Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Correlated timing errors and terrain dependent uncertainty in synthetic slope microseismic localisation

Lai Wei

wl20061002@qq.com

China University of Geosciences

Bin Yan

China University of Geosciences

Xiangli Zhang

China University of Geosciences

Yu Chang

China University of Geosciences

Zehao Li

China University of Geosciences

Research Article

Keywords: Microseismic monitoring, Surface arrays, Correlated timing errors, Hypocentre localisation, Terrain, Uncertainty calibration

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11335856/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Correlated timing errors and terrain dependent uncertainty in synthetic slope microseismic localisation

Lai Wei, Bin Yan, Xiangli Zhang, Yu Chang and Zehao Li

Department of Electronic Information Engineering, School of Mechanical Engineering and Electronic Information, China University of Geosciences (Wuhan), Wuhan, China

Corresponding author: Lai Wei, wl20061002@qq.com

Author emails: wl20061002@qq.com; 150038048@qq.com; 505066723@qq.com; 1138328141@qq.com; 3014106782@qq.com

Article type: Research Article

Abstract

Surface microseismic arrays can support monitoring of subsurface fracture activity, but increasing receiver density does not necessarily improve the reliability of reported locations. We present a reproducible synthetic benchmark that evaluates terrain coordinates, velocity calibration and timing correlation together. Six surface geometries share a 34.5 m horizontal envelope, with 16–576 receivers assigned to four timing groups. An independently implemented two-layer ray solver generates arrivals for 200 paired test events; a separate set of 12 known sources calibrates layer velocities. Generalised least squares is compared with diagonal weighting, homogeneous propagation and flattened receiver coordinates. At 64 receivers and a 0.15 ms array-shared clock standard deviation, covariance-aware inversion gives 90th-percentile location errors of 0.78–0.82 m and empirical 95% ellipsoid coverage of 96.5–97.5%. Diagonal weighting gives errors of 2.06–2.12 m and coverage of 33.0–34.5%. With 576 receivers, diagonal coverage falls to 3.5–4.0%, whereas propagating calibration uncertainty improves covariance-aware coverage to 94.0–96.0%. Additional waveform tests use unknown event times, noise-only detection calibration and automatic arrival picking. These tests expose sensitivity to coherent transients and reference contamination. A weak lateral slowness anomaly also produces severe undercoverage despite modest location errors. The results identify conditional accuracy and uncertainty-calibration requirements for slope-array design, without establishing field performance or an early-warning threshold.

Article Highlights

- More receivers can produce false precision when array timing errors are shared.
- Terrain coordinates and velocity uncertainty must accompany location accuracy tests.
- Noise-only and interference tests expose limits hidden by ideal arrival-time simulations.

Keywords: Microseismic monitoring; Surface arrays; Correlated timing errors; Hypocentre localisation; Terrain; Uncertainty calibration

1 Introduction

Microseismic observations provide information about fracture, frictional slip and other internal processes that can be difficult to observe at the surface. In slope monitoring, they complement displacement measurements by recording elastic disturbances generated within the ground. Associations between microseismic activity and stages of slope failure have been studied in soft soils and an active rock glacier (Yfantis et al. 2021; Faillettaz et al. 2019). These associations motivate monitoring, but an observed event is not by itself evidence of impending failure. Both source localisation and discrimination of unrelated activity are needed before changes in a seismic catalogue can inform a geotechnical interpretation.

Passive seismic monitoring also differs from optical or electromagnetic observation. A ground-coupled sensor measures ground motion produced by elastic-wave propagation. Tree crowns do not obstruct this propagation as they obstruct an optical line of sight. Nevertheless, wind, roots, moving vegetation, vehicles and machinery can generate ground motion that enters the sensors. Vegetation therefore remains a potential noise source. Studies of ambient-noise correlation and combined passive seismic and inclinometer measurements illustrate that slow slope evolution can require observables beyond impulsive microseismic events (Le Breton et al. 2021; Fiolleau et al. 2023). A design intended to detect microfracture activity should not automatically be presented as a direct measurement of creep displacement.

