Journal of Applied Computational Intelligence

eISSN: 3136-6937

Vol. 4, No. 1, January, 2026

A Gaussian Process Model for Ocean Sound Speed Profile Inversion Incorporating Ray Travel Time Kernels

Tahani Al-Ghafli[1,*], Hessa Al-Mansoori[1], Karima Al-Riyami[2]

[1] College of Technological Innovation, Zayed University, Dubai, PO Box 4783, United Arab Emirates
[2] College of Computing and Engineering, British University in Dubai, Dubai, PO Box 345015, United Arab Emirates

Received 10 November 2025; Revised 02 March 2026; Accepted 09 March 2026; Published 12 March 2026

Abstract

Ocean sound speed profile inversion from acoustic travel times is limited by uneven depth sensitivity and the small number of distinguishable propagation paths. This study develops a Gaussian process formulation in which ray travel time kernels act directly on a depth covariance model, allowing profile estimates and conditional uncertainty to be computed within the same linear inverse problem. The evaluation uses a one-kilometre water column, a ten-kilometre transmission range, twenty-four idealized reflected paths, and eighty sound speed anomaly profiles. The observation noise standard deviation is 0.3 milliseconds. The proposed model is compared with independent depth regularization, an excessively smooth Gaussian process, a short correlation model, and an oracle covariance reference. Reconstruction error, interval coverage, timing residuals, and depth dependent uncertainty are evaluated using identical observations for every method. The kernel-based model attains a root mean square error of 1.227 metres per second and pointwise interval coverage of 86.68 percent; the independent depth reference produces an error of 1.450 metres per second. Sensitivity experiments examine six noise levels, six prior correlation lengths, and six retained path counts. The resulting analysis separates prior assumptions from acoustic information and identifies conditions under which small timing residuals coexist with substantial profile error. The model provides an interpretable probabilistic baseline for profile reconstruction, while its present validation remains restricted to fixed straight reflected paths and Gaussian perturbations. The controlled benchmark uses a deliberately mismatched prior; empirical coverage below the nominal 95 percent level identifies a calibration limitation alongside the reconstruction improvement.

Keywords: Ray Travel Time Kernels, Ocean Sound Speed Profiles, Gaussian Process Inversion, Acoustic Tomography, Posterior Uncertainty

1. Introduction

Ocean acoustic tomography retrieves information about large-scale ocean interior variability from path-integrated acoustic observations [1]. Ray-based ocean acoustic tomography estimates sound speed profile variations using eigenray travel times and linearized forward models [2]. Seismic ocean thermometry uses travel-time sensitivity kernels to relate acoustic delays to path-integrated temperature variations [3]. Sparse observations and limited environmental information make sound speed profile inversion an ill-posed problem [4]. A sensitivity kernel can be used to localize the support of an acoustic observation in the environment more precisely. The difference will be more significant if several arrivals have travelled along the same path and thus provide less independent information. Boundary interactions, arrival overlap and incomplete depth coverage influence the information available in shallow-water waveguides. Research on the physical constraints of shallow-water travel-time tomography is related to the above problems of recoverability and

* Corresponding author: tahani.gh@zu.ac.ae

http://dx.doi.org/10.68406/jaci.2026.vol4iss1no9:111-128

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

observation design. An engineering inversion needs to find the vertical structures that are constrained by the available travel time and quantify the uncertainty of these constraints. This assessment requires separate evaluation of vertical information and conditional uncertainty. A profile can reproduce the observations by placing physically different anomalies at depths where their sensitivities are compensated. Operationally, this ambiguity will alter the interpretation of the inferred thermocline and reduce the reliability of propagation predictions based on the estimated profile. The present study treats this ambiguity as an intrinsic property of the measurement system.

Traditional acoustic inversions often address this ambiguity by reducing the number of profile parameters or by penalizing deviations from a fixed reference. Empirical orthogonal functions (EOFs) restrict the reconstruction to a subspace learned from environmental samples. Such a restriction can be effective within the sampled variability range but may suppress structures poorly represented in the training archive. Variational regularization incorporates prior smoothness through penalty terms in linear inverse problems [5]. Its effectiveness depends on compatibility between the imposed scale and the environmental structures present. Excessive smoothing can suppress localized anomalies that contribute measurable travel-time changes. Conversely, an overly flexible model may attribute the observed noise to oscillations in depth. Bayesian inverse formulations combine prior information with an observation model to obtain posterior distributions that describe reconstruction uncertainty [6]. The resulting intervals are determined by both the assumed forward operator and the prior model. Errors of arrival association, receiver position and reference path geometry need to be added to the uncertainty model. Therefore, the observation operator and the prior depth covariance and error model need to be explicitly introduced in the probabilistic inversion. Gaussian process regression can represent depth-dependent sound speed uncertainty and its gradient in stochastic ray propagation [7]. They transform the depth covariance using the same linear functionals that describe the acoustic observations. Therefore, the modification of that structure requires separating the constrained data structure from the prior-supplied structure.

Recent machine-learning methods have expanded the approaches available for inferring environmental properties from acoustic measurements. Differentiable neural operators provide data-driven surrogate models that can be embedded in optimization for inverse problems [8]. Evaluation must therefore consider training coverage, observation geometry and uncertainty transfer. A model trained on a single family of reference environments may have a different arrival-profile perturbation relationship in the deployed environment. PDE-informed Gaussian process priors incorporate physical model structure into surrogate-based Bayesian inversion [9]. An integral sensitivity operator is used to connect the depth structure with the observations in acoustic profile inversion. Gaussian process regression requires an explicit observation operator to connect path labels with the contribution of depth perturbations to observed delays. The mean and covariance of a Gaussian process can be propagated through closed, densely-defined linear operators under suitable regularity conditions [10]. Thus, a natural way to introduce travel-time kernels has been provided. A further research need is a transparent evaluation that distinguishes profile accuracy, timing fit and uncertainty calibration under controlled observation conditions. In linearized electromagnetic source localization, Bayesian resolution analysis characterizes localization bias and ambiguities associated with the observation operator [11]. We use this perspective to examine the structures constrained by a kernel-informed Gaussian process and the sensitivity of its conditional intervals to covariance assumptions. It is a reproducible formulation and diagnostic benchmark that can be used to test the effect of acoustic support on prior correlation directly.

To study the joint effect of acoustic-path sensitivity and vertical environmental correlation on profile recovery, a Gaussian process inversion has been established that directly incorporates ray travel-time kernels into the relationship between depth-dependent sound-speed anomalies and observed arrival delays. The analysis addresses uneven depth support, correlation among arrival sensitivities, and the relationship between posterior precision and reconstruction error. Reference anomalies at all depths are used for the evaluation of profile error independently of acoustic residuals. The evaluation deliberately uses different reference and inversion covariances. A corresponding covariance reference is provided to offer a diagnostic comparison in the assumed statistical family. The experimental design isolates linear inversion in a fixed-path-geometry scenario. This design isolates conditioning from additional sources of error. The contribution lies in an inspectable acoustic benchmark and its diagnostics; conditioning a Gaussian prior on linear observations supplies the established inference framework. Cross-domain inverse-problem studies provide methodological context for covariance propagation and resolution, while the acoustic benchmark supplies the domain-specific evaluation. The article contains five sections. Section 2 reviews travel-time sensitivity methods, probabilistic profile reconstruction and evaluation practices. Section 3 introduces the observation operator, covariance construction, posterior calculation and reproducible computational procedure. Section 4 describes the evaluation data and analyzes reconstruction accuracy, interval calibration, sensitivity to prior scale and timing noise, path availability, residual structure, and depth-dependent error. Section 5 concludes the paper, identifies limitations of the current model, and outlines directions for future research, such as expanding the propagation model and collecting environmental data independently.

2. Related Work

2.1. Travel time sensitivity and environmental resolution

The connection between acoustic travel time and environmental changes has motivated the development of several tomography methods. Ray formulations are local variations of slowness along the reference propagation path that produce observation operators which are relatively simple to interpret when the arrival can be reliably identified. Liu et al. [12] constructed Gaussian process kernels from normal-mode depth functions for acoustic-field denoising and interpolation across receiving depths. Path integration can obscure vertical fluctuations whose positive and negative contributions cancel. Independent resolution of a depth interval depends on contrasts among the complete observation sensitivities. Recoverability is based on the differences in complete sensitivity patterns of the retained observations. Based on the above observations, the full observation matrix can be regarded as a set of linear functionals, each with full depth sensitivity. Increasing the number of arrivals with similar sensitivities may provide only a modest improvement in profile reconstruction.

A finite-frequency method replaces an infinitesimally thin path with a spatially varying distribution obtained from a wave-based description. Frequency-dependent travel-time sensitivity kernels distinguish contributions from different vertical temperature modes in seismic ocean tomography [13]. Such kernels can account for the effects outside the geometrical ray when the high-frequency approximation is inadequate. However, their interpretations are based on the observable features of the received waveform. Peak delays, phase estimates and correlation-based time shifts have different sensitivity distributions. The current model employs an idealized ray kernel, and wave-based effects have been omitted from the observation operator. A general linear kernel can be used for the statistical formulation, but its validity must be verified independently. This distinction separates verification of covariance propagation from validation of the underlying propagation physics.

Laboratory experiments in a scaled shallow-water ultrasonic waveguide have compared travel-time inversions based on ray theory and diffraction-based sensitivity kernels [14]. Such experiments permit prescribed environmental changes and repeatable measurement geometry. They can also be used to identify which parts of the acquisition process have been simplified in a purely numerical way. Waveform interference, missed detections and peak association errors are all problems of the acquisition model. A controlled experiment can study the effect of inversion conditioning alone. Separation is required for the interpretation of improved reconstruction accuracy. An estimator may perform well for correctly labelled delays but is vulnerable to a small number of misidentified arrivals in an operational deployment. The research gap concerns the interpretability of information and uncertainty as the measurement model becomes more realistic.

2.2. Prior models and Gaussian process inverse inference

Low-dimensional environmental representations are computationally efficient and can encode recurring vertical patterns. Empirical modes are suitable for archives with significant seasonal and regional variations, and spline and layered parameterizations can be used to control the depth resolution directly. Bayesian geoacoustic inversion combines prior information with acoustic backscattering measurements and evaluates parameter uncertainty and posterior correlations [15]. A central interpretive issue is separating prior information from information supplied by the observations. The posterior estimate can vary gradually in an unobserved depth interval due to the prior's correlation with observed regions in that interval. The covariance model supplies the corresponding smooth conditional transition. It is based on the assumption of an interval correlation. Reporting conditional variance alongside the estimated profile exposes this distinction when the underlying uncertainty assumptions are specified.

A Gaussian process represents a random function through its mean and covariance and permits conditioning on linear observations. Gaussian measure conditioning supports sequential assimilation of linear operator observations and updates the corresponding posterior mean and covariance [16]. For sound speed reconstruction, it is possible to obtain the acoustic observation covariance from the depth covariance via the travel-time-sensitivity operator. The cross-covariance of a depth location and the observation set is then used to update the reference profile. A related interpretation is Bayesian regularization of a discretized inverse problem. The Gaussian process formulation determines how prior correlation distributes information from the observation operator across depth. Covariance families can be compared by examining their effects on reconstruction accuracy and conditional uncertainty. A long correlation scale may reduce variance at the cost of increased bias for localized structures, and a short scale can lead to significant profile uncertainty due to path-integrated measurements of depth cells.

Covariance parameters can be set based on independent hydrographic knowledge or estimated from the acoustic data. Log-marginal-likelihood optimization estimates Gaussian process kernel hyperparameters, while the selected covariance smoothness influences the recovered field [17]. Parameter identifiability depends on the available path combinations: changes in prior scale and noise variance can produce similar observation covariances. A sharply defined numerical optimum requires independent assessment of depth identifiability. Evaluation should examine sensitivity to plausible

covariance parameters while separating parameter selection from test profile assessment. The proposed prior is fixed before test reconstruction, and its length scale is varied only in the designated sensitivity analysis. The matched covariance model is referred to as the oracle reference and has privileged access to the exact reference covariance.

2.3. Learning based reconstruction and validation practice

Saha et al. [18] combined generative learning with ray-based ocean acoustic tomography through a variational autoencoder representation of sound speed profiles and dynamical regularization. Direct mapping can be computationally efficient after training; comparison with model-based inversion requires an account of the size and distribution of the training archive. Differences in geometry, noise, reference profiles and output parameterization can change the ranking. Therefore, the engineering evaluation should distinguish between an algorithmic advantage and an advantage resulting from additional environmental information. An exactly known reference covariance is a type of privileged information. If it is to be considered a learned model, then the training process that provides this knowledge must be specified. The same acoustic observations and the same prior marginal variance are used for the comparisons in this paper to make the effects of depth correlation more prominent.

Uncertainty validation requires repeated assessment across profiles in addition to individual reconstruction plots. In inverse imaging, posterior variance estimates can be combined with conformal prediction to construct error bounds with coverage guarantees [19]. In environmental inversion, the unit of replication is affected by depth samples in the same profile. Depth locations within a profile are correlated, making the profile the appropriate unit of independent replication. The distributions of the profile-level errors and the pointwise interval coverages are different, so they are presented separately. Ensemble methods are used to perform both types of tests under the specified variability family. Transfer to a sea trial would require independently measured profiles, the uncertainty of those measurements, and acquisition conditions that are sufficiently well-documented to distinguish between environmental errors and forward-model errors. The present benchmark provides a reproducible evaluation of a kernel-integrated Gaussian process under controlled covariance mismatch.

3. Kernel Integrated Gaussian Process Formulation

3.1. Observation geometry and linear travel time kernels

The unknown field is the depth dependent sound speed anomaly relative to a prescribed reference profile. We assume horizontal uniformity, a flat surface and seabed, and fixed propagation paths during inversion. Each observation represents an arrival delay relative to the reference travel time. The reference path may be obtained from a ray solver; the present benchmark uses straight specularly reflected paths constructed through the image source method. This configuration isolates linear functional conditioning under fixed geometry. Refracted propagation requires a corresponding path solver. Sound speed is expressed in metres per second, depth and path length in metres, and travel time in seconds. Consistent units are essential because the observation kernel connects acoustic delays with environmental perturbations.

$$t_i(c) = \int_{\Gamma_i} \frac{1}{c(z(s))} ds, \delta t_i \approx -\int_{\Gamma_i} \frac{\delta c(z(s))}{c_0(z(s))^2}\, ds \tag{1}$$

The linearized travel time equation follows from differentiating reciprocal sound speed while holding the propagation path fixed. Its negative sign expresses the earlier arrival associated with a positive sound speed perturbation. The approximation retains first order anomaly contributions under fixed path geometry. Larger anomalies and changes in path identity require a separate validity assessment. The present calculation uses a common linear operator throughout the evaluation, so the reconstruction errors characterize the statistical inverse problem under that observation model. Assessing nonlinear effects would require independent travel time calculations and a path association procedure. Because the evaluation retains the same fixed linear operator, the reported errors assess conditioning and covariance mismatch; independent nonlinear travel-time calculations are required to quantify linearization error.

The computational depth grid spans the water column at uniform spacing. Path length contributions are accumulated into the nearest depth cell, with half weights assigned to the first and last path samples for endpoint quadrature. Each cell contribution is multiplied by the negative reciprocal square of the reference sound speed. The resulting matrix maps grid anomalies to arrival delays. Its columns describe sensitivity to individual depth cells, while its rows describe the complete depth footprint of individual arrivals. The matrix is rectangular, with more depth unknowns than travel time observations. Regularization therefore selects among admissible perturbations, and the prior covariance supplies a statistical preference among profiles with similar acoustic signatures. For the main experiment, the observation matrix has 24 rows and 101 columns; its rank is at most 24, so depth-grid refinement alone cannot increase the independent information supplied by these delays. Each accumulated path-length weight is measured in metres; the corresponding sensitivity entry therefore has units of seconds per metre per second of sound-speed perturbation.

$$\boldsymbol{y} = \boldsymbol{Hf} + \varepsilon, \varepsilon \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}), H_{ij} = -\frac{w_{ij}}{c_0(z_j)^2} \quad (2)$$

In the observation equation, **f** contains the depth anomalies, **y** contains the measured delays, and w_{ij} denotes the accumulated length of path i within depth cell j. Observation errors are independent in the main benchmark, giving a diagonal covariance matrix **R** with a common variance. Shared clock offsets or correlated arrival picking errors would require a structured covariance. The sign and units of each sensitivity entry can be checked independently of the inference procedure. A uniform positive anomaly shortens the modeled travel times, while doubling a cell's path length contribution doubles its local first order sensitivity.

Figure 1 presents the computational architecture connecting the reference water column, image source paths, depth sensitivity accumulation, covariance transformation, and posterior outputs. The diagram separates the physical observation operator from the statistical prior and distinguishes delay inputs from reconstructed profile outputs. This separation permits the propagation operator to be updated independently of the covariance assumptions. Changes to the prior alter the covariance update while preserving the units and physical meaning of the observations.

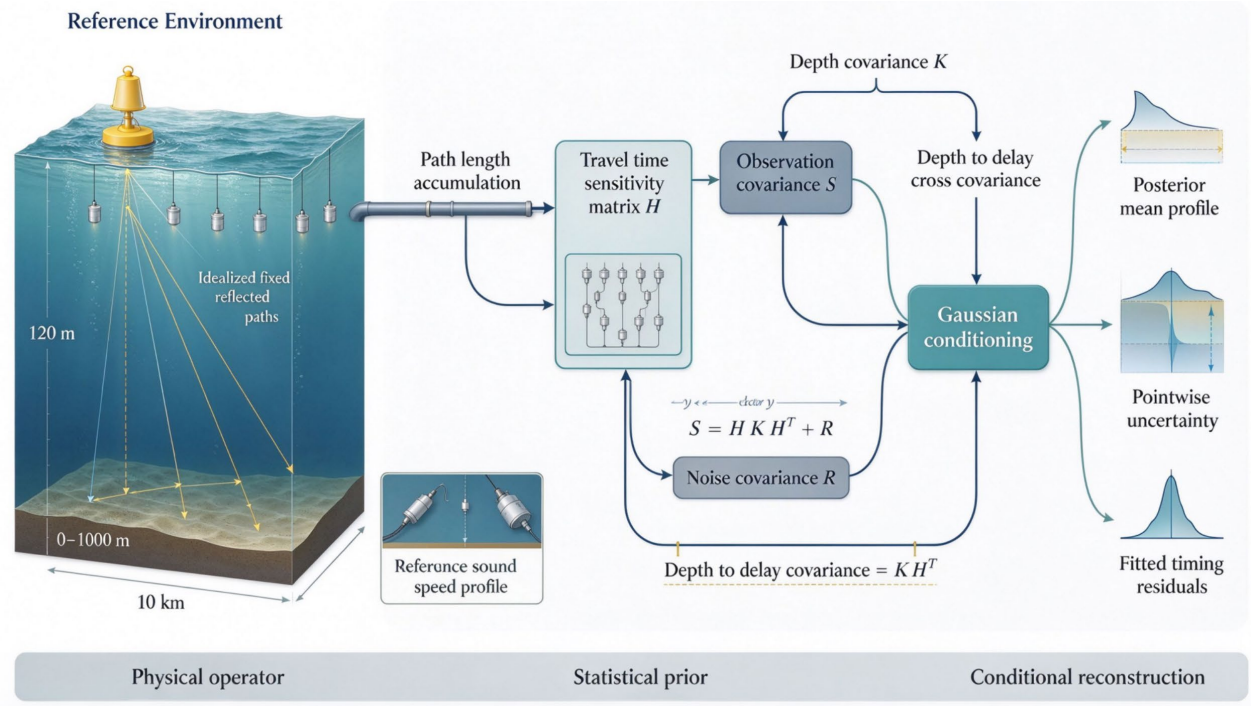


Figure 1. Architecture of the ray travel time kernel integrated Gaussian process inversion

3.2. Depth covariance and conditional profile inference

The anomaly field has a zero prior mean because the reference profile carries the deterministic background structure. A nonzero prior mean can be accommodated by subtracting its predicted delay before conditioning and restoring it to the reconstructed field afterward. A Matérn covariance with moderate local roughness represents vertical correlation. Its amplitude specifies the expected anomaly magnitude, and its length parameter controls the decay of covariance with depth separation. The covariance is positive definite on the depth grid and can be propagated through linear observation operators. The length parameter describes prior correlation, while acoustic resolving power also depends on the observation geometry and noise. The selected covariance is the Matérn family with smoothness parameter three halves; its length scale describes prior vertical dependence, while acoustic resolution also depends on geometry and observation noise.

$$k(z, z') = \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\ell}\right) exp\left(-\frac{\sqrt{3}r}{\ell}\right), r = |z - z'| \quad (3)$$

The covariance equation defines a stationary relationship between depth separation and environmental dependence. Here, σ_f is the prior anomaly standard deviation and ℓ is the vertical correlation length. The selected covariance accommodates local slope variations and vertically structured anomalies. Hydrographic stability and thermodynamic relationships would require additional prior constraints. The present design assigns an identical marginal variance to every depth. Extensions could introduce a depth varying amplitude or a change in correlation scale near a thermocline,

provided independent information supports the additional parameters. The stationary model supplies a controlled baseline for separating the effects of geometry and vertical correlation.

The depth covariance matrix **K** is obtained by evaluating $k(z, z')$ at the grid nodes. The corresponding observation covariance combines transformed environmental variability with measurement error. Correlated depth perturbations induce correlated arrival delays, so two observations can remain strongly dependent even when their receiver depths differ. This dependence is particularly pronounced when paths traverse similar depth intervals. A positive observation noise variance makes the observation covariance nonsingular in the present experiments. The update is evaluated through linear system solutions; inverse notation expresses the conditioning relationship compactly.

$$\boldsymbol{S} = \boldsymbol{HKH}^{\top} + \boldsymbol{R}, \boldsymbol{\mu} = \boldsymbol{KH}^{\top}\boldsymbol{S}^{-1}\boldsymbol{y} \tag{4}$$

Through the posterior mean equation, the delay vector updates the profile according to the depth to observation cross covariance and the dependence among arrival measurements. The resulting estimate is a weighted combination of these cross covariances. Both path sensitivity and observation correlation determine the weights. Prior dependence can transmit an update from sampled regions to depths outside direct path support, producing a conditional extrapolation. An excessively long correlation scale can spread a locally supported correction across a broad interval. An excessively short scale can leave localized structures weakly constrained by integrated delays. These effects motivate evaluating reconstruction accuracy together with uncertainty across several covariance choices.

Posterior covariance describes the remaining uncertainty under the assumed linear Gaussian model. With fixed geometry, covariance parameters, and observation noise, it is independent of the observed delay values. Different profile realizations consequently share the same conditional standard deviation in the main benchmark. This property is essential to interpreting uncertainty bands around individual reconstructions: their width is determined by the measurement geometry and statistical assumptions. Realized errors vary among profiles even when the conditional covariance is correctly specified. With fixed covariance parameters and geometry, conditional variance is common across profiles, whereas realized errors depend on each anomaly and delay realization.

$$\boldsymbol{C} = \boldsymbol{K} - \boldsymbol{KH}^{\top}\boldsymbol{S}^{-1}\boldsymbol{HK}, I_j = \left[\mu_j - 1.96\sqrt{C_{jj}}, \mu_j + 1.96\sqrt{C_{jj}}\right] \tag{5}$$

The uncertainty equation provides the posterior covariance and the associated pointwise 95 percent intervals. Simultaneous profile bands would require a joint probability construction across the depth grid. Interval calibration is evaluated empirically over the profile ensemble. Differences between the proposed prior and reference covariance can produce discrepancies between nominal and empirical coverage. The oracle model checks the conditioning implementation: when the assumed covariance and noise match the reference distribution, repeated sampling coverage should approach the nominal pointwise probability, subject to finite ensemble variability. This comparison assesses the internal consistency of the conditioning calculation. Pointwise coverage is the proportion of depth-profile pairs whose reference anomaly lies within its corresponding interval; simultaneous coverage would require a separate joint-profile criterion.

3.3. Computation diagnostics and reproducible evaluation

The implementation constructs the observation matrix once and applies each candidate covariance to the same delay vectors. Covariance parameters and evaluation profiles remain fixed during the main comparison. For each covariance, the observation system is solved for all profiles, followed by calculation of the conditional covariance. The observation matrix factorization is small because the number of arrival delays is limited. At finer depth resolution, storage and multiplication of the full depth covariance become more demanding. Structured covariance methods or basis approximations could reduce this cost. The present grid permits a direct implementation whose numerical operations remain accessible for inspection. Inverse notation in the conditioning equations denotes linear-system solutions; a numerical implementation should retain the observation covariance and solve for the required right-hand sides.

Profile error is evaluated against the reference anomaly at every depth. The principal accuracy statistic aggregates squared error over depth nodes and ensemble members. Additional metrics include mean absolute error, pointwise coverage, mean interval width, and fitted timing residual. Each metric addresses a distinct aspect of performance. A small residual describes agreement with the observations used for conditioning, while independent prediction requires additional observations. A small profile error indicates recovery within the specified variability family. High coverage can result from calibrated uncertainty or broad intervals, making interval width essential to its interpretation.

$$RMSE = \sqrt{\frac{1}{NM}\sum_{n=1}^{N}\sum_{j=1}^{M}\left(\mu_{jn} - f_{jn}\right)^2} \tag{6}$$

For the error equation, N denotes the number of independent profiles and M the number of depth nodes. The quantities μ_{jn} and f_{jn} are the reconstructed and reference anomalies at depth node j in profile n. Correlation among depth nodes makes the profile the appropriate unit of independent replication. Profile specific error distributions are retained to expose variability beyond the aggregate statistic. Standardized errors are also inspected to assess whether conditional standard deviations describe the observed reconstruction spread. Population uncertainty in an improvement would require resampling at the independent profile level.

Figure 2 describes the workflow connecting geometry assembly, profile and delay inputs, five fixed covariance models, diagnostic extraction, and figure export. One branch varies observation noise and prior correlation length, while another retains progressively larger subsets of the ordered path collection. These branches join at the reporting stage. Separating them from the main comparison preserves the distinction between sensitivity analysis and parameter selection. The principal model is fixed before the sensitivity surface is inspected, and the oracle branch is identified by its access to the exact reference covariance.

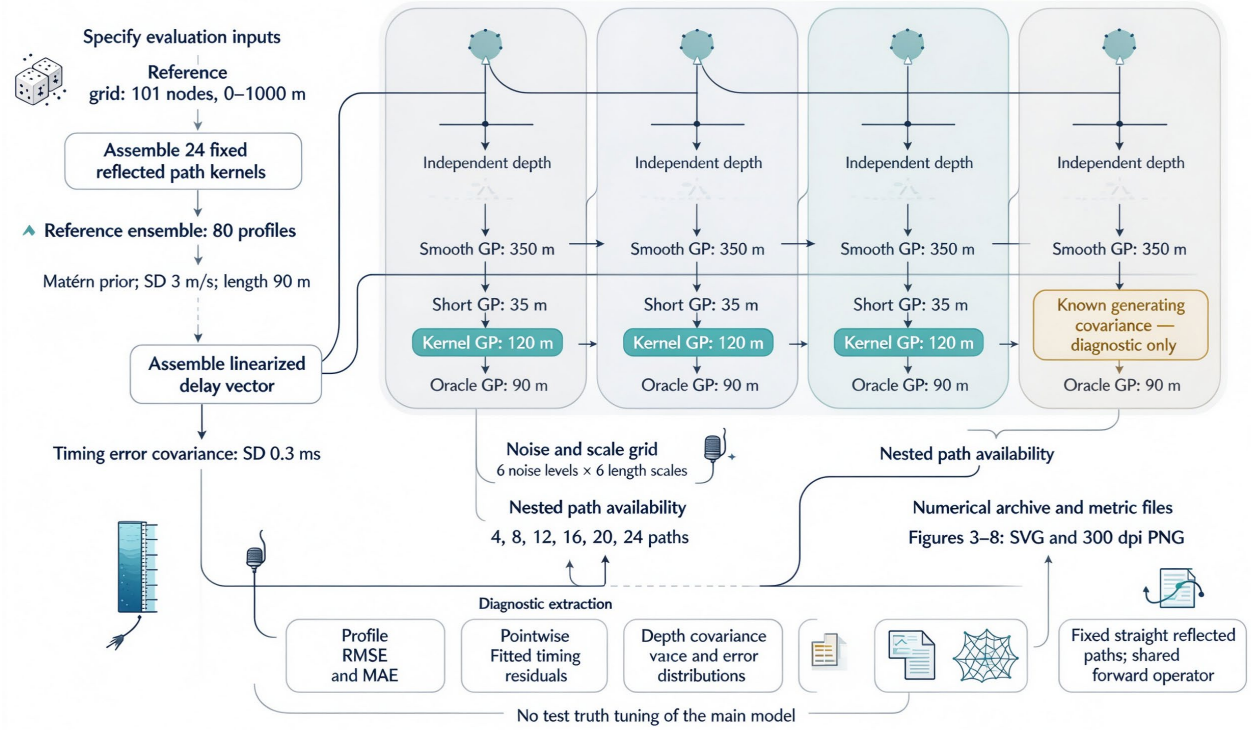


Figure 2. Evaluation and diagnostic workflow with separated sensitivity branches

The numerical archive stores the reference profile, depth grid, observation operator, anomaly ensemble, delay vectors, proposed covariance, posterior mean, and posterior standard deviation. Summary metrics and sensitivity arrays are retained in separate machine-readable files. The data figures use these arrays and their associated diagnostic calculations, maintaining a consistent relationship between the reported statistics and plotted quantities. Retaining profile and depth indices permits independent recalculation of ensemble errors, interval coverage, and depth dependent diagnostics. The archived operator also supports subsequent comparisons using alternative covariance models under identical observation conditions. Independent recalculation requires the depth grid, reference profile, path operator, anomaly ensemble, delay vectors, covariance settings and noise specification to be retained together.

4. Experiments and Analysis

4.1. Experimental data and profile reconstruction

A water column from the surface down to a depth of 1 km is taken for the evaluation, and 101 grid nodes with a spacing of 10 m are used. A deterministic reference profile is composed of a background vertical gradient and a large-scale subsurface sound-speed reduction. The source is at a depth of 120 metres, and the range of horizontal transmission is 10 kilometres. Receivers are positioned at six depths through the water column. Four image-receiver branches are arranged for each receiver, and thus 24 path functionals result. Specular reflection is represented by folding straight image-source paths into the water column. The construction has several depth weights and avoids the numerical problems of solving a nonlinear ray boundary value problem. The construction idealizes propagation geometry to isolate statistical conditioning. Reference-profile-based underwater acoustic positioning can use a precomputed ray grid to calculate travel times efficiently [20]. Consistency is maintained in the current experiment by using the same operator

for all evaluations. The marginal standard deviation of the 80 independent anomaly profiles in the ensemble is 3 m/s and the vertical correlation length is 90 m. The proposed inversion is based on the same marginal variance but has a correlation length of 120 metres. This deliberate mismatch separates the proposed prior from the exact reference covariance. The selected anomaly magnitude is small compared to the reference sound speed; an independent forward calculation is required to quantify nonlinear travel time errors. Each profile is associated with a vector of linearized delays, and the standard deviation of the observation error is 0.3 ms. All profiles are included in the total, and the first realization provides an example of the plot. Wavelet packet features, hidden Markov models and mixture density networks have been combined to infer marine environmental parameters from acoustic signals in noisy conditions [21]. The present evaluation retains the data underlying every plotted element in numerical arrays for diagnostic traceability. The main setup comprises six receiver depths, four image branches per receiver, 24 delay observations, 101 depth nodes and 80 independent profile realizations. The first archived realization supplies the single-profile illustration, while all 80 realizations contribute to the aggregate metrics.

Figure 3(a) displays five realized sound speed profiles and the reference on the 101-node depth grid. The 3 metre per second anomaly standard deviation determines the scale of the deviations, while the 90-metre reference correlation length permits localized vertical changes within the one-kilometre column. The plotted profiles retain the broad background gradient and subsurface depression but differ in the location and sign of smaller anomalies. These differences provide a controlled test of whether the inverse model recovers departures from the reference across several depths. The first five realizations are displayed in their archived order, preserving a reproducible connection between the figure and the full eighty profile ensemble. Their local excursions remain part of the specified Gaussian variability, and the deterministic reference carries the mean vertical structure.

Figure 3(b) displays sensitivity magnitudes for paths 1, 5, 9, 13, 17, and 21, expressed in microseconds per metre per second. Each curve describes the contribution of a depth cell to an integrated arrival delay. Differences in their vertical support identify the combinations of depth intervals available to the inversion. Shared peaks indicate that several observations respond to similar perturbations, whereas differences between curves supply information that can separate vertical contributions. The folded image paths cross intermediate intervals repeatedly, so the receiver depth and the largest sensitivity peak describe only part of the complete observation footprint. The stored sensitivity matrix retains the negative derivative associated with earlier arrivals at higher sound speed; the figure uses magnitudes to compare depth support on a common positive scale.

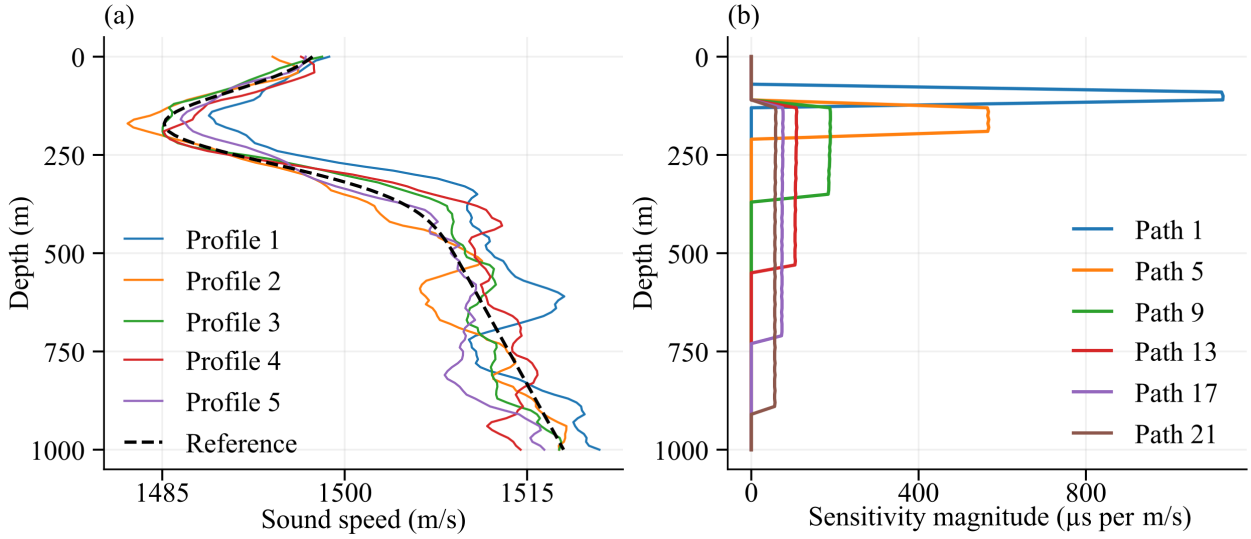


Figure 3. Environment and observation support. (a) Five realized profiles and the reference. (b) Six selected depth sensitivity magnitudes

Five covariance models are compared using identical observations. The independent depth reference uses equal marginal variances and a diagonal depth covariance. The smooth Gaussian process uses a squared exponential covariance with a three hundred and fifty metre scale, representing an intentionally broad prior. The short scale model uses an exponential covariance with a thirty-five-metre scale. The proposed kernel model uses the Matérn covariance at one hundred and twenty metres. The oracle model uses the ninety-metre reference Matérn covariance. These labels describe covariance choices within the same linear Gaussian conditioning algorithm. The comparison isolates covariance choices within a common conditioning algorithm. Convergence analyses of learned regularization relate the observation operator, denoiser structure and regularization strength to stable recovery as noise decreases [22]. Holding marginal variance and observations fixed here isolates a specific effect: how depth correlation changes the distribution of acoustic information across the profile. The independent depth reference isolates the effect of treating neighbouring nodes as

uncorrelated. A path delay constrains a weighted sum of many cell anomalies, leaving much of their relative arrangement unspecified. The posterior can reduce error in that weighted sum while retaining uncertainty in the corresponding vertical pattern. The broad smooth model represents the opposite difficulty. It links distant depths strongly and can represent large scale changes efficiently, yet it has limited freedom to reproduce shorter vertical structures. The short scale model supplies more local flexibility, with a covariance whose roughness differs from the reference variability model. The proposed model occupies an intermediate position and also differs from the truth in correlation length. The oracle isolates the penalty associated with this mismatch within the chosen Gaussian family. It provides a diagnostic reference with explicitly privileged covariance information. All five estimators use the same observation vectors, operator and Gaussian conditioning procedure; the comparison isolates covariance structure within that inference family. The oracle is a matched-covariance diagnostic with privileged environmental information, whereas the proposed prior retains the declared correlation-length mismatch. The short-scale reference uses an exponential covariance, whereas the proposed model uses a Matérn covariance of smoothness three halves; both correlation scale and covariance regularity differ.

The experiment evaluates delay observations against the corresponding reference profiles. The noise standard deviation describes perturbations added to already associated arrival delays. Arrival tracking, environmental drift during transmission, and uncertain source and receiver coordinates require additional error components. Errors of those kinds can be structured across paths and can invalidate a diagonal observation covariance. Segmented Bayesian acoustic-tomography calibration separates relative clock-offset estimation from the joint estimation of transducer positions and sound speed parameters [23]. A practical implementation would need either to estimate those components jointly or to marginalize them with a defensible prior. The present design separates acquisition error modeling from covariance induced profile uncertainty. The evaluation uses a common grid and path quadrature throughout the calculation. This removes numerical operator mismatch and is appropriate for checking the conditioning equations, but it favors every tested estimator relative to a practical acquisition chain. A subsequent benchmark could compare a finer forward grid or a different propagation model with the present inversion discretization. That experiment would distinguish statistical regularization error from discretization and forward modeling error. It would also require a structured observation error model to represent discrepancies that depend on path geometry. Reporting this distinction is necessary before attributing any future loss of coverage solely to the Gaussian process prior. Using a common grid and path quadrature isolates estimator behavior under operator consistency; an independently discretized forward calculation would assess the additional effect of numerical mismatch.

Figure 4(a) shows the first reference anomaly, its posterior mean, and the pointwise 95 percent interval at 101 depth nodes. The reconstruction follows the vertical components supported by the acoustic kernels, while the conditional mean attenuates structures whose contribution to the delay vector is weak relative to the declared prior variability. The interval width averages 3.547 metres per second across the depth grid. Its local shape depends on both the path collection and the covariance model, so a narrow region identifies strong conditional constraint under those assumptions. The first realization is retained as a fixed illustration, and the ensemble metrics establish how representative its error is of the complete experiment. The mean retains the components supported by the observations and the assumed prior. Conditioning minimizes expected squared error under the assumed Gaussian model and therefore suppresses components whose acoustic evidence is weak relative to their prior variability. This suppression can improve average reconstruction while producing visibly attenuated peaks in individual cases. The degree of attenuation depends on whether a fluctuation projects onto a well observed combination of depths. Two anomalies with similar maximum amplitudes can have different reconstruction quality if one produces a distinct delay pattern and the other lies near a weak singular direction of the observation matrix. A plot of truth and estimate should consequently be read together with the observation operator and uncertainty diagnostics. Identifying the source of improvement requires joint inspection of the reconstruction, observation operator and uncertainty diagnostics.

Figure 4(b) compares true and estimated anomalies for every fourth profile, producing 2020 depth samples from twenty realizations. The identity line provides a reference for the direction and magnitude of reconstruction error. The proposed model has an ensemble RMSE of 1.227 metres per second and a mean absolute error of 0.936 metres per second. Deviations toward the centre of the scatter distribution reflect attenuation of uncertain anomaly components; they arise when multiple depth patterns produce similar integrated delays. Colour encodes depth and makes it possible to inspect whether larger departures cluster within particular parts of the column. Depth samples within each realization share environmental correlation, so the twenty profiles supply the independent experimental units represented in this panel.

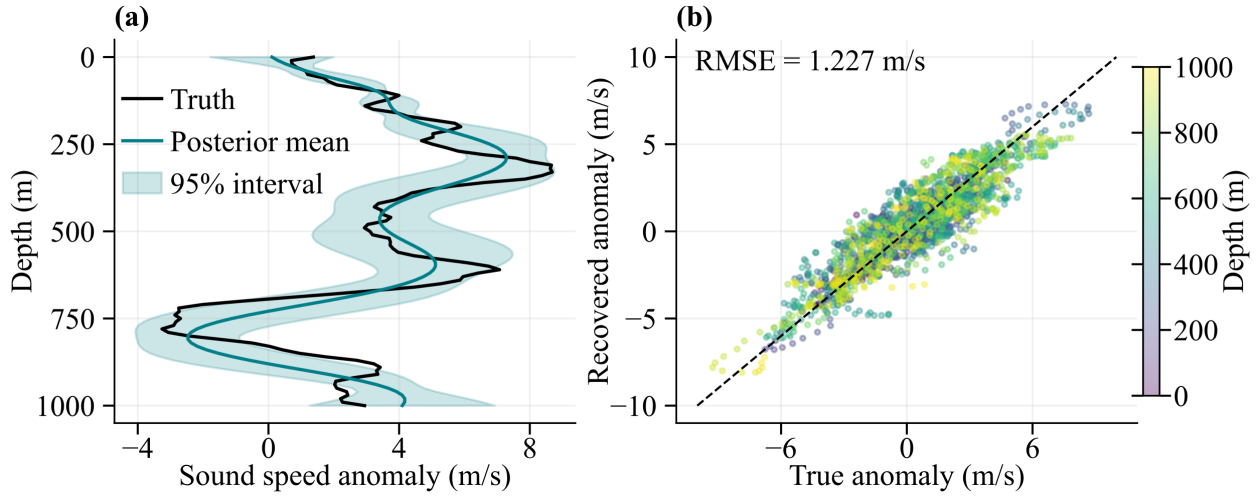


Figure 4. Profile recovery and ensemble agreement. (a) First reference anomaly with the posterior mean and pointwise interval. (b) True and reconstructed anomalies for twenty profiles

Profile specific variation remains relevant when the average error is favorable. A realization may contain a local anomaly in a region that shares nearly identical sensitivity across many paths. Another may have most of its energy in a broadly observed vertical mode. The first can exhibit substantial reconstruction error alongside a small timing residual. The second may be reconstructed accurately even with the same noise level and posterior interval width. In the fixed parameter Gaussian model, this variation arises because actual errors depend on the realization while conditional covariance depends only on the declared geometry and statistics. The ensemble box plots retain this distinction. Reporting the first profile alone would hide it and could encourage a misleading interpretation of the interval as a direct measure of realized local error. Several implementation checks are relevant to interpreting these values. The conditional covariance is symmetrized conceptually by its analytical construction, and tiny negative diagonal values caused by floating point cancellation are clipped only at the numerical precision floor before taking square roots. This numerical safeguard permits real-valued standard deviations while preserving the need to inspect materially negative covariance values. A production implementation should also inspect the spectrum and condition number of the observation covariance when arrival sensitivities become nearly redundant. The current calculations use direct linear solves and a strictly positive noise variance. The calculations use the declared noise variance and direct linear system solutions throughout. Because the depth dimension is modest, the numerical archive permits independent verification of both the covariance update and the reported intervals with a different linear algebra implementation.

4.2. Accuracy calibration and covariance sensitivity

The main quantitative comparison appears in Table 1. The proposed model obtains an aggregate profile error of 1.227 metres per second, compared with 1.450 metres per second for the independent depth reference. Its mean absolute error is 0.936 metres per second, pointwise coverage is 86.68 percent, and average pointwise interval width is 3.547 metres per second. The fitted timing residual is 0.252 milliseconds. These statistics use all eighty profiles and the complete depth grid. Mean absolute error gives a less tail sensitive description than root mean square error, while coverage and interval width describe the conditional uncertainty. The underdetermined observation matrix permits small timing residuals to coexist with substantial depth reconstruction error. In ocean acoustic inversion, a model resolution matrix describes how estimated parameters represent weighted combinations of the underlying true parameters [24]. The table therefore reports both domains and retains the oracle label for the matched covariance reference. Using the displayed rounded RMSE values, the decrease from 1.450 to 1.227 metres per second corresponds to approximately 15.4 percent relative to the independent-depth reference.

Table 1. Ensemble performance with identical observations. Coverage is pointwise; the oracle uses the reference covariance

Model	RMSE (m/s)	MAE (m/s)	Coverage (%)	Width (m/s)	Residual (ms)
Independent	1.450	1.107	99.901	11.301	0.270
Smooth GP	1.790	1.377	10.483	0.438	1.174
Short GP	1.269	0.972	99.926	8.929	0.252
Kernel GP	1.227	0.936	86.683	3.547	0.252
Oracle GP	1.221	0.934	96.052	4.623	0.252

Figure 5(a) compares eighty profile RMSE values for each covariance model. The proposed aggregate error is 1.227 metres per second, compared with 1.450 for the independent depth model, 1.790 for the smooth model, 1.269 for the short scale model, and 1.221 for the oracle. The difference between the proposed and independent depth errors is approximately 15.4 percent. The box plots retain the spread among realizations and expose cases in which a favorable average coexists with a comparatively large individual error. Figure 5(b) shows the proposed point error histogram across 8080 depth profile combinations. Its centre and tails describe complementary parts of the error distribution: the mean absolute error is 0.936 metres per second, while squared error gives larger deviations more influence on the RMSE.

Figure 5(c) compares forty-nine quantiles of standardized errors with a standard normal reference. The deviations in the tails indicate how the observed error distribution differs from the conditional scales supplied by the covariance model. Repeated prior-predictive experiments and suitable test quantities provide a basis for checking computed posterior distributions [25]. Figure 5(d) shows a proposed pointwise coverage of 86.68 percent for nominal 95 percent intervals. The oracle reaches 96.05 percent coverage, while the smooth model reaches 10.48 percent. These contrasts identify a pronounced effect of covariance assumptions on declared precision, even though the proposed and oracle mean errors are close. The independent and short scale models exceed 99 percent coverage with considerably wider intervals, making interval width essential for interpreting their apparent conservatism.

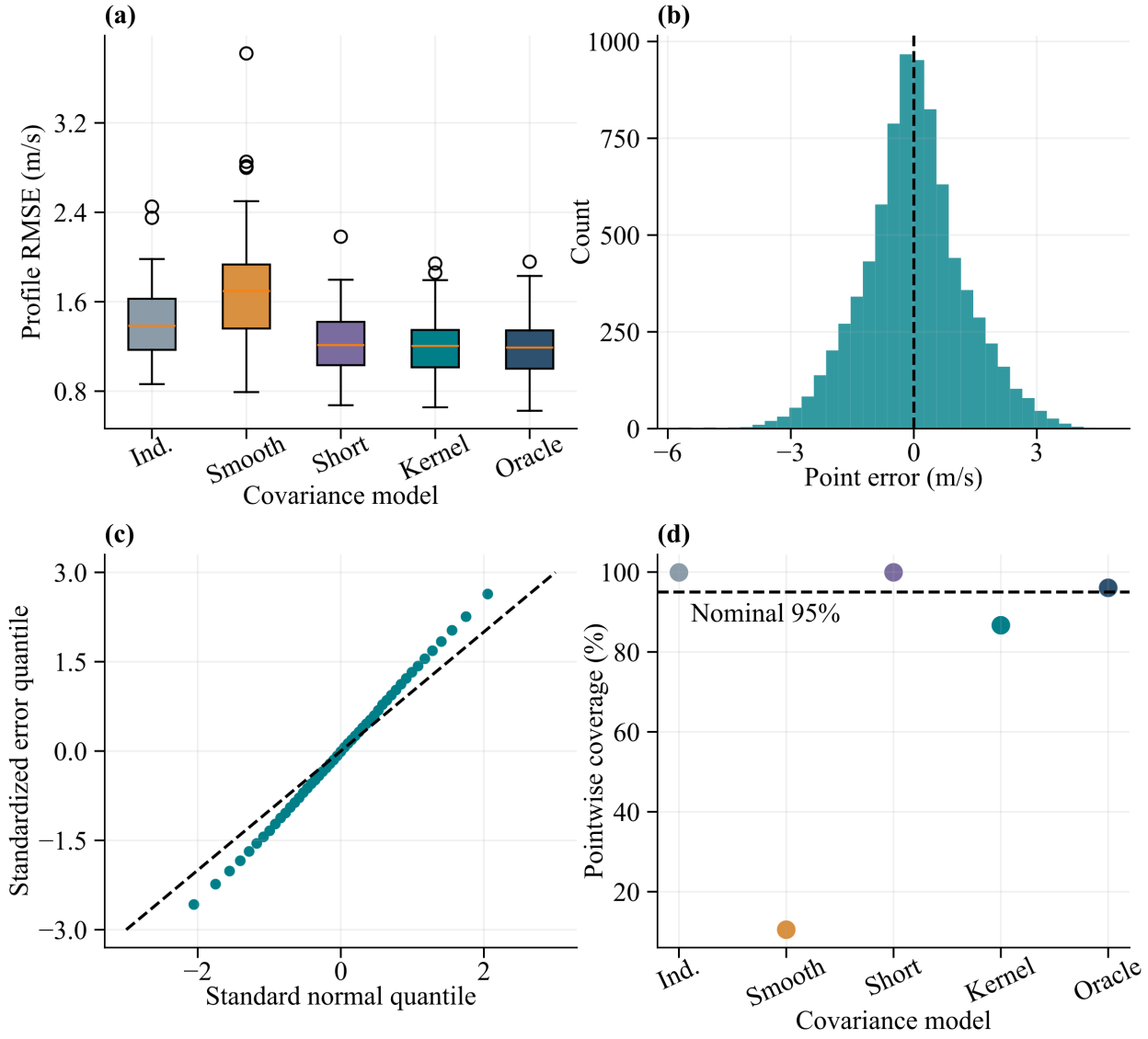


Figure 5. Accuracy and calibration diagnostics. (a) Profile error distributions. (b) Point error histogram. (c) Standardized error quantile comparison. (d) Pointwise interval coverage

Covariance mismatch changes both recovery and calibration. The proposed scale exceeds the reference scale, so it assigns greater correlation to some depth separations than the reference covariance does. That assumption can distribute a local acoustic correction over a broader region and can reduce conditional uncertainty more strongly than warranted for short vertical fluctuations. The direction of coverage error depends on the interaction between the complete covariance and observation operator. The empirical calculation is therefore necessary. The oracle comparison helps

distinguish numerical conditioning from prior misspecification. Differences in calibration between the matched and proposed models indicate the sensitivity of the result to assumptions about environmental variability. If both behave unexpectedly, the observation covariance, ensemble specification, and implementation require inspection before a physical explanation is proposed. The oracle result makes this limitation especially informative. Its reconstruction error is close to that of the proposed model, while its interval coverage is much nearer the declared probability. A modest difference in mean error can thus accompany a material difference in uncertainty assessment. This occurs because covariance mismatch influences the amount of uncertainty removed by the measurements even when the posterior means remain similar. Evaluating only reconstruction error would obscure that distinction. The matched model checks consistency under the specified reference distribution; independent environmental evidence is needed to assess its covariance family in the ocean. Its role is to reveal how much of the proposed model behavior can be explained by the selected prior scale, with independent covariance information required to identify that scale in an operational application. The displayed intervals condition on fixed prior amplitude, correlation length and observation noise; integrating hyperparameter uncertainty would require an additional inference layer.

The reported intervals condition on fixed covariance parameters. They omit uncertainty in the prior amplitude, correlation length, and noise standard deviation. Estimating these parameters from a small collection of arrivals and then substituting point estimates would generally retain that omission. Hierarchical Gaussian process modeling assigns priors to hyperparameters and propagates both latent-function and hyperparameter uncertainty into experimental design [26]. The distinction matters especially where the acoustic data provide little information about vertical correlation. Interval width should correspond to a stated source of uncertainty. The present interval width is interpretable precisely because its scope is narrow. It describes depth uncertainty induced by the assumed prior and observation noise, with propagation and acquisition errors requiring separate uncertainty components. The numerical comparison reveals a specific uncertainty tradeoff that deserves attention beyond the ranking of mean errors. The independent depth model covers almost every reference anomaly value, but its intervals remain broad because path integration leaves most independent cells unresolved. The short correlation model has a similar coverage pattern despite a lower reconstruction error. Their broad intervals produce conservative conditional descriptions under the specified variability model. The broad smooth model behaves differently: it produces very narrow intervals and misses much of the true depth structure. Its low coverage reflects a prior that assigns little variance to the fluctuations present in the evaluation ensemble. The proposed model improves accuracy relative to these alternatives, yet its coverage also falls below the nominal level. Its performance in this benchmark combines reduced mean reconstruction error with an identifiable calibration limitation.

The noise and correlation length analysis varies six timing noise standard deviations from 0.1 to 1.2 milliseconds and six prior scales from 40 to 300 metres. Each row of the risk surface uses a distinct observation error array shared across the six covariance choices at that level. This common noise within each row permits a cleaner comparison of length scale effects. Noise arrays differ between rows, introducing finite ensemble variability into comparisons across noise levels. The anomaly ensemble is held fixed, and the observation noise supplied to the inversion matches the specified error level for that row. Gaussian process validation depends on covariance-parameter estimation, with poorly estimated parameters affecting predictive uncertainty and motivating robust validation criteria [27]. The experiment evaluates sensitivity while retaining the main model parameters fixed before inspection of the surface.

Figure 6(a) presents the thirty-six RMSE values obtained from six noise levels and six prior length scales. The length grid contains 40, 70, 100, 140, 200, and 300 metres; the noise grid contains 0.1, 0.2, 0.3, 0.5, 0.8, and 1.2 milliseconds. Every cell uses the same eighty true profiles, and all six scales in a noise row share a common noise realization. Numerical labels permit direct comparison between adjacent parameter combinations. The surface therefore exposes the response of reconstruction accuracy to prior correlation and observation precision under controlled conditions. The proposed 120 metre scale remains fixed for the main comparison; the grid serves as a descriptive sensitivity experiment. An excessively long scale suppresses short vertical features even when the observation noise is small. Prior representations that assign insufficient variance to the reference environmental structures retain a source of bias as observation noise decreases. Conversely, a short scale can allocate substantial variance to poorly observed local variations. It may fit the delays while retaining a larger uncertainty across depth. The central part of the scale grid tests the tradeoff between these effects. Its location relative to the reference ninety metre scale influences performance together with the acoustic observation geometry. Because the acoustic operator measures weighted depth combinations, the scale that minimizes finite ensemble error depends on its interaction with the observation geometry. The risk surface resolves the interaction between prior scale and observation geometry for the specified evaluation. The sensitivity grid is descriptive: reference-profile errors are used to evaluate parameter effects, while the main 120 metre prior remains fixed throughout the principal comparison.

Figure 6(b) displays profile error densities for the five fixed covariance models, with medians marked within each violin. Each density represents eighty profile specific errors from the 0.3 millisecond noise condition. The independent depth and short scale models retain broad uncertainty despite their different error distributions, whereas the smooth covariance imposes a restrictive vertical structure that increases reconstruction error. The proposed and oracle distributions are close in aggregate accuracy, while their interval coverage differs by approximately 9.37 percentage

points. This comparison shows why the error distribution and interval diagnostics answer distinct engineering questions. The density width visualizes the finite ensemble sample concentration and depends on the smoothing used to draw the violin, while the median retains a direct sample-based reference. Each violin summarizes 80 profile RMSE values; its width represents a smoothed sample density rather than an additional uncertainty interval or additional independent observations.

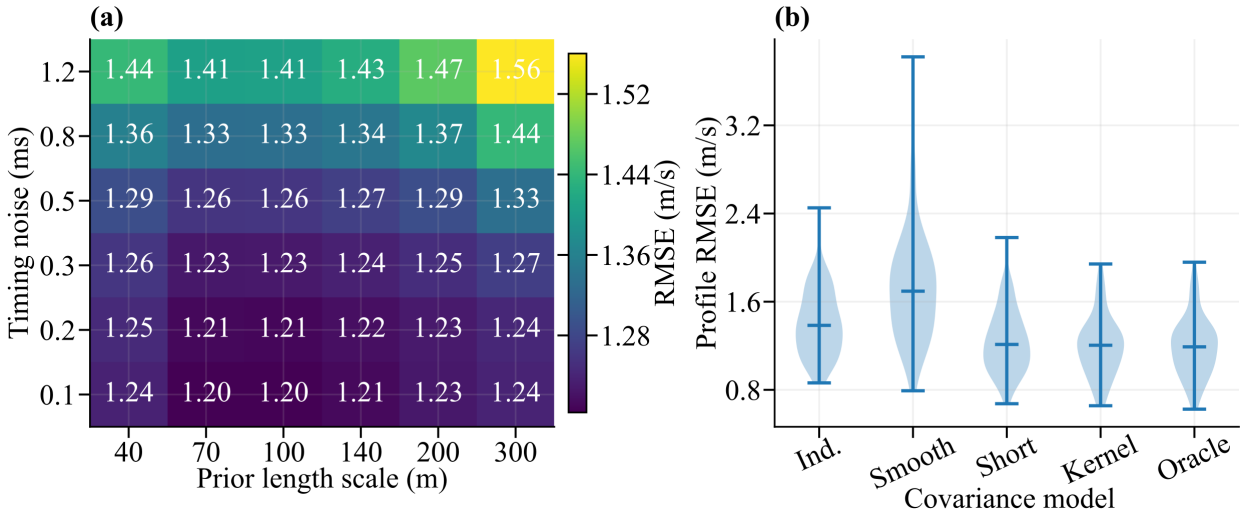


Figure 6. Sensitivity to statistical assumptions. (a) Reconstruction error across six noise levels and six prior correlation lengths. (b) Profile error densities for the five fixed covariance models

Noise affects the relative weight given to the acoustic data. As the declared noise increases, the posterior gives less emphasis to fitting individual delays and retains more of the prior covariance. The response varies with the information content of each observation combination. The effect depends on the signal to noise ratio of each observation combination. Strongly observed combinations can remain informative while weak combinations become effectively prior dominated. The associated vertical resolution varies through the water column. Interpreting the surface through observation combinations accounts for the depth dependence of resolving power. A future acquisition analysis could connect these combinations to singular vectors or averaging kernels, provided it also retained the observation error structure. A low error region on the sensitivity grid reflects compatibility among the specified observation support, anomaly family, and covariance scale. Selecting an operational scale requires independent environmental evidence or a validation procedure that separates parameter selection from evaluation. The covariance length regulates how updates propagate between depths, while the noise variance regulates the weight given to each observation combination. These effects can compensate in the observation domain and still produce different depth reconstructions. A scale chosen exclusively from fitted delay residuals may therefore favour an environmental description with excessive confidence in unresolved vertical components. The recorded risk matrix preserves the full parameter comparison for subsequent evaluation with alternative geometries.

4.3. Path availability residual structure and depth dependent error

Path availability is examined by retaining the first 4, 8, 12, 16, 20, and 24 rows of the ordered observation matrix. The experiment uses a deterministic nested sequence. Because rows are ordered by receiver depth and image branch, increasing the count changes both the number of observations and the range of depth footprints. The sequence combines the effects of measurement count and receiver geometry. Bayesian experimental design can select observation configurations by minimizing the trace of posterior covariance under a linearized measurement model [28]. The sequence describes information accumulation for the specified array; evaluating receiver placement requires comparisons between alternative subsets at equal observation counts. The declared row order fixes every subset before reconstruction. The path-count sequence is conditional on the fixed row order; equal-count comparisons between alternative subsets would be required to establish a geometry-selection advantage.

Figure 7(a) reports RMSE values of 2.339, 2.231, 2.088, 1.829, 1.484, and 1.227 metres per second for 4, 8, 12, 16, 20, and 24 retained paths, respectively. The complete sequence reduces error by approximately 47.5 percent relative to the four-path case. Because the rows are ordered by receiver depth and image branch, the reduction reflects both additional observations and broader depth support. Figure 7(b) pairs these errors with mean interval widths of 8.389, 7.689, 6.612, 5.396, 4.267, and 3.547 metres per second. The paired trajectory connects realized accuracy with conditional precision and reveals how additional path information narrows the posterior under the fixed prior. Nested conditioning reduces posterior covariance in the positive semidefinite sense when the prior, noise model, and added observations are consistent. Realized reconstruction error under covariance mismatch can fluctuate across successive

steps in a finite ensemble. Additional noisy observations can move the posterior mean closer to or farther from an individual true profile. The ensemble average is less variable than a single realization but remains a finite sample quantity. This distinction explains why a path count curve should be inspected alongside interval width and the exact subset definition. A monotonic uncertainty reduction is a statement within the assumed model; an observed error reduction is an empirical property of the evaluated ensemble. Each quantity consequently requires its own empirical interpretation.

Figure 7(c) displays the full path fitted timing residual distribution, with reference lines at plus and minus 0.3 milliseconds. Its RMSE is 0.252 milliseconds, slightly below the injected noise standard deviation. The difference follows from conditioning the profile on those same delay observations: the update explains part of the observed variation, including components contributed by noise. Posterior predictive checks compare data summaries with a predictive reference distribution, and calibrated posterior p-values support Bayesian model assessment [29]. Independent noisy repeats or withheld paths would provide a complementary assessment of prediction. The current residual histogram characterizes agreement with the fitted observations and helps identify whether a covariance imposes an acoustic mismatch, as illustrated by the smooth model residual of 1.174 milliseconds. The 0.252 millisecond residual describes observations used for conditioning; predictive assessment would require withheld paths or independent noisy repeats.

Figure 7(d) displays the 101 by 101 posterior covariance matrix as filled contours. Diagonal values describe the remaining pointwise variance, while off diagonal values describe conditional dependence between depth errors. Coherent contour regions arise because path integration constrains combinations of depths and the prior transfers information across vertical separation. Oppositely signed covariance regions indicate compensating uncertainty: increasing one part of the profile can be offset by changes elsewhere while maintaining similar delays. The contour scale retains covariance units, allowing the apparent spatial extent of dependence to be distinguished from its magnitude. The associated uncertainty structure is fixed across all eighty profiles because geometry, covariance parameters, and observation noise are held constant; realized errors vary with the anomaly and delay vector. A broad covariance region can therefore persist even when the fitted delay residual is small, because compensating depth perturbations remain compatible with the observations. Conditioning constrains the path weighted combinations most strongly, leaving uncertainty in combinations that alter local sound speed while preserving those integrated quantities. Inspecting both covariance magnitude and spatial extent helps identify this distinction. A subsequent comparison could hold the number of receivers fixed and exchange one arrival branch for another, testing whether the resulting change reduces uncertainty in a specific vertical interval. Such a comparison would directly connect a covariance feature to an acquisition choice under the same statistical prior.

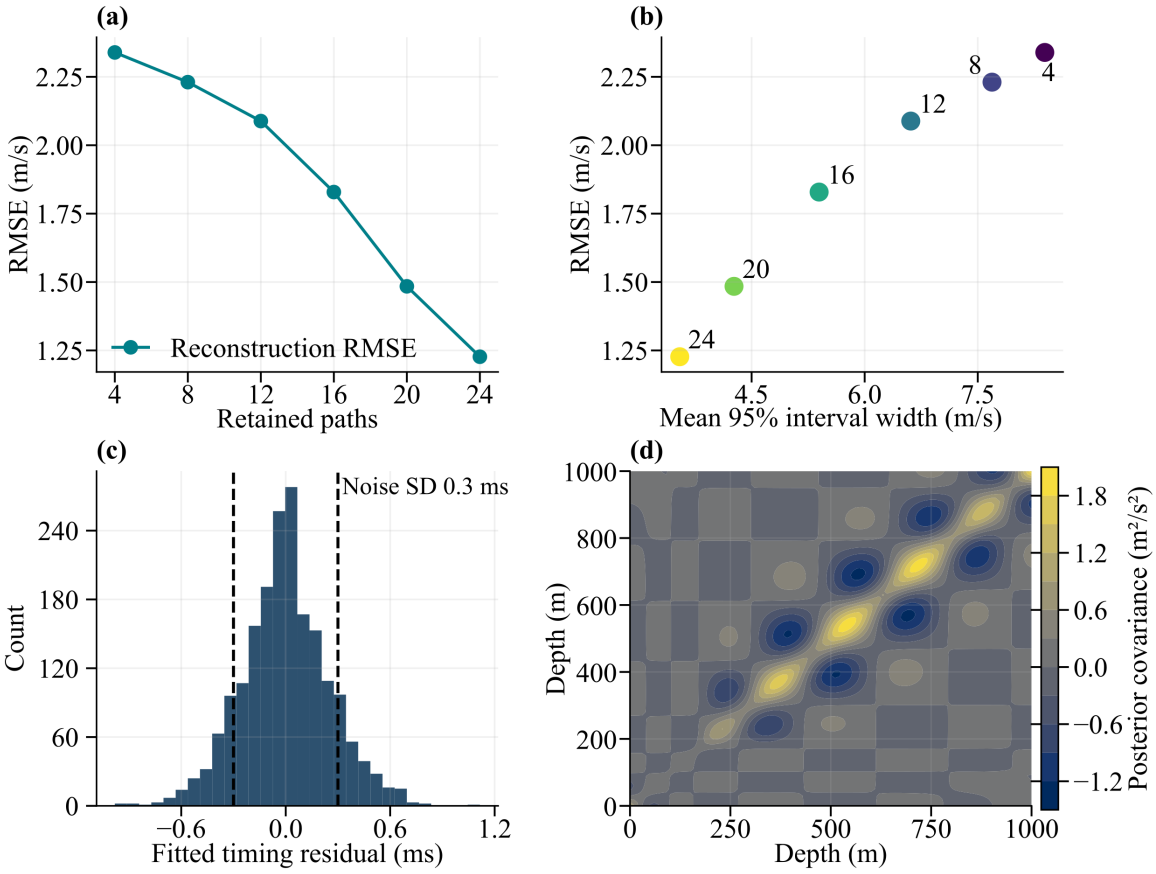


Figure 7. Path availability and residual structure. (a) Error versus retained paths. (b) Interval width versus error. (c) Fitted timing residual distribution. (d) Posterior depth covariance contours

Analysis of geophysical posterior ensembles uses distribution-aware summaries to represent central tendencies and non-uniqueness in nonlinear inverse solutions [30]. Figure 8(a) presents five empirical cumulative RMSE distributions with eighty realizations per curve and a reference probability of 0.9. At a fixed error threshold, a larger cumulative probability indicates that more profiles have been reconstructed within that threshold. The proposed aggregate RMSE is 1.227 metres per second, close to the oracle value of 1.221 metres per second. The cumulative curves retain differences in profile specific performance that may be obscured by these nearby averages. Their upper portions identify cases with relatively large errors and support inspection of whether a covariance improves most profiles or primarily changes the tail. Every curve uses the same anomaly ensemble and observations, preserving the paired comparison across methods.

Figure 8(b) shows absolute errors in five depth bands spanning 0–190, 200–390, 400–590, 600–790, and 800–990 metres. Each band contains twenty depth nodes and eighty profiles, yielding 1600 correlated point errors per box plot. The 1000 metre boundary node remains part of the aggregate metrics and falls outside these five equal sized bands. Absolute errors expose local magnitudes, while the box centres and whiskers identify how their distribution varies through the water column. A receiver near a band can collect arrivals that also integrate anomalies in other intervals, so its proximity alone gives an incomplete account of local recovery. The depth grouping tests this consequence directly using the proposed posterior mean and the known controlled truth. The five bands each contain 20 depth nodes; the 1000 metre node is excluded from this equal-count grouping and retained in the whole-column metrics.

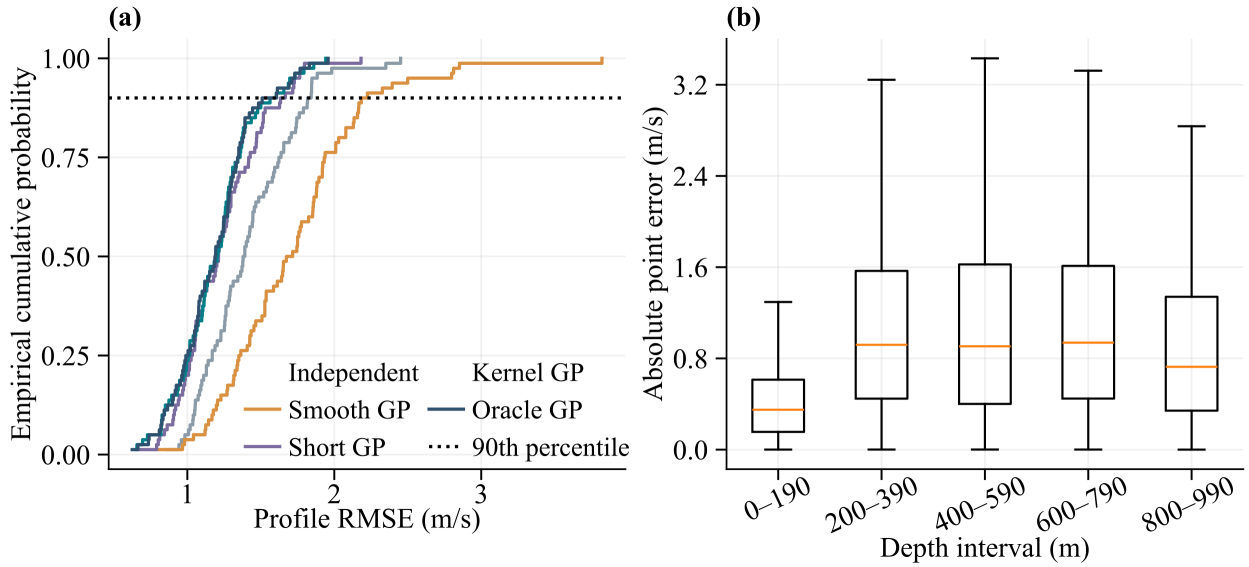


Figure 8. Distributional and depth dependent errors. (a) Empirical cumulative profile error distributions. (b) Absolute errors in five depth intervals

Depth dependent error can arise from both geometry and prior coupling. Local error depends on the complete reflected path footprint, which integrates anomalies across several depth intervals. The posterior covariance can connect neighboring intervals strongly enough to produce broad coherent error structures. Banded absolute errors reveal local magnitudes but discard their signs and cross depth dependence. The covariance contours supply the complementary dependence information. Feature specific recovery can be evaluated through targeted perturbation tests and averaging kernels defined for a thermocline displacement or a localized anomaly. The present outputs remain profile estimates with conditional uncertainty, evaluated against the known realization at each grid node. The observation model also makes horizontal representativeness a material limitation. A range dependent ocean can produce delays that a one-dimensional profile model explains through an effective vertical anomaly. That anomaly can support prediction of similar integrated observations while differing from pointwise hydrographic profiles along the track. The evaluation isolates vertical recovery within a horizontally uniform environment. Comparison with field casts would additionally require a definition of spatial representativeness along the acoustic track. A field evaluation would need to define whether the target is a local profile, a path averaged vertical structure, or a parameterization of a range dependent field. The statistical covariance and observation operator would then need to match that target. This requirement affects the meaning of accuracy before any numerical performance comparison is made. For an operational extension, the target must be specified as a local profile, a path-averaged vertical representation or a range-dependent field before hydrographic validation is interpreted.

The banded error samples share two dependence structures: neighbouring depths are correlated within a profile, and each posterior estimate combines the same set of path sensitivities. Resampling independent profiles would preserve both structures when quantifying uncertainty in a band median or an upper error quantile. Treating the 1600 entries in a band as independent replicates would artificially increase the effective sample size. The independent depth covariance imposes a specific assumption about local variability. It encodes a specific lack of correlation at a grid spacing of ten

metres, assigning variability to structures that an integrated travel time may be unable to distinguish. Refining the grid while holding that covariance unchanged would alter the implied random field and could change the comparison. The correlated models have a clearer physical scale because their covariance is defined by actual depth separation. Even so, a grid convergence study would be needed before treating the discretized result as an approximation to a continuous inverse solution with known numerical accuracy. The present experiment uses one grid to preserve comparability among methods. Its reference values apply to the declared grid and quadrature. The stored archive retains the profile dimension, enabling paired profile resampling across covariance models and depth bands. For each resampled ensemble, a band specific median error could be calculated together with the difference between competing estimators. Retaining the same resampled profiles for every method would preserve the paired design and avoid confounding estimator differences with changes in anomaly composition. The resulting distribution would quantify uncertainty in the comparative statistic while maintaining the actual within profile depth dependence. A similar calculation could examine an upper quantile when the engineering priority is limiting large errors in a sensitive interval. Its interpretation would remain conditional on the specified anomaly family and fixed path geometry. It also preserves the observation matrix needed to investigate whether a difficult band corresponds to weak sensitivity or to a covariance mismatch. In a targeted follow-up experiment, an anomaly centred within that band could be varied in width and amplitude while holding the acquisition geometry fixed, separating the effect of local environmental scale from the distribution of path support. Any uncertainty interval for a comparative performance statistic should resample whole profiles and preserve the paired observations across methods.

5. Conclusion

A Gaussian process formulation for ocean sound-speed-profile inversion was developed in this paper, and ray-travel-time kernels were used as linear observation operators. A depth covariance model is added to address acoustic delay and return a reconstructed anomaly along with conditional uncertainty. The assessment examined profile recovery, timing agreement, calibration and the influence of observation support. Therefore, the above characteristics need to be considered independently in the analysis of weakly constrained regions of the environment. The formulation offers a transparent framework for investigating the joint effects of acoustic information and prior correlation, along with the physical operator and statistical assumptions. The engineering outcome is a reproducible means of examining how path support and prior depth correlation jointly determine profile estimates and conditional uncertainty.

The present study is limited by idealized propagation geometry, fixed reference paths, horizontal uniformity and Gaussian observation errors. The conditional intervals account for the prescribed prior and observation noise; covariance parameters, arrival association, acquisition geometry and propagation approximation require additional uncertainty components. The comparison evaluates selected covariance alternatives within a common inference family. These restrictions limit the interpretation of the reported performance and prevent the direct application of the reported results to operational environmental monitoring. Establishing independent resolution of the recovered vertical structure requires sensitivity and uncertainty assessment alongside profile agreement and timing residuals.

Future work should evaluate refracted-ray and finite-frequency propagation with independent observations before operational performance is inferred. Independent acoustic and hydrographic measurements are required to assess calibration under actual arrival tracking and environmental fluctuations. Future work should examine hierarchical covariance inference, structured acquisition error models and geometry selection using validation protocols that separate parameter tuning from assessment. Expanding the unknown field beyond a horizontally uniform profile can enable the model target to have the spatial averaging properties of acoustic observations and maintain an interpretable connection between physical sensitivity and probabilistic reconstruction. A staged assessment would first replace the fixed operator with independently computed arrivals, then test held-out acoustic observations and finally compare with independently measured hydrographic profiles.

6. Declarations

6.1 Author Contributions

Tahani Al-Ghafli contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Hessa Al-Mansoori contributes to validation, analysis, investigation, data collection, draft preparation. Karima Al-Riyami contributes to conceptualization, validation, analysis. All authors have read and agreed with the manuscript before its submission and publication.

6.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3 Institutional Review Board Statement

Not applicable.

6.4 Informed Consent Statement

Not applicable.

6.5 Date Availability

Not applicable.

6.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Xu, W., Chen, Y., Jiang, Y., Li, Z., Wang, X., Li, B., Li, N., & Zhang, W. (2026). A framework for inversion and assimilation of deep-sea mirror ocean acoustic tomography observations. International Journal of Applied Earth Observation and Geoinformation, 154, Article 105626. https://doi.org/10.1016/j.jag.2026.105626

[2] Ollivier, E., Touret, R. X., McKinley, M., Jin, J., Bracco, A., & Sabra, K. G. (2024). Performance study of ray-based ocean acoustic tomography methods for estimating submesoscale variability in the upper ocean. The Journal of the Acoustical Society of America, 155(2), 1315–1335. https://doi.org/10.1121/10.0024819

[3] Peng, S., Callies, J., Wu, W., & Zhan, Z. (2024). Seismic ocean thermometry of the Kuroshio Extension region. Journal of Geophysical Research: Oceans, 129(2), Article e2023JC020636. https://doi.org/10.1029/2023JC020636

[4] Fan, H., Xie, S., & Xue, S. (2026). Inversion of sound speed profile controlled by sparse observations: Research background, current status and technical analysis. Oceans, 7(3), Article 45. https://doi.org/10.3390/oceans7030045

[5] Bortolotti, V., Landi, G., & Zama, F. (2025). Uniform multipenalty regularization for linear ill-posed inverse problems. SIAM Journal on Scientific Computing, 47(2), A790–A810. https://doi.org/10.1137/23M1600943

[6] Horst, A., Maboudi Afkham, B., Dong, Y., & Lemvig, J. (2025). Uncertainty quantification for linear inverse problems with Besov prior: A randomize-then-optimize method. Statistics and Computing, 35(4), Article 101. https://doi.org/10.1007/s11222-025-10638-2

[7] Paine, T., & Bhatt, E. (2024). Geometric stochastic ray propagation using the special Euclidean group. JASA Express Letters, 4(4), Article 046002. https://doi.org/10.1121/10.0025522

[8] Azizzadenesheli, K., Kovachki, N., Li, Z., Liu-Schiaffini, M., Kossaifi, J., & Anandkumar, A. (2024). Neural operators for accelerating scientific simulations and design. Nature Reviews Physics, 6(5), 320–328. https://doi.org/10.1038/s42254-024-00712-5

[9] Bai, T., Teckentrup, A. L., & Zygalakis, K. C. (2024). Gaussian processes for Bayesian inverse problems associated with linear partial differential equations. Statistics and Computing, 34(4), Article 139. https://doi.org/10.1007/s11222-024-10452-2

[10] Matsumoto, T., & Sullivan, T. J. (2024). Images of Gaussian and other stochastic processes under closed, densely-defined, unbounded linear operators. Analysis and Applications, 22(03), 619–633. https://doi.org/10.1142/S0219530524400025

[11] Lahtinen, J. (2024). On bias and its reduction via standardization in discretized electromagnetic source localization problems. Inverse Problems, 40(9), 095002. https://doi.org/10.1088/1361-6420/ad5f53

[12] Liu, Y., Hu, R., Dou, D., Niu, H., Chen, D., & Xu, L. (2024). Gaussian processes with normal-mode-based kernels for matched field processing. Applied Acoustics, 220, Article 109954. https://doi.org/10.1016/j.apacoust.2024.109954

[13] Peng, S., Callies, J., Nagura, M., & McPhaden, M. J. (2025). Yanai waves in the deep East Indian Ocean observed with seismic ocean tomography. Journal of Physical Oceanography, 55(4), 471–488. https://doi.org/10.1175/JPO-D-24-0107.1

[14] Roux, P., Iturbe, I., Nicolas, B., Virieux, J., & Mars, J. I. (2011). Travel-time tomography in shallow water: Experimental

demonstration at an ultrasonic scale. The Journal of the Acoustical Society of America, 130(3), 1232–1241. https://doi.org/10.1121/1.3621271

[15] Zheng, Y., Yu, S., Qin, Z., Liu, X., Xie, C., Liu, M., & Zhao, J. (2024). Bottom multi-parameter Bayesian inversion based on an acoustic backscattering model. Journal of Marine Science and Engineering, 12(4), Article 629. https://doi.org/10.3390/jmse12040629

[16] Travelletti, C., & Ginsbourger, D. (2024). Disintegration of Gaussian measures for sequential assimilation of linear operator data. Electronic Journal of Statistics, 18(2), 3825–3857. https://doi.org/10.1214/24-EJS2262

[17] Chlingaryan, A., Leung, R., & Melkumyan, A. (2024). Augmenting stationary covariance functions with a smoothness hyperparameter and improving Gaussian process regression using a structural similarity index. Mathematical Geosciences, 56(3), 605–637. https://doi.org/10.1007/s11004-023-10095-5

[18] Saha, P., Touret, R. X., Ollivier, E., Jin, J., McKinley, M., Romberg, J., & Sabra, K. G. (2025). Leveraging sound speed dynamics and generative deep learning for ray-based ocean acoustic tomography. JASA Express Letters, 5(4), Article 040801. https://doi.org/10.1121/10.0036312

[19] Narnhofer, D., Habring, A., Holler, M., & Pock, T. (2024). Posterior-variance–based error quantification for inverse problems in imaging. SIAM Journal on Imaging Sciences, 17(1), 301–333. https://doi.org/10.1137/23M1546129

[20] Yang, W., Xue, S., Ma, Y., Xiao, Z., Li, B., & Wang, K. (2026). An ultra-efficient ray-tracing method based on a precomputed full-element ray grid for underwater acoustic positioning. Measurement, 258, Article 119546. https://doi.org/10.1016/j.measurement.2025.119546

[21] Smaragdakis, C., Taroudaki, V., & Taroudakis, M. I. (2024). Using machine learning techniques in inverse problems of acoustical oceanography. Studies in Applied Mathematics, 153(2), Article e12704. https://doi.org/10.1111/sapm.12704

[22] Hauptmann, A., Mukherjee, S., Schönlieb, C.-B., & Sherry, F. (2025). Convergent regularization in inverse problems and linear plug-and-play denoisers. Foundations of Computational Mathematics, 25(4), 1087–1120. https://doi.org/10.1007/s10208-024-09654-x

[23] Han, X., Liu, C., Li, S., Gao, D., Li, G., Zhu, M., & Xu, P. (2026). A segmented Bayesian method for parameter calibration in shallow-water acoustic tomography. Journal of Marine Science and Engineering, 14(19), Article 1850. https://doi.org/10.3390/jmse14191850

[24] Bonnel, J., Dosso, S. E., Hodgkiss, W. S., Ballard, M. S., Garcia, D. D., Lee, K. M., McNeese, A. R., & Wilson, P. S. (2024). Trans-dimensional inversion for seafloor properties for three mud depocenters on the New England shelf under dynamical oceanographic conditions. The Journal of the Acoustical Society of America, 155(3), 1825–1839. https://doi.org/10.1121/10.0025176

[25] Modrák, M., Moon, A. H., Kim, S., Bürkner, P., Huurre, N., Faltejsková, K., Gelman, A., & Vehtari, A. (2025). Simulation-based calibration checking for Bayesian computation: The choice of test quantities shapes sensitivity. Bayesian Analysis, 20(2), 461–488. https://doi.org/10.1214/23-BA1404

[26] Jones, K. D., & Dowling, A. W. (2026). BITS for GAPS: Bayesian information-theoretic sampling for hierarchical Gaussian process surrogates. Computers & Chemical Engineering, 211, Article 109650. https://doi.org/10.1016/j.compchemeng.2026.109650

[27] Marrel, A., & Iooss, B. (2024). Probabilistic surrogate modeling by Gaussian process: A review on recent insights in estimation and validation. Reliability Engineering & System Safety, 247, Article 110094. https://doi.org/10.1016/j.ress.2024.110094

[28] Eberle-Blick, S., & Hyvönen, N. (2024). Bayesian experimental design for linear elasticity. Inverse Problems and Imaging, 18(6), 1294–1319. https://doi.org/10.3934/ipi.2024015

[29] Paganin, S., & de Valpine, P. (2025). Computational methods for fast Bayesian model assessment via calibrated posterior p-values. Journal of Computational and Graphical Statistics, 34(2), 462–473. https://doi.org/10.1080/10618600.2024.2374585

[30] Mills, C. L., Rudolph, M. L., & Lekić, V. (2025). Exploring and summarizing ensemble solutions to geophysical inverse problems. Journal of Geophysical Research: Machine Learning and Computation, 2(4), Article e2024JH000472. https://doi.org/10.1029/2024JH000472