the  $i$ th diagonal element of  $\mathbf{C}_d$  (i.e., the

131 variance associated to the  $i$ th datum); the likelihood then simplifies to

$$\mathcal{L}(\mathbf{m}) = \frac{1}{(2\pi)^{N/2} \prod_{i=1}^N \sigma_i} \exp \left[ -\frac{1}{2} \sum_{i=1}^N \left( \frac{r_i(\mathbf{m})}{\sigma_i} \right)^2 \right]. \quad (8)$$

132 It is often practical, for numerical reasons, to consider the natural logarithm of likelihood:

$$\log [\mathcal{L}(\mathbf{m})] = -\frac{N}{2} \log(2\pi) - \sum_{i=1}^N \log \sigma_i - \frac{1}{2} \sum_{i=1}^N \left( \frac{r_i(\mathbf{m})}{\sigma_i} \right)^2. \quad (9)$$

133 The last term on the right side of Eq. (9) is a sum of the squares of Gaussian-distributed random variables

134 with unit variance, so the expected value for the log-likelihood is [8]

$$\mathbb{E} [\log(\mathcal{L})] = -\frac{N}{2} \log(2\pi) - \sum_{i=1}^N \log \sigma_i - \frac{N}{2}, \quad (10)$$

135 with variance

$$\text{Var} [\log(\mathcal{L})] = 2N. \quad (11)$$

### 136 3.3 Priors, proposals, and acceptance probabilities

137 Prior probability densities play an important role in Bayesian inference and, together with proposal  
138 densities and likelihood, define the acceptance ratios in MCMC sampling, e.g., Eq. (6). In this paper,  
139 priors are set to constrain the inversion to physically-realistic regions of the trans-D parameter space  
140 while allowing the data information to predominately determine the solution. Three types of priors must  
141 be defined: the prior for the number of interfaces, the prior for the partition of interface depths, and the  
142 prior for the geoacoustic parameters.

143 The prior for the number of interfaces is chosen here to be a Poisson distribution [6],

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (12)$$

144 with Poisson's parameter  $\lambda=4$  chosen to slightly favor models with low complexity ( $k \approx 3$  or 4, which have  
145 equal probability). Note that this is not a highly-informative prior, as the variation in prior probabilities  
146 for  $0 \leq k < 10$  is only about a factor of 10, while likelihood values differentiating between the data fit  
147 of various models often vary over many orders of magnitude. Hence, data information predominates in  
148 defining the number of interfaces.

149 The prior for sediment-layer interface depths is taken to be a uniform probability density over the  
150 interval  $[0, z_{\max}]$ , with  $z_{\max}$  set deep enough that the geoacoustic parameters below this depth do not  
151 influence the acoustic data. The corresponding prior for the depth partition  $\mathbf{z}_k$  (set of  $k$  interface depths)

<sup>152</sup> is conveniently expressed using the Dirichlet distribution [9, 10]

$$P(\mathbf{z}_k|k) = k! z_{\max}^{-k}. \quad (13)$$

<sup>153</sup> To allow a depth-dependent prior for the geoacoustic parameters  $c$ ,  $\rho$ , and  $\alpha$ , given the requirement  
<sup>154</sup> of uniform priors in the partition-modeling approach, the unknown model parameters for each layer are  
<sup>155</sup> taken to be the deviation from a predefined depth-dependent reference model, with depth-independent  
<sup>156</sup> priors set for the deviations. As an example, if  $c_R(z)$  is the reference model for sound speed, then the  
<sup>157</sup> unknown sound-speed parameter for the  $i$ th layer (associated with interface depth  $z_i$ ) is  $\delta c_i$ , with the  
<sup>158</sup> layer sound speed used in data prediction given by  $c_i = c_R(z_i) + \delta c_i$ . Defining  $\mathbf{g}_k$  to be the set of  
<sup>159</sup> geoacoustic parameter deviations for a model with  $k$  interfaces, and  $\Delta c$ ,  $\Delta \rho$ , and  $\Delta \alpha$  to be the (depth-  
<sup>160</sup> independent) prior bound widths for deviation parameters  $\delta c$ ,  $\delta \rho$ , and  $\delta \alpha$ , respectively, the prior for the  
<sup>161</sup> geoacoustic parameter deviations (including the halfspace) is

$$P(\mathbf{g}_k|k) = \frac{1}{(\Delta c \Delta \rho \Delta \alpha)^{k+1}}. \quad (14)$$

<sup>162</sup> Combining the above results, the prior for state  $(k, \mathbf{m}_k)$  is given by the product of the priors on  $k$ ,  
<sup>163</sup>  $\mathbf{z}_k$ , and  $\mathbf{g}_k$  [given by Eqs. (12)–(14), respectively], leading to

$$P(k)P(\mathbf{m}_k|k) = \frac{\lambda^k}{e^\lambda z_{\max}^k (\Delta c \Delta \rho \Delta \alpha)^{k+1}}. \quad (15)$$

<sup>164</sup> To formulate acceptance probabilities, Eq. (6), for perturbation, birth, and death moves also requires  
<sup>165</sup> specification of the corresponding proposal densities. Beginning with perturbation moves (no change in  
<sup>166</sup> number of interfaces,  $k' = k$ ), it is clear from Eq. (15) that the prior ratio  $P(k)P(\mathbf{m}'_k|k) / [P(k)P(\mathbf{m}_k|k)]$   
<sup>167</sup> is unity. Further, using a symmetric proposal density for geoacoustic parameters,  $Q(\mathbf{g}'_k|\mathbf{g}_k) = Q(\mathbf{g}_k|\mathbf{g}'_k)$ ,  
<sup>168</sup> leads to a unit proposal ratio (the symmetric parameter proposal used in this work is described in  
<sup>169</sup> Sec. 3.4). Hence, for perturbation moves the acceptance probability simplifies to the likelihood ratio

$$A_P(k, \mathbf{m}'_k|k, \mathbf{m}_k) = \min \left[ 1, \frac{\mathcal{L}(k, \mathbf{m}'_k)}{\mathcal{L}(k, \mathbf{m}_k)} \right]. \quad (16)$$

<sup>170</sup> For birth steps ( $k' = k + 1$ ), the prior ratio, using Eq. (15), is

$$\frac{P(k') P(\mathbf{m}'_{k'}|k')}{P(k) P(\mathbf{m}_k|k)} = \frac{\lambda}{z_{\max} \Delta c \Delta \rho \Delta \alpha}. \quad (17)$$

<sup>171</sup> The proposal ratio for a birth step is the ratio of the proposal for the reverse step [death of the proposed  
<sup>172</sup> layer, with probability  $1/(k+1)$ ] to that of the proposed birth step given uniform probability of  $1/z_{\max}$   
<sup>173</sup> for randomly inserting an interface over  $(0, z_{\max})$  with a new geoacoustic parameter set  $\mathbf{g}'_{k'}$  which differ

<sup>174</sup> from the current values  $\mathbf{g}_k$  only over the new layer:

$$\frac{Q(k, \mathbf{m}_k | k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k)} = \frac{z_{\max}}{k+1} \frac{Q(\mathbf{g}_k | \mathbf{g}'_{k'})}{Q(\mathbf{g}'_{k'} | \mathbf{g}_k)}. \quad (18)$$

<sup>175</sup> Dosso et al. [10] showed that proposing geoacoustic parameters for a new layer from the prior can be  
<sup>176</sup> more efficient than using proposal densities designed to remain near current values. Here, values for new  
<sup>177</sup> geoacoustic parameters are generated randomly from the uniform parameter priors, i.e.,  $Q(\mathbf{g}'_{k'} | \mathbf{g}_k) =$   
<sup>178</sup>  $1/(\Delta c \Delta \rho \Delta \alpha)^{k'}$ , so the proposal ratio for birth moves becomes

$$\frac{Q(k, \mathbf{m}_k | k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k)} = \frac{z_{\max} \Delta c \Delta \rho \Delta \alpha}{k+1}. \quad (19)$$

<sup>179</sup> Hence, the acceptance probability for birth moves [product of prior ratio given by Eq. (17), proposal  
<sup>180</sup> ratio given by Eq. (19), and likelihood ratio] becomes

$$A_B(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k) = \min \left[ 1, \frac{\lambda}{k+1} \frac{\mathcal{L}(k', \mathbf{m}'_{k'})}{\mathcal{L}(k, \mathbf{m}_k)} \right]. \quad (20)$$

<sup>181</sup> A similar procedure leads to the acceptance probability for death moves ( $k' = k - 1$ ) of

$$A_D(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k) = \min \left[ 1, \frac{k}{\lambda} \frac{\mathcal{L}(k', \mathbf{m}'_{k'})}{\mathcal{L}(k, \mathbf{m}_k)} \right]. \quad (21)$$

<sup>182</sup> The term  $\lambda/(k+1)$  in Eq. (20) represents the slight preference to accept birth moves if  $k < \lambda - 1$  and to  
<sup>183</sup> reject birth moves if  $k \geq \lambda$  due to the Poisson prior on  $k$  in Eq. (12). Likewise, the term  $k/\lambda$  in Eq. (21)  
<sup>184</sup> represents the preference to accept death moves if  $k > \lambda$  and to reject death moves if  $k < \lambda$ .

<sup>185</sup> Finally, as suggested by [9], a joint prior for  $\rho$  and  $c$  has been defined from a set of laboratory and  
<sup>186</sup> environmental measurements [11] [12] to avoid unreasonable sound speed/density combinations.

### <sup>187</sup> 3.4 Unit-lag principal component proposal density

<sup>188</sup> The proposal density for perturbation moves is a key component of the efficiency of MCMC sampling,  
<sup>189</sup> which is considered in this section (since the model dimension does not change in perturbation moves, the  
<sup>190</sup> explicit dependence on hyper-parameter  $k$  is omitted here for simplicity). The optimal proposal density  
<sup>191</sup> is given by the PPD itself; in this case the MH criterion, Eq. (4), reduces to unity so all perturbations  
<sup>192</sup> are accepted. In practical problems the PPD is not available during sampling; however, it is possible to  
<sup>193</sup> make use of a local linear approximation to define an efficient proposal scheme.

<sup>194</sup> For a fixed-dimensional inverse problem with Gaussian-distributed data errors of covariance matrix  
<sup>195</sup>  $\mathbf{C}_d$  and a Gaussian prior of covariance  $\mathbf{C}_p$ , a local linear approximation to the PPD is given by Gaussian

196 probability density with posterior model covariance matrix

$$\mathbf{C}_m = [\mathbf{J}^\top \mathbf{C}_d^{-1} \mathbf{J} + \mathbf{C}_p^{-1}]^{-1}, \quad (22)$$

197 where  $\mathbf{J}$  is the Jacobian matrix of partial derivatives  $J_{ij} = \partial d_i / \partial m_j$ . To apply this PPD approximation  
 198 in MCMC sampling, perturbations can be carried out in a principal-component (PC) parameter space  
 199 where the PC parameters are uncorrelated. The orthogonal transformation (rotation) between physical  
 200 parameters  $\mathbf{m}$  and PC parameters  $\tilde{\mathbf{m}}$  is given by

$$\tilde{\mathbf{m}} = \mathbf{U}^\top \mathbf{m}, \quad \mathbf{m} = \mathbf{U} \tilde{\mathbf{m}}, \quad (23)$$

201 where  $\mathbf{U}$  is the column-eigenvector matrix of the posterior covariance matrix,

$$\mathbf{C}_m = \mathbf{U} \mathbf{W} \mathbf{U}^\top, \quad (24)$$

202 and  $\mathbf{W} = \text{diag}[w_i]$  is the eigenvalue matrix, with  $w_i$  representing the variance of PC parameter  $\tilde{m}_i$ . The  
 203 PC decomposition provides both directions and length scales for effective parameter perturbations: PC  
 204 parameters are perturbed individually, with the perturbation for  $\tilde{m}_i$  drawn from a Gaussian distribution  
 205 with variance  $w_i$ , and the perturbed models rotated back to physical space for likelihood evaluation.  
 206 This procedure is initiated using the linearized model covariance approximation given by Eq. (??),  
 207 with Jacobian partial derivatives computed numerically at a high-probability model (determined via a  
 208 preliminary burn-in stage). Uniform bounded priors of width  $\Delta m_i$  are approximated by taking  $\mathbf{C}_p$  to  
 209 be a diagonal matrix with variances equal to those of the uniform distributions, i.e.,  $(\Delta m_i)^2 / 12$ .

210 Following initialization, it can be advantageous to replace the linearized posterior covariance matrix  
 211 with a nonlinear estimate computed using the  $S$  samples to that point in Markov chain

$$\mathbf{C}_m = \sum_{s=1}^S (\mathbf{m}_s - \bar{\mathbf{m}})(\mathbf{m}_s - \bar{\mathbf{m}})^\top, \quad (25)$$

212 where  $\bar{\mathbf{m}}$  represents the sample mean. Although this procedure formally violates the requirement for  
 213 MCMC steps to depend only the current state, it is allowable in practice as it represents a diminishing  
 214 adaptation [13]. However, Dosso et al. [10] found that using the nonlinear covariance estimate was less  
 215 successful in trans-D inversion than fixed-D inversion because of the common multi-modal character of  
 216 parameter sub-spaces, particularly at high  $k$  values. In such cases, the covariance estimated about the  
 217 sample mean may not be a good representation of the sampling behaviour. An alternative is to compute

218 the covariance of the changes along the Markov chain, referred to as the unit-lag covariance matrix

$$\mathbf{C}_m^1 = \sum_{s=1}^{S-1} (\mathbf{m}_{s+1} - \mathbf{m}_s)(\mathbf{m}_{s+1} - \mathbf{m}_s)^\top. \quad (26)$$

219 The unit-lag covariance quantifies the direction and length scales of the random-walk history of the  
 220 MCMC chain, and can be used to propose efficient perturbations for future steps. For trans-D inversion,  
 221 distinct proposal densities are applied for models with differing numbers of interfaces  $k$ . When a birth  
 222 or death step results in a model with a value of  $k$  which has not been encountered previously, the  
 223 linearized covariance approximation, Eq. (22), is computed at the current model and applied initially; this  
 224 covariance matrix is subsequently updated with the unit-lag covariance estimated from MCMC sampling.  
 225 For numerical stability, covariance estimation is carried out for non-dimensionalized parameters scaled  
 226 to vary over the unit interval

### 227 3.5 Parallel tempering

228 An efficient rjMCMC algorithm requires a well-mixed Markov chain that effectively samples the trans-D  
 229 parameter space, achieving reasonable acceptance rates for birth/death moves and parameter perturba-  
 230 tions in potentially multi-modal subspaces. Parallel tempering [14, 15, 16] provides a method to achieve  
 231 this, based on running a series of parallel interacting Markov chains for which the acceptance criterion  
 232 is successively relaxed by raising the likelihood to powers  $1/T$  (where  $T \geq 1$  is referred to as the sam-  
 233 pling temperature). High- $T$  chains have an increased probability of accepting low-likelihood models, and  
 234 hence provide a wider sampling of the trans-D space with increased probability of accepting dimension  
 235 jumps and large perturbations. Conversely, low- $T$  chains provide concentrated sampling but are prone  
 236 to become trapped in localized regions of the space. Since chains at  $T > 1$  provide biased sampling  
 237 of the PPD (sample from distributions that are increasingly similar to the prior as  $T$  increases), only  
 238 the samples collected at  $T=1$  are retained to characterize the PPD. Parallel tempering improves PPD  
 239 sampling by providing probabilistic interchange between chains with different temperatures, ensuring  
 240 that low- $T$  chains can access all regions of the space, providing a robust ensemble sampler. Applying the  
 241 MH criterion to a proposed interchange between chains consisting of model  $\mathbf{m}_i$  at temperature  $T_i$  and  
 242  $\mathbf{m}_j$  at  $T_j$  leads to acceptance probability

$$A[(\mathbf{m}_j, T_i), (\mathbf{m}_i, T_j) | (\mathbf{m}_i, T_i), (\mathbf{m}_j, T_j)] = \min \left[ 1, \left( \frac{\mathcal{L}(\mathbf{m}_i)}{\mathcal{L}(\mathbf{m}_j)} \right)^{1/T_j - 1/T_i} \right]. \quad (27)$$

243 Parallel tempering requires an appropriate choice of the number and temperatures of chains, which is  
 244 discussed in Sec. 3.6. Interchange attempts are not computationally expensive since forward modelling  
 245 (data prediction) is not required.

246 To take advantage of the combination of parallel tempering and the PC proposal scheme described  
 247 in Sec. 3.4 within trans-D inversion, a different temperature-dependent proposal is applied to match the  
 248 tempered acceptance criterion for each chain [10]. In particular, the linearized covariance estimate used  
 249 to initialize the PC proposal for temperature  $T$  is taken to be

$$\mathbf{C}_m(T) = [\mathbf{J}^\top (T \mathbf{C}_d)^{-1} \mathbf{J} + \mathbf{C}_p^{-1}]^{-1}. \quad (28)$$

250 After this initialization, the unit-lag covariance matrix  $\mathbf{C}_m^1(T)$  is computed from the MCMC sampling  
 251 with temperature  $T$  and used to replace the linearized approximation. This results in a different PC  
 252 proposal for each number of interfaces  $k$  and for each temperature  $T$ , based on the covariance esti-  
 253 mated for that  $k$ - $T$  combination. This generally produces larger, less-directional perturbations at higher  
 254 temperatures, which is consistent with the correspondingly-relaxed acceptance criterion.

### 255 3.6 Sequential Bayesian Inversion

256 This section describes the sequential Monte Carlo strategy used to invert the large sequential data set  
 257 described in Sec. 2. In sequential inversion, the idea is to use the information available from one data  
 258 set (e.g., reflectivity data for one source ping) to initialize the inversion for the next data set (next  
 259 ping). Assuming the PPD for the current data  $\mathbf{d}_i$  is available, the goal is to efficiently estimate the  
 260 PPD for a new data set  $\mathbf{d}_{i+1}$ . Defining the  $j$ th particle as a model  $\mathbf{m}_j$  plus its probabilistic weight  
 261  $w_{ji} = \mathcal{L}(\mathbf{m}_j, \mathbf{d}_i)$ , a particle cloud (set) of  $C$  particles is drawn from the PPD for  $\mathbf{d}_i$ . The first stage of  
 262 updating the cloud to represent the PPD of the new data consists of resampling. The likelihoods for the  
 263 existing particles with respect to the new data set  $\mathcal{L}(\mathbf{m}_j, \mathbf{d}_{i+1})$  are computed, updating the weights of  
 264 the particles to reflect their ability to represent the PPD of the new data. A new cloud of  $C$  particles  
 265 is then built by randomly drawing particles (with replacement) from the existing cloud with probability  
 266 proportional to the updated particle weights. Resampling replaces low-likelihood particles with copies of  
 267 higher-likelihood particles, and is carried out separately for each number of interfaces  $k$  in the original  
 268 sample to preserve the distribution of  $k$  values at this stage.

269 The second stage of the sequential Monte algorithm, referred to as burn-in, is designed to moving  
 270 the particles into high-probability regions of the trans-D parameter space for the new data. Particles are  
 271 updated via a set of rjMCMC steps, including parameter perturbations and interface births and deaths,  
 272 employing unit-lag PC parameter proposals and parallel tempering, as described in Sec. 3.4 and 3.5.  
 273 The tempering schedule (set of temperatures) for each particle is chosen so that there are  $C/2$  parallel  
 274 chains sampled at temperature  $T = 1$  and  $C/2$  chains sampled at closely-spaced temperatures  $T > 1$ . To  
 275 avoid overly-complex models (i.e., models with too many interfaces) at this stage, the MHG acceptance  
 276 criterion is modified by replacing the likelihood in Eq. (6) with the exponential of the negative of the

277 Bayesian information criterion (BIC), which applies a penalty for model complexity in terms of number  
278 of parameters:

$$\text{BIC}(\mathbf{m}) = -2 \log [\mathcal{L}(\mathbf{m})] + M \log N. \quad (29)$$

279 Applying the BIC in sampling reduces model complexity (favors models that fit the data with a smaller  
280 number of interfaces) in the cloud at this stage but introduces a bias in the PPD which is accounted  
281 for in the subsequent stage. The burn-in stage is terminated when 5% of the particles at  $T = 1$  achieve  
282 log-likelihoods greater than the expected value  $\mathbb{E}[\log(\mathcal{L})]$  for the new data set, as given by Eq. (10). In  
283 case an insufficient number of particles reach the expected value in a predefined maximum number of  
284 tempered rjMCMC steps, the burn-in stage is repeated setting as the target  $\mathbb{E}[\log(\mathcal{L})] - \sqrt{\text{Var}[\log(\mathcal{L})]}$ .

285 After burn-in, one or more stages consisting of resampling and parallel-tempering rjMCMC sampling  
286 based on the likelihood (rather than the BIC) are applied to remove the bias introduced with BIC  
287 sampling, retain particle diversity in the cloud, and balance the cloud to the new PPD. Finally a sample  
288 of the PPD is drawn from the untempered ( $T = 1$ ) particles in the cloud by saving samples from rjMCMC  
289 steps until the sample is complete. The whole procedure can then be repeated for the next data set,  
290  $\mathbf{d}_{i+2}$ , and so on.

## 291 4 Simulation study

292 This section illustrates the sequential Monte Carlo inversion using realistic simulated data. Simulation  
293 provides several advantages in the test phase. First, the unknowns of the problems are available as well as  
294 the most appropriate parameterization (the number of seabed layers) and the error statistics. Moreover  
295 the knowledge of error statistics means the target likelihood is known accurately and the comparison  
296 with the true value straightforward.

297 Algorithms developed in previous studies [2] inverting plane-wave reflection coefficients highlighted  
298 the advantages of sequential Bayesian inversions in terms of computing time and quality of results. In this  
299 section, the inversion algorithm developed is applied to simulated spherical-wave reflection coefficients  
300 along a range-dependent track. The true model used to generate the data is taken from [2] and is  
301 presented in Fig. ???. The simulation is based on the geometry of the Clutter09 experiment and provides,  
302 for each of 170 pings, a set of 32 reflection coefficients for grazing angles over  $[29, 67]^\circ$  for 6 frequency bands  
303 centered at 975, 1100, 1250, 2100, 2400, and 2700 Hz with bandwidths of 37.5 Hz. The most challenging  
304 features of the profile are located between pings 45 and 46, where a geological fault is simulated, and  
305 between pings 148 to 155, where the profile mimics an erosional channel. Data were computed for  
306 the true model by solving the Sommerfeld integral [Eq. 1] using Levin integration as described by [4].  
307 Gaussian-distributed errors with zero mean and standard deviation 0.03 were added to the data.

308 To demonstrate the efficiency of the unit-lag PC proposal density, a comparison of marginal proba-

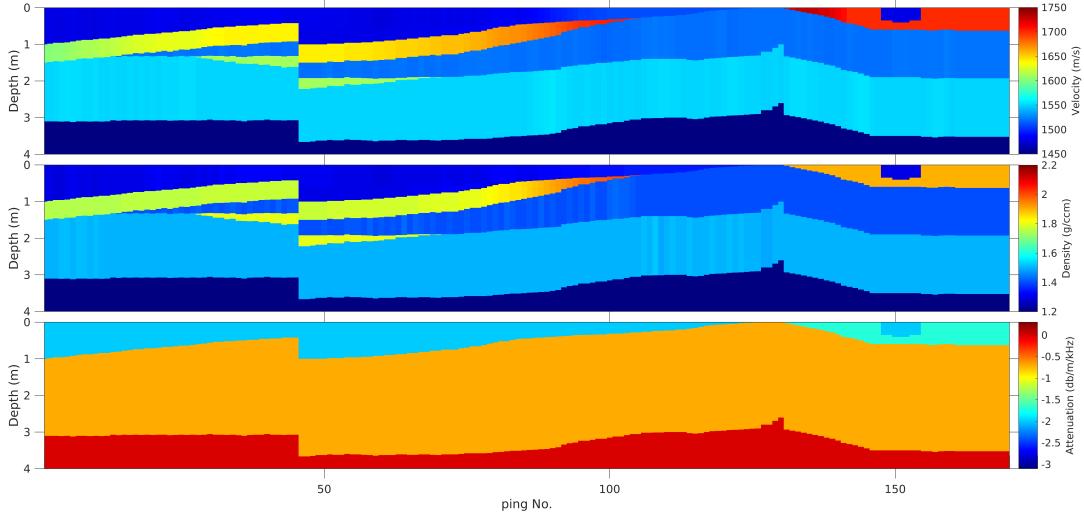


Figure 4: Environment used for simulation.

309 bility density for sediment sound speed profiles obtained for the first ping of the simulation is presented  
 310 in Fig. 5. The top row shows the sound-speed marginal estimated from sets of 100, 316, 1000, 3162,  
 311 and 10000 samples obtained by sampling using a Cauchy proposal density as in [ ]. The Cauchy pro-  
 312 posal density perturbs physical (unrotated) parameters with a scale parameter chosen equal to 1/40 of  
 313 the prior distribution width for the specific parameter. The Cauchy proposal density is preferred to  
 314 Gaussian proposal as it occasionally allows large perturbation lengths. The middle row of Fig. 5 shows  
 315 marginal densities obtained using the unit-lag PC proposal density for the same number of samples, and  
 316 the third row shows the PPD estimated from 100,000 samples collected with Cauchy proposal which  
 317 represents a well converged sample for comparison. Visual inspection of results indicates that the unit-  
 318 lag PC proposal density significantly improves the convergence rate such that fewer samples are needed  
 319 to produce a reasonable estimate of the PPD. The improved efficiency in sampling is also evident when  
 320 the perturbations are examined. In Fig. 6, a set of 15 consecutive steps for both the Cauchy proposal  
 321 density (blue) and the PC proposal density (red) projected on the space  $c_1, \rho_1$  are plotted above the  
 322 joint marginal probability density for  $c_1$  and  $\rho_1$ , estimated from the PPD. For the Cauchy proposal the  
 323 steps are relatively small and overall the chain section does not explore the space efficiently. For the  
 324 PC proposal, step lengths are about the size of the covariance semi-axes and steps are taken along the  
 325 covariance ellipsoid semi-axes directions, and explores the parameter space much more efficiently.

326 The result of the full sequential inversion consists of a set of 170 trans-D PPDs. The marginal PPDs  
 327 for profile of geoacoustic parameters are estimated using a cloud of 200 particles with the algorithm  
 328 described in Sect. 3.6, the burn-in was set up such that in case the convergence (i.e. at least 5% of  
 329 the particles do not reach target likelihood) is not reached after 15 iterations of burn-in procedure or  
 330 1800 s the ping is skipped. From the cloud at every ping, a sample of set of 20000 samples was collected.

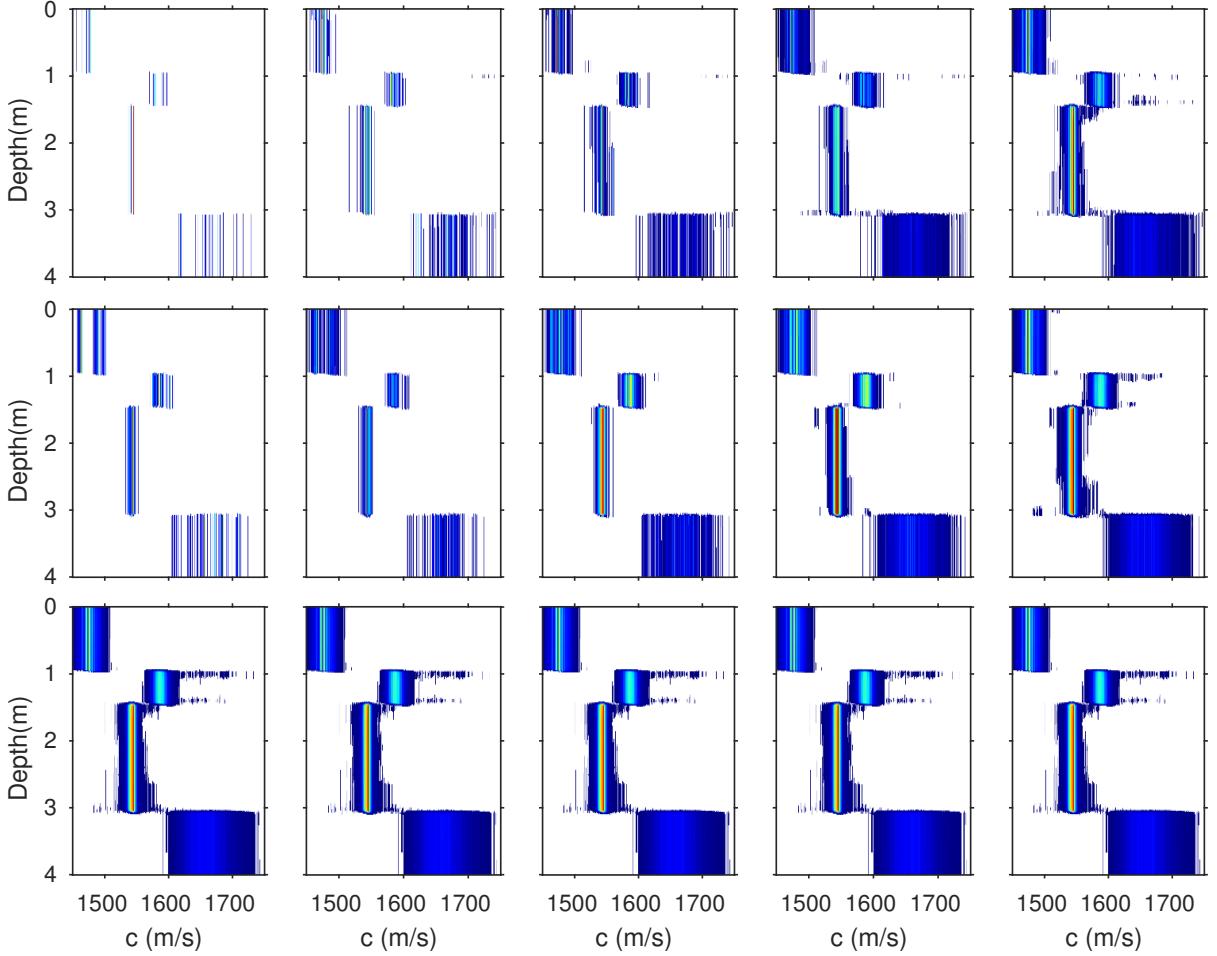


Figure 5: Marginal probability distribution for sound speed  $c$  from the first ping of the simulated track. Top panels, left to right: marginal distribution obtained by sample collected using Cauchy proposal density estimated from 100, 316, 1000, 3162 and 10000 samples respectively. Middle panels, left to right: marginal distribution obtained by sample collected using principal component unit-lag proposal density estimated from 100, 316, 1000, 3162, and 10000 samples respectively. Bottom panels: marginal distribution estimate from a 100,000 samples.

331 The mean of the PPD marginalized for the sound speed and density is summarized in Fig. ???. A  
 332 complete parameter uncertainty analysis for pings 44, 45, 46 and 47 is presented in Fig. ???. From the  
 333 comparison of the marginal distributions it is possible to understand how the algorithm can recover from  
 334 a non-converged sample. In this case the environment changes abruptly at ping 45, and the information  
 335 provided by the particle cloud that samples ping 44 is not a good candidate to initialize the cloud for  
 336 ping 45. In the time given to the algorithm to find a good model, the sampler does not finish the burn-  
 337 in, so it collects samples in an unlikely region of the model space. Nonetheless, the parallel tempering  
 338 strategy combined with the unit-lag PC proposal density results in efficient exploration of the model  
 339 space and quickly recovers from the abrupt change at ping 47. Data fit relative to pings 44 to 47 are  
 340 shown in Fig. ???. Even though a rigorous benchmark analysis was not carried out, the inversion time  
 341 for inversions that use unit-lag PC proposal and Cauchy proposal was monitored. The Cauchy proposal-  
 342 based inversion took  $\approx 18.1$  hours against the  $\approx 9.7$  used for the PC unit-lag proposal based inversion.

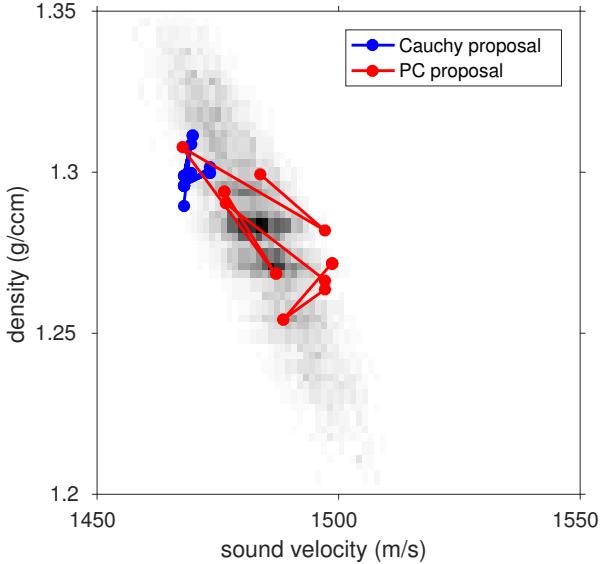


Figure 6: Values for parameter couples  $[c_1\rho_1]$  collected for 15 subsequent steps in the Markov chain for the first ping of simulation. Blue dots represent model collected with the Cauchy density proposal, red dots represents models obtained with the principal component unit-lag proposal density.

343 The main difference in time was given by the fact that the Cauchy-based algorithm failed to converge  
 344 quickly in correspondence to abrupt changes in the environment (pings 45 and ping 148).

## 345 5 Malta Plateau inversion

346 In this section results from the dataset from Malta Plateau described in Section 2 are reported. For this  
 347 inversion a set of 200 particles was used and PPD were estimated from a set of 10000 samples collected  
 348 from the particle cloud after the burn-in. PPDs for all the pings considered were sampled using both the  
 349 unit-lag PC proposal density and Cauchy proposal density, retrieving in both cases similar geoacoustic  
 350 profiles. In this section some aspects of the inference from the estimated set of PPD are explored. Fig. 10  
 351 illustrates the means for the range-dependent PPDs marginalized with respect to sound speed, density  
 352 and attenuation as sampled using both the Cauchy proposal density (Fig. 10a) and PC proposal density  
 353 (Fig. 10b). Some regions of the profile are of particular interest. The topmost layer of the sediments,  
 354 in the North-West (left) section of the profile displays a low-velocity wedge-shaped layer that eventually  
 355 disappear about 36.48 Lat N. This layer is characterized by a unusually low-value for sound speed and  
 356 consistently a low value for attenuation. These results support the idea of a surficial mud layer, as  
 357 suggested by Holland and Dettmer [17]. Both the shape and position of the low-velocity wedge are  
 358 consistent with information provided by the seismic profile collected along the survey (see below and  
 359 Fig. 12). Another interesting layer is located along the profile at a depth of  $\approx 150 - 155$  m. Density  
 360 and sound speed for this layer appear to be particularly high in the portion of the track located between  
 361 36.43 and 36.46 Latitude North, in which the top margin of the layer appears particularly rough. This

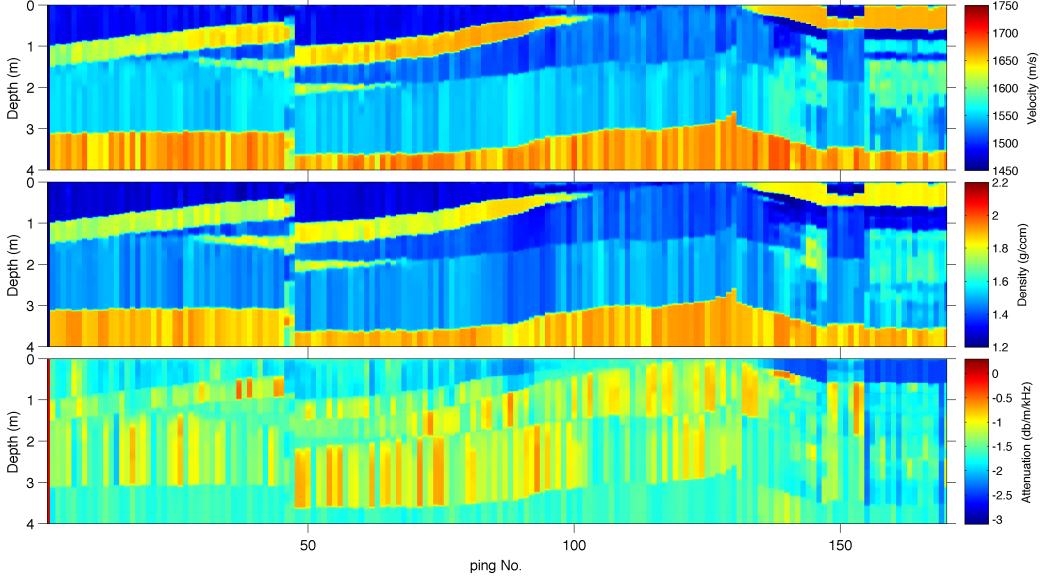


Figure 7: Inversion result in terms of posterior density distribution mean for velocity and density along the track.

roughness is unusual for sediments, and it suggests the possibility of exposure to atmospheric agents during an earlier period of low sea level which may have eroded the surface that, at a later stage, filled with soft sediments. These observations are compatible with results obtained by the means of unit-lag PC proposal density, which are reported in Fig. 10b. Sound speed models from the same data sets were produced by Pinson et al. [18] using the image source method and are reported in Fig. 11. Models from [18] explore the sediments at deeper depth and are less sensitive to thin layers, the main velocity structure for the shallower part of model from [18] is nonetheless highly compatible with the one retrieved here. In particular the North-West low-velocity wedge and the high-velocity layer running 4 to 6 m below the water-sediment interface result consistent in results from both methods. Fig. 12 shows the probability density of interfaces as function of depth along the track compared with high resolution seismic profile for the survey, which is used as reference for results appreciation. To appreciate results from inference, it is important to keep in mind that the seismic profile does not properly define interfaces depth, it is simply a set of seismograms (i.e. time series of displacement field) which is converted from travel time to depth by assuming an arbitrary constant sound speed as scale factor. Since the seabed sound speed is unknown and variable, this scaling process introduces a distortion in the layering structure displayed. Alternatively, the interface depth marginal density from the PPD represent a quantitative map of interface-depth probabilities which is properly scaled using sediment sound speeds sampled in the inversion. Hence, results from marginalization return significant higher information content. Here it is evident that the main structural interfaces match. For example, the match between the wedging layer (marked “A” in Fig. 12) in both seismic profile and the inversion result, as well as the correspondence

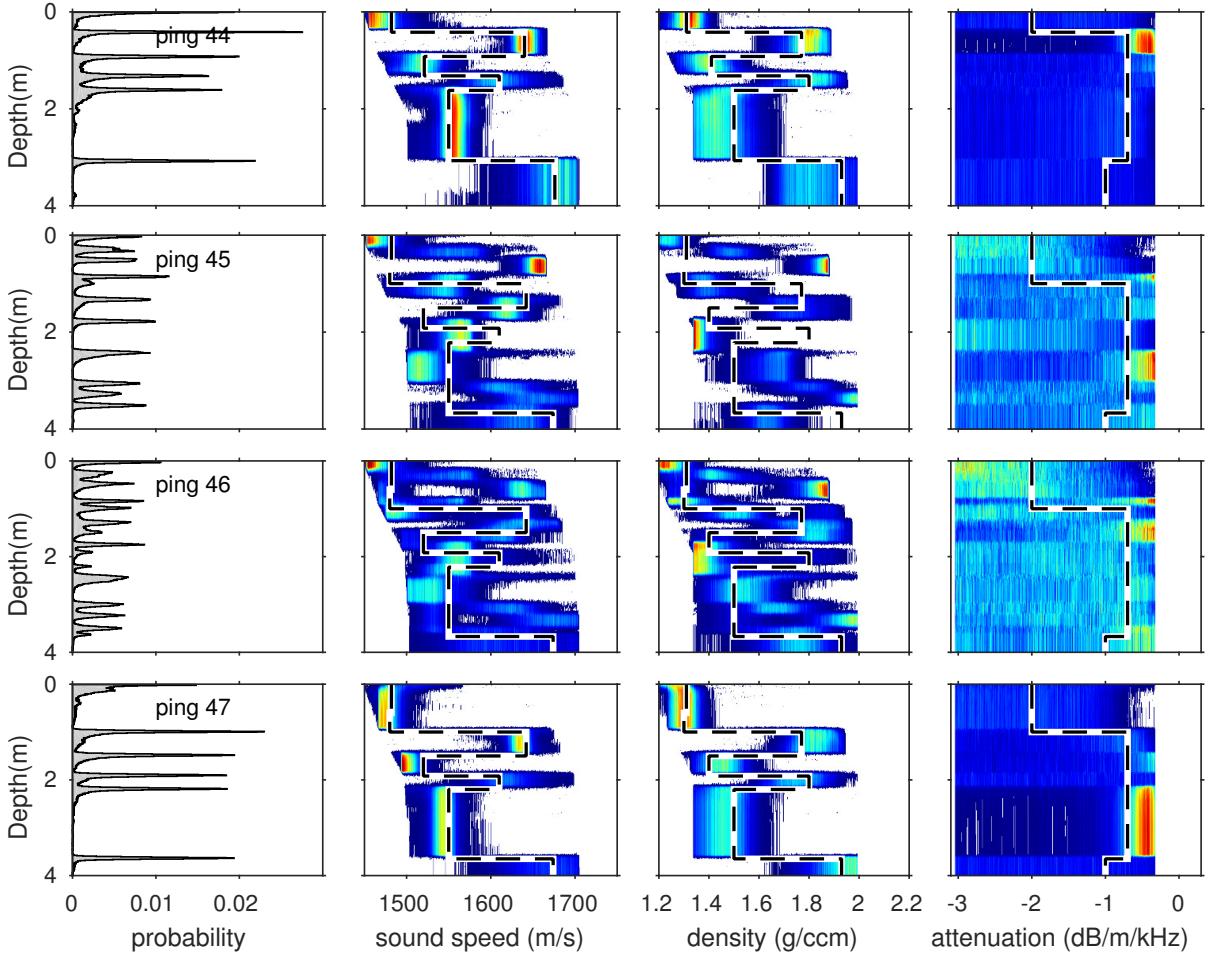


Figure 8: Marginal probability distribution of (left to right) interface depth, sound speed, density and attenuation for the fault pings in the simulation. Top to bottom are shown ping 44, 45, 46 and 47. Dashed lines mark the simulated environment.

of the deep interface marked “B” in Fig. 12. A set of similar regions of interest have been highlighted with shadowed areas in Fig. 12, North to South (left to right) correspondences are indicated for a deep interface not associated with high acoustic impedance contrast, a steep interface, a region in which a new sediment layer becomes evident and a set of three rough and steep interfaces that match closely in the two models. Qualitatively, the two models are highly compatible, although the seismic profile shows better resolution of more layer interfaces that remain stable along the track, the marginal of interfaces depths from reflection coefficient inversion provides quantitative information inferred from measured data. In Fig. 13 the width of the 95% confidence interval for these marginal distribution is shown. Confidence intervals (Figs. 13a and 13b), are smaller in correspondence of peaked marginals and highlight decreased resolution for geoacoustic parameters in proximity of layer interfaces, in area below reflector or where variation of acoustic impedance results limited; notwithstanding the marginal attenuation distribution results constrained by data for the low attenuation values corresponding to the low-velocity wedge, the lack of sensitivity of reflection coefficient data with respect to attenuation results evident. A complete

395 uncertainty estimation can be inferred from individual ping PPDs; an example is given in Figs. 14a and  
396 14b relative to ping 262. Fig. 15 shows the distribution of  $\log(\mathcal{L})$  and  $k$  along the track. The algorithm  
397 is able to fit all the data to the required precision (i.e., the distribution of the likelihood reach the target,  
398 marked in Fig. 15 as a dashed line), and the distribution of the number of interfaces  $k$  seems to be  
399 compatible with the fit, i.e., the algorithm is not over-interpreting datasets with the introduction of  
400 many layers. An example of data fit for three different pings – two subsequent one and an isolated one –  
401 is summarized in Fig. 16 and proves the excellent fit of the predicted data. It appears from the results of  
402 these inversions that the Malta Plateau dataset supports complicated (i.e., many layers) models in which  
403 the number of sediment layers and the value of geoacoustic parameters vary dramatically with range,  
404 nonetheless data for every ping are fitted to the level required by the assumptions made on noise (cf.  
405 Fig. 15). In these inversions sound attenuation result to be the less sensitive parameter and it is retrieved  
406 with some precision only in area in which its value is very small. Although estimation of attenuation  
407 results vague, it plays a very minor role in modelling reflection coefficient (to the scale examined in  
408 this work) which are, in the experiment geometry examined here, mainly sensitive to sound speed and  
409 density which are estimated accurately. Attenuation estimations for sediments present in the Malta  
410 Plateau region were presented by Holland & Dosso [17], which retrieved values of  $\approx 0.0009$  estimated  
411 from reverberation measurements. Here a compatible result is estimated, limited to the low-velocity  
412 wedge in which the attenuation is estimated in a limited-length confidence interval. Results retrieved  
413 using Cauchy proposal density and unit-lag PC proposal density are in good qualitative agreement. As  
414 further test, the whole track has been inverted along the inverse direction (from south-east to north-west)  
415 with compatible geoacoustic results for both sound speed, density and attenuation. Well log data are  
416 available for the track initial and end sites. As further improvement in model validation, the PPD related  
417 to these sites are shown in Fig. 17 in comparison with well log data.

## 418 6 Summary

419 Prior work has reported the effectiveness of Bayesian Filter in inverting, and in inversion result appre-  
420 ciation, large geoacoustic data volumes. Dettmer et al. [2] for example documented the effectiveness  
421 of a non-linear filter that make use of AIS as filtering distribution in a deep water (i.e. plane-wave ap-  
422 proximation for the pressure field). However the efficiency for methods explored in these studies conflict  
423 with the necessity from field operators to obtain accurate-geoacoustic models and uncertainty estimates  
424 urgently. In this paper a Bayesian trans-D algorithm was developed to invert large sequential data vol-  
425 umes and applied to both simulated data and data from an AUV survey in the Mediterranean Sea. The  
426 inversion scheme combines methods commonly used in particle filters, such as the use of probabilistic  
427 weights to track the evolution of a particle cloud, and resampling. An efficient proposal scheme is based

428 on the unit-lag PC proposal density and the parallel tempering. Application to simulated data shows  
429 that the inversion algorithm handles datasets from environments in which parameters values and the  
430 parametrization itself changes abruptly. A large dataset from Malta Plateau was also inverted, and the  
431 resulting model for this region allows a preliminary interpretation in terms of sediments structure. The  
432 layering structure retrieved for this area matches in its main feature the layering structure as obtained  
433 with high-resolution seismic profiling. These results extends findings from Dettmer and Dosso to shallow  
434 water environments, confirming the efficiency of PT and unit-lag proposal density in Bayesian inversion  
435 in complex environment and when spherical-wave propagation is taken in account. Moreover this study  
436 produced an accurate model from a large region which results in good agreement with independent  
437 geophysical information about the same region collected by mean of different techniques. This paper  
438 indicates therefore that the computing time invested in the exploration of distribution of limited inter-  
439 est, like the high temperature distribution sampled by the particles along the PT, or in sampling the  
440 different unit-lag covariance matrices, which on their own are of scarce interest, is profitably spent as  
441 it allows accurate results in a limited time frame. Most notably this is the first study to the author's  
442 knowledge to investigate efficiency of an SMC method to track environmental changes as estimated from  
443 spherical reflection coefficients in shallow waters. The presented algorithm provides evident advantages  
444 for a potentially broad set of ocean scientists who need an efficient tool to explore large datasets and  
445 for those field operators, as the ones setting sonar experiments for the detection and classification of  
446 unexploded ordinances (UXO) who benefits from the precedent knowledge of geoacoustic parameter in  
447 the regions of interest. As noted earlier, it will be important to have a better understanding of the error  
448 statistics, which result to be key factor in inversion procedure developed, and studies on this aspect are  
449 advisable for the future.

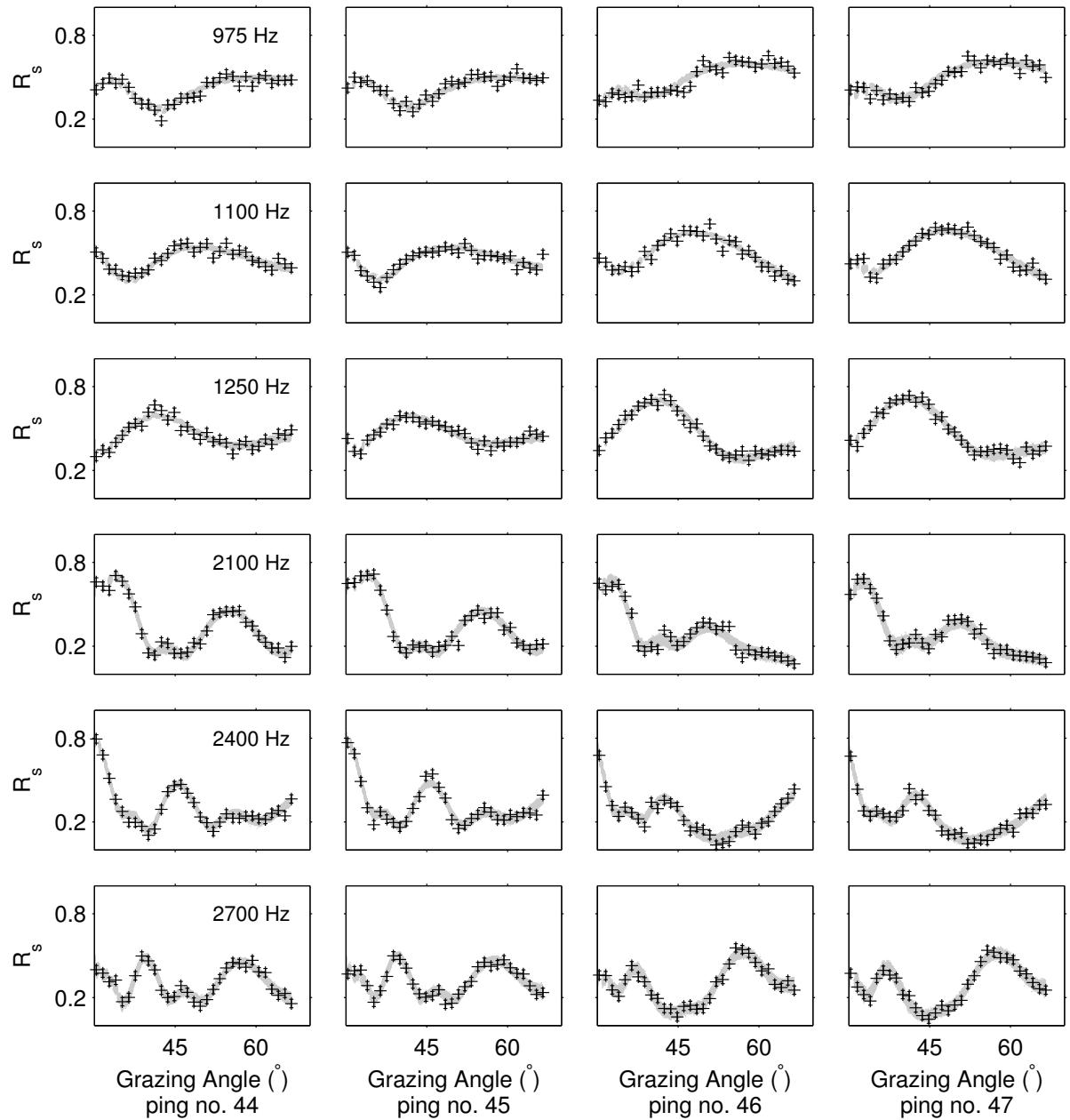
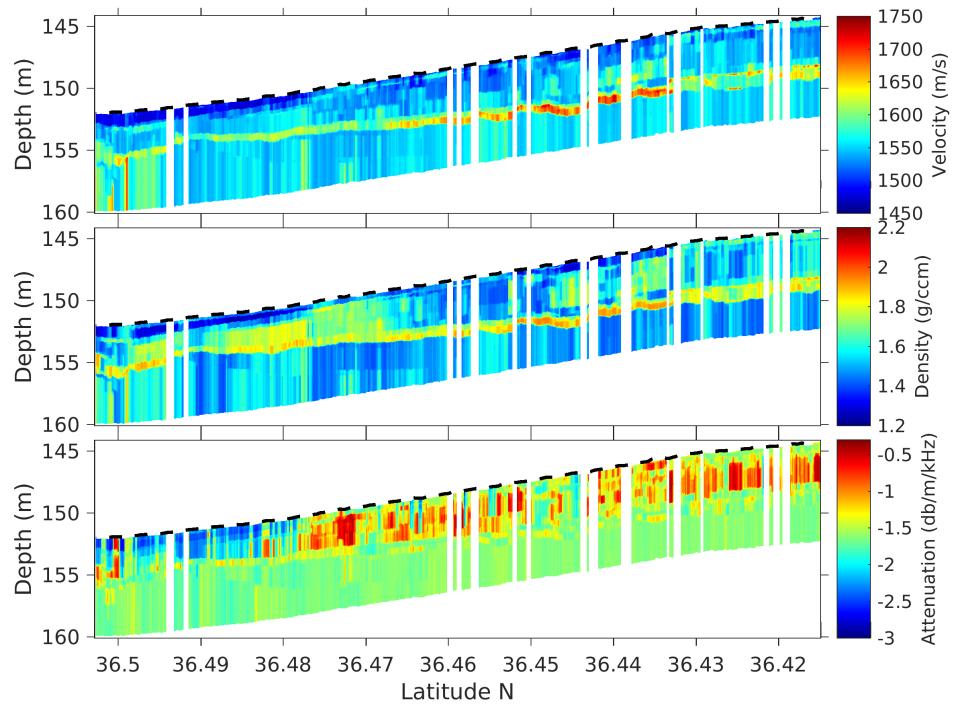
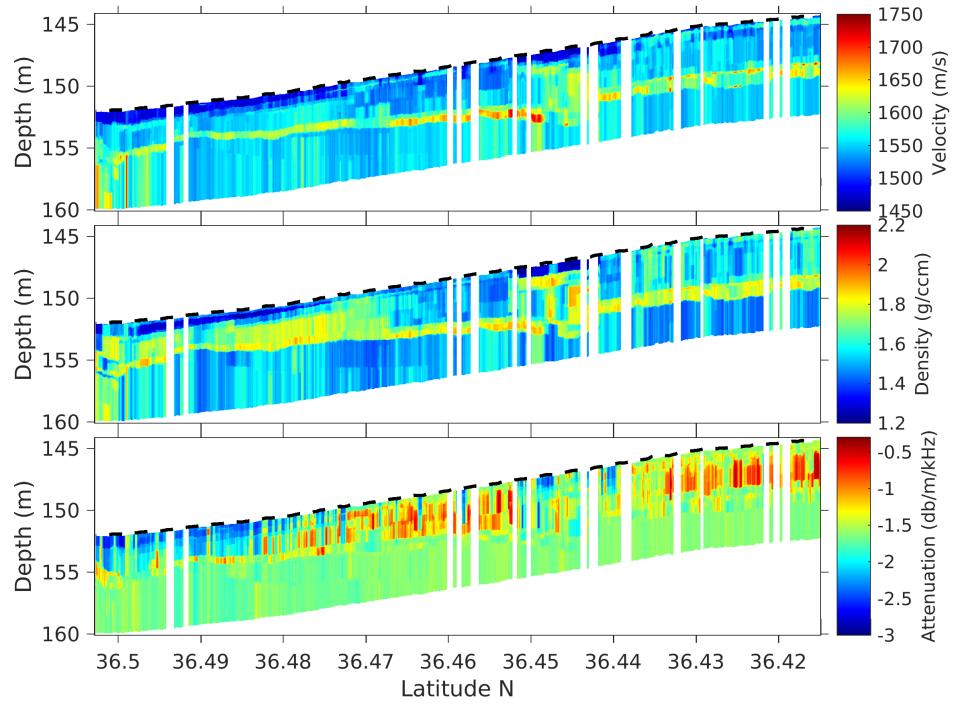


Figure 9: Data fit for ping 44, 45, 46 and 47 for the six frequency bands considered in the simulation. Frequency bands are labeled on the left column.



(a) Cauchy proposal.



(b) unit-lag PC proposal

Figure 10: Mean of marginal distributions for geoacoustic parameters plotted consistently with bathymetry.

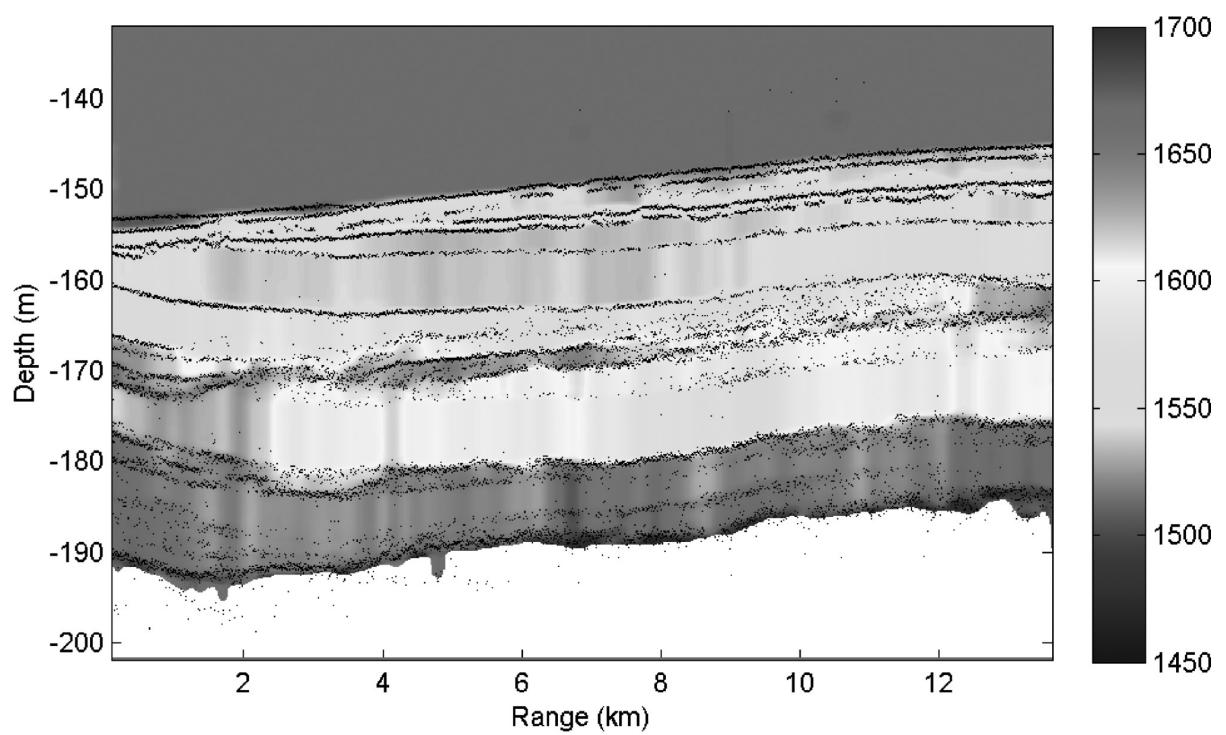


Figure 11: Sound velocity profile inferred with the image source method (from [18]).

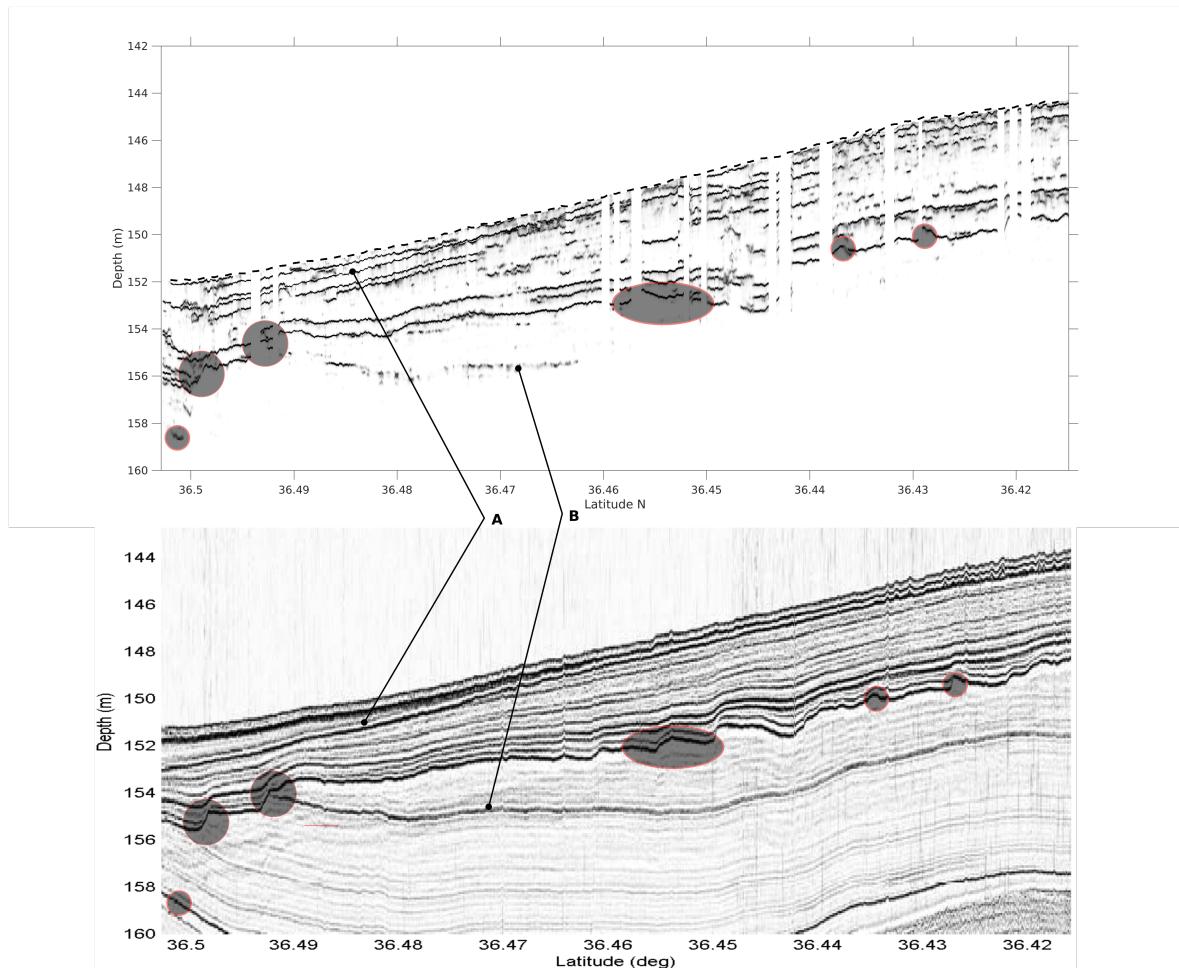
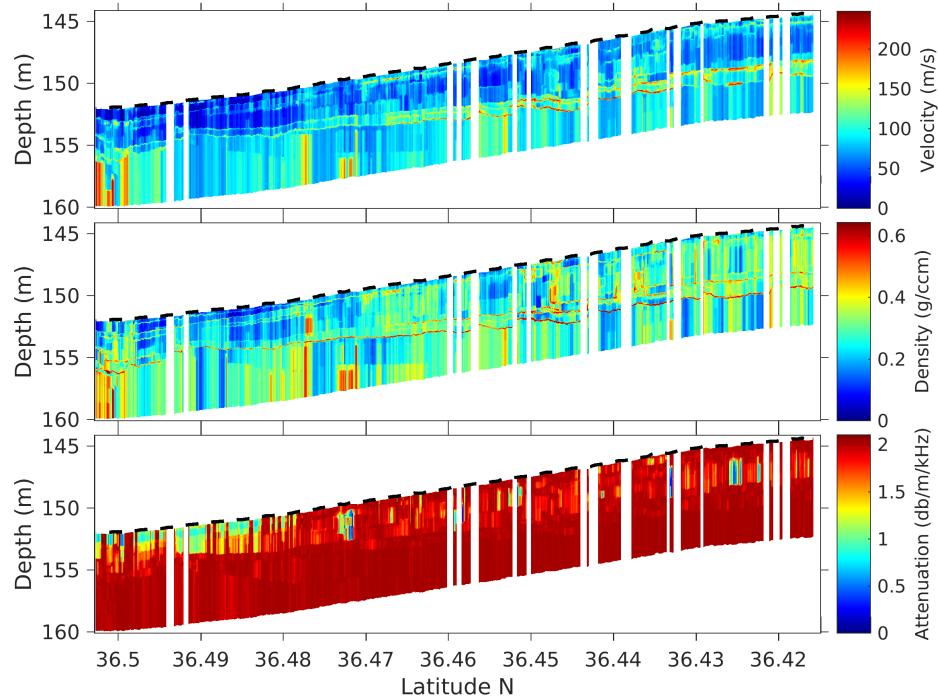
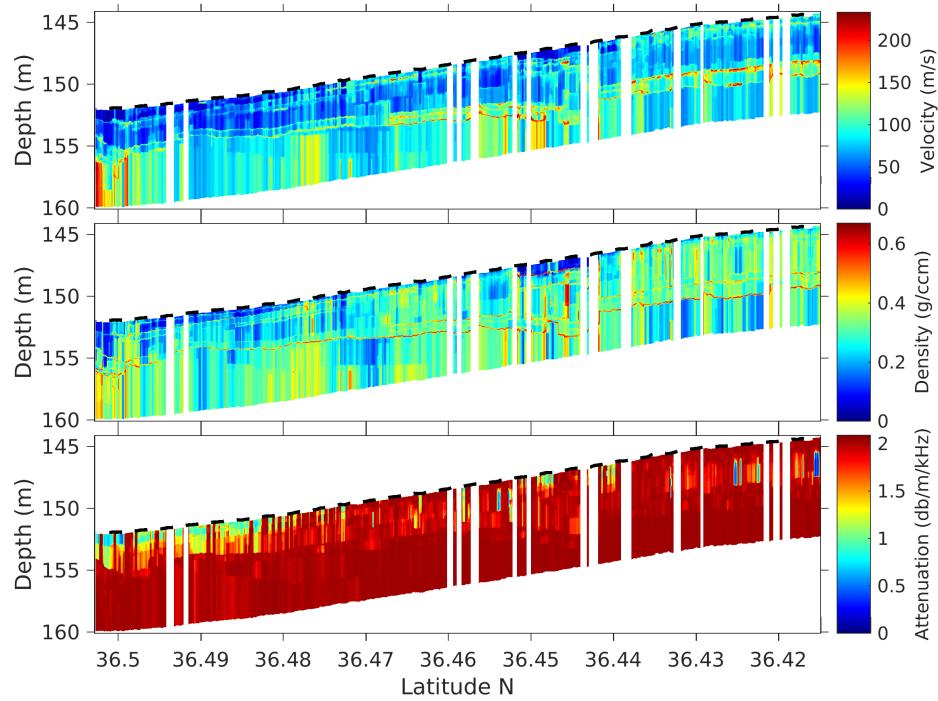


Figure 12: Comparison between interface depth distribution (top panel) and high resolution seismic profile (bottom panel).



(a) Cauchy proposal.



(b) unit-lag PC proposal

Figure 13: Length of the 95% confidence intervals of marginal distributions for geoacoustic parameters plotted consistently with bathymetry.

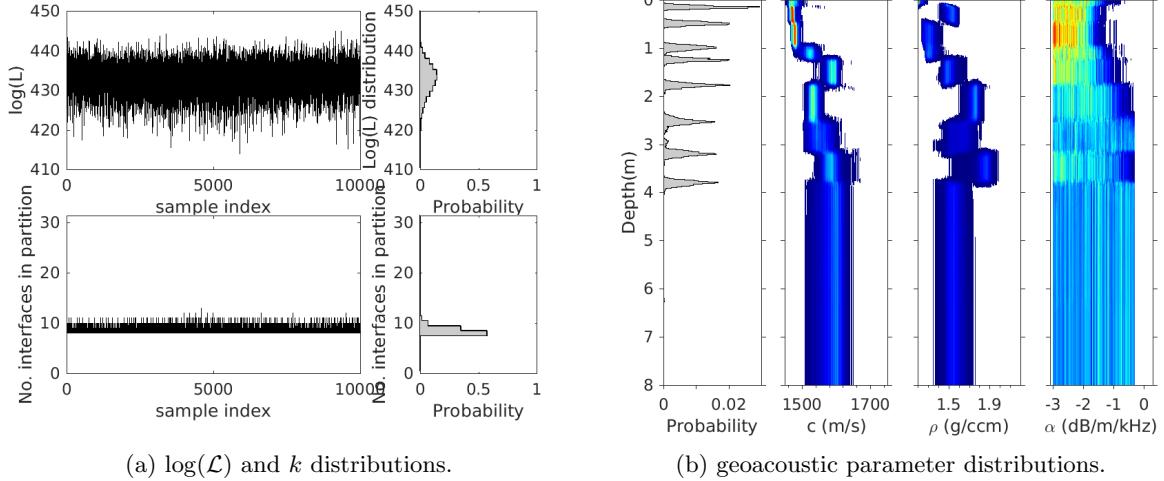


Figure 14: **(a)**:  $\log(\mathcal{L})$  (top row) and number of interfaces (bottom row) time series (left side) and distributions (right side) relative to the 262<sup>nd</sup> ping along the track. **(b)**: marginal probability distribution with depth for (left to right) interfaces position, sound velocity ( $c$ ), density ( $\rho$ ) and attenuation ( $\alpha$ ).

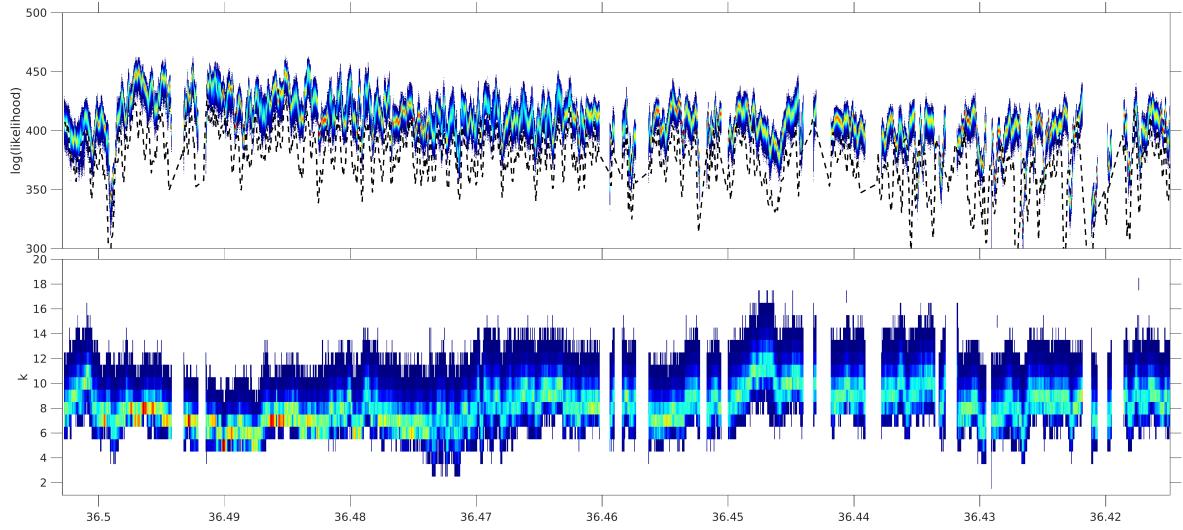


Figure 15:  $\log(\mathcal{L})$  (top panel) and marginal probability distribution for the number of interfaces (bottom panel) along the track. The dashed line marks the expected value for the  $\log(\mathcal{L})$  as computed by Eq. (10)

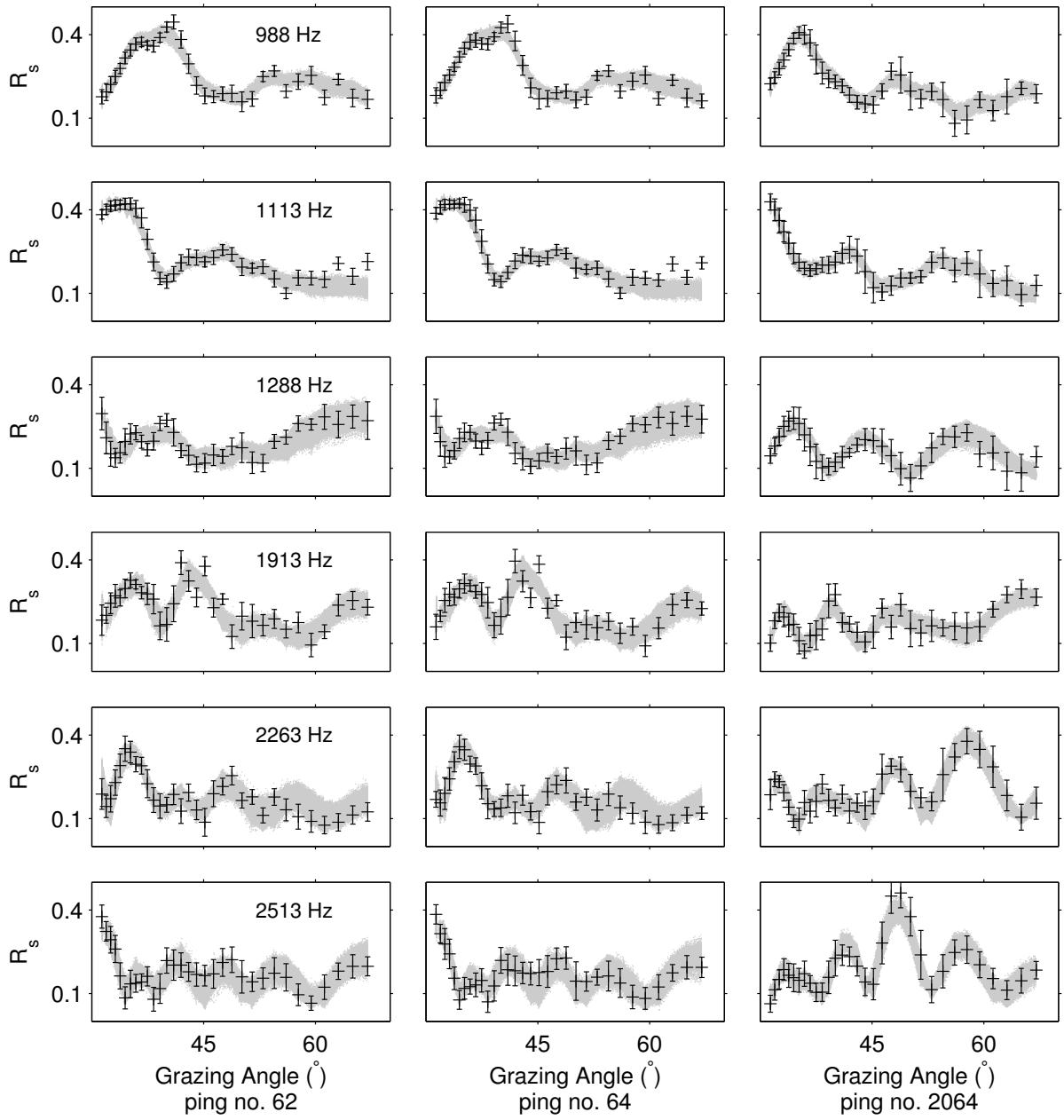


Figure 16: Data fit for ping 62 (left column of panels), 64 (central column of panels) and 2064 (right column of panels) as obtained from distribution estimated with Cauchy proposal density for the six frequency bands considered (labelled in left column of panels).

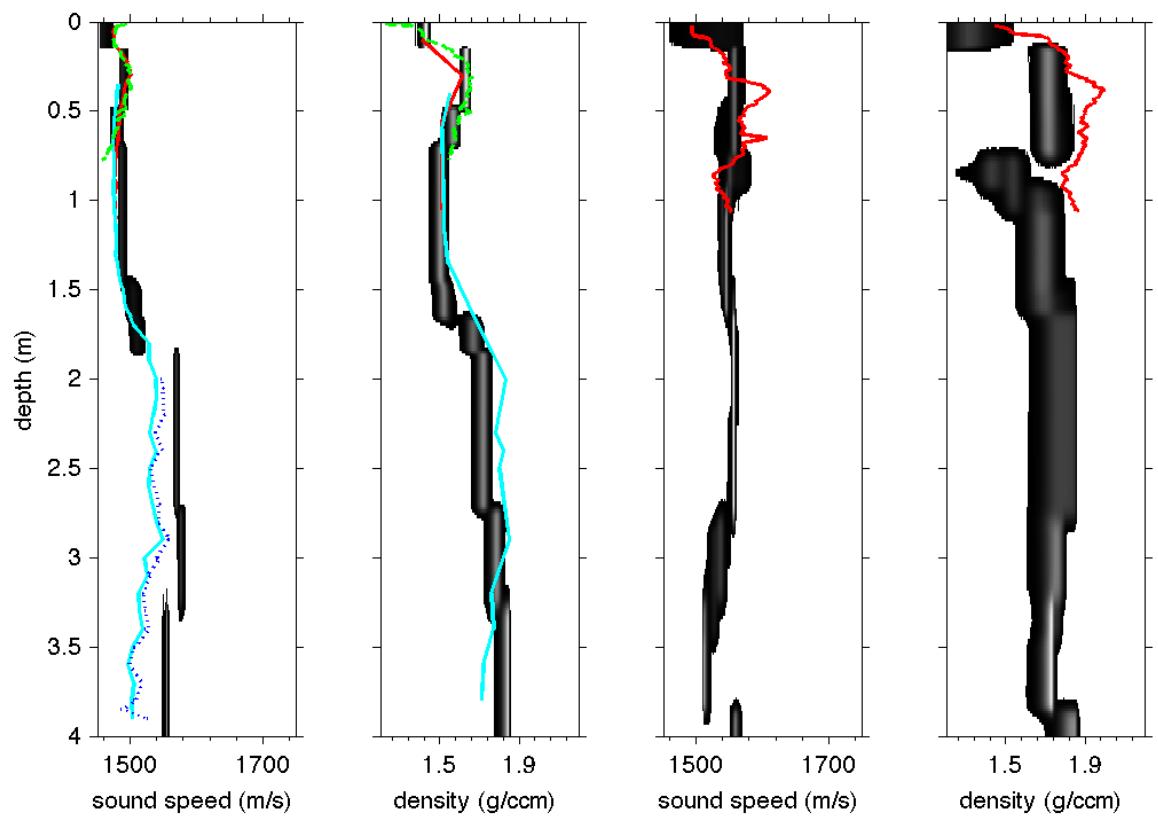


Figure 17: Sound velocity and density marginal distributions compared with well log data collected at the NW (left panels) and SE (right panels) extremes of the track.

450 **References**

- 451 [1] Caglar Yardim, Peter Gerstoft, and William S Hodgkiss. Tracking of geoacoustic parameters using  
452 kalman and particle filters. *The Journal of the Acoustical Society of America*, 125(2):746–760, 2009.
- 453 [2] Jan Dettmer, Stan E Dosso, and Charles W Holland. Sequential trans-dimensional monte carlo  
454 methods for range-dependent geoacoustic inversion. *The Journal of the Acoustical Society of Amer-*  
455 *ica*, 129(4):2599–2599, 2011.
- 456 [3] Charles W Holland, Peter L Nielsen, Jan Dettmer, and Stan Dosso. Resolving meso-scale seabed  
457 variability using reflection measurements from an autonomous underwater vehicle. *The Journal of*  
458 *the Acoustical Society of America*, 131(2):1066–1078, 2012.
- 459 [4] J E Quijano, S E Dosso, J Dettmer, and C W Holland. Fast computation of seabed spherical-wave  
460 reflection coefficients in geoacoustic inversion. *The Journal of the Acoustical Society of America*,  
461 138(4):2106–2117, 2015.
- 462 [5] F B Jensen, W A Kuperman, M B Porter, and H Schmidt. *Computational ocean acoustics*. Springer  
463 Science & Business Media, 2000.
- 464 [6] Peter J Green. Reversible jump markov chain monte carlo computation and bayesian model deter-  
465 mination. *Biometrika*, pages 711–732, 1995.
- 466 [7] N. Piana Agostinetti and A. Malinverno. Receiver function inversion by trans-dimensional monte  
467 carlo sampling. *Geophysical Journal International*, 181(2):858, 2010.
- 468 [8] William H Press. *FORTRAN Numerical Recipes: Numerical recipes in FORTRAN 90*. Cambridge  
469 University Press, 1996.
- 470 [9] Gavin Steininger, Jan Dettmer, Stan E Dosso, and Charles W Holland. Trans-dimensional joint  
471 inversion of seabed scattering and reflection data. *The Journal of the Acoustical Society of America*,  
472 133(3):1347–1357, 2013.
- 473 [10] Stan E Dosso, Jan Dettmer, Gavin Steininger, and Charles W Holland. Efficient trans-dimensional  
474 bayesian inversion for geoacoustic profile estimation. *Inverse Problems*, 30(11):114018, 2014.
- 475 [11] Edwin L Hamilton. Geoacoustic modeling of the sea floor. *The Journal of the Acoustical Society of*  
476 *America*, 68(5):1313–1340, 1980.
- 477 [12] Edwin L Hamilton and Richard T Bachman. Sound velocity and related properties of marine  
478 sediments. *The Journal of the Acoustical Society of America*, 72(6):1891–1904, 1982.
- 479 [13] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of markov chain monte*  
480 *carlo*. CRC press, 2011.

- 481 [14] CJ Geyer. Markov chain monte carlo maximum likelihood. In *Computing science and statistics:*  
482       *Proceedings of 23rd Symposium on the Interface Interface Foundation, Fairfax Station, 1991*, pages  
483       156–163, 1991.
- 484 [15] Stan E Dosso, Charles W Holland, and Malcolm Sambridge. Parallel tempering for strongly non-  
485       linear geoacoustic inversion. *The Journal of the Acoustical Society of America*, 132(5):3030–3040,  
486       2012.
- 487 [16] Jan Dettmer and Stan E Dosso. Trans-dimensional matched-field geoacoustic inversion with hi-  
488       erarchical error models and interacting markov chains. *The Journal of the Acoustical Society of*  
489       *America*, 132(4):2239–2250, 2012.
- 490 [17] Charles W Holland and Jan Dettmer. In situ sediment dispersion estimates in the presence of  
491       discrete layers and gradients. *The Journal of the Acoustical Society of America*, 133(1):50–61, 2013.
- 492 [18] Samuel Pinson, Laurent Guillon, and CW Holland. Range dependent sediment sound speed profile  
493       measurements using the image source method. *The Journal of the Acoustical Society of America*,  
494       134(1):156–165, 2013.