Array processing improves sensitivity by exploiting the spatial coherence of arrivals. Delay-and-sum beamforming, travel-time stacking and matched-field processing are established tools in array seismology (Rost and Thomas 2002; Mao and Gubbins 1995). Dense patch arrays have been used for sensitive detection and localisation of subsurface sources (Chmiel et al. 2019), while stacking velocity analysis and cross-correlation imaging provide other localisation routes (Zhang et al. 2021; Yang et al. 2023). Conventional aperture and sidelobe calculations remain useful for describing directional selectivity, but they do not specify the accuracy of an unknown three-dimensional source inferred from noisy waveforms. Detection, picking and inversion introduce distinct error mechanisms.

Receiver geometry and the propagation model affect those errors. Sensor-network optimisation has been investigated for underground and mine monitoring (Wang et al. 2021, 2022; Hou et al. 2024; Huang et al. 2026). Coordinate errors can alter estimated microseismic locations (Li et al. 2022), and velocity information is particularly relevant on structurally complex slopes (Zhang et al. 2020). Joint or sequential velocity updating and event localisation, and anisotropic propagation models, address further sources of model error (Tian et al. 2022; Shen et al. 2023). Arrival-time uncertainty and probabilistic location studies also show why a single error distance is an incomplete performance description (Abakumov et al. 2020; Liang 2024). The inverse-problem framework itself is well established (Tarantola and Valette 1982; Lienert et al. 1986; Lomax et al. 2000).

A further distinction is whether errors are independent at the receiver level. Receivers connected to one acquisition unit can share a timing perturbation. Treating that perturbation as many independent observations exaggerates the information supplied by additional channels. This is a statistical issue even when the individual marginal error variance is correct. It is especially relevant when dense arrays are assessed with a small number of acquisition groups, when arrival times from both P and S phases share a clock, or when velocity uncertainty is omitted from a location ellipsoid. A metre-scale point estimate can coexist with an unreliable confidence region.

The question addressed here is when increasing receiver density improves both hypocentral accuracy and the reliability of a reported confidence region. We test this in a fully synthetic slope-array benchmark with six terrain surfaces, equal outer aperture, four receiver counts and four timing groups. Covariance-aware and diagonal inversions receive the same arrivals and marginal variances, isolating the effect of shared timing

errors. Separate velocity calibration and repeated calibration datasets test whether improved precision survives uncertainty in the fitted propagation parameters. Coordinate flattening, propagation misspecification and waveform-derived arrivals then identify conditions under which that inference fails. The contribution is a reproducible error-budget comparison with explicit failure boundaries, using established estimators rather than proposing a new beamforming algorithm or an optimal field layout.

2 Methods

2.1 Surface geometries and test sources

Coordinates are expressed in metres, with positive Z upward. The receiver domain spans −17.25 to 17.25 m along both horizontal axes, giving a 34.5 m envelope. Six terrain height functions represent a horizontal plane, a gentle plane, a saddle, a steep plane, a convex surface and a stepped surface. Table 1 gives their explicit definitions. They are controlled mathematical surfaces, not digital elevation models of particular mines. The horizontal receiver positions and source identities are matched across surfaces so that differences can be attributed to the specified coordinate model rather than an uncontrolled change of site.

Table 1 Terrain definitions and principal simulation parameters

Item	Definition or value
Flat	$Z = 4$
Gentle plane	$Z = 4 + X/6 + Y/3$
Saddle	$Z = 4 + X/6 + Y/3 + XY/162$
Steep plane	$Z = 4 + 0.5X + 0.3Y$
Convex	$Z = 4 + 0.22X + 0.16Y + 0.006(X^2 + Y^2)$
Step	$Z = 4 + 0.15X + 0.12Y + 4I(X \geq 0)$
Receiver counts	16, 64, 144, 576; four timing groups
Horizontal envelope	34.5 m × 34.5 m
Independent P/S timing SD	0.12/0.18 ms
Shared timing SD	0, 0.03, 0.15, 0.50 ms
Known calibration sources	12 per terrain, separate from tests
Timing / waveform identities	200 / 100 paired identities

The dense layout contains 24 × 24 receivers at a projected 1.5 m spacing. Splitting the domain into four quadrants gives four groups of 12 × 12 receivers. The 16-, 64- and 144-receiver layouts preserve the same outer envelope, using respectively two, four and six equally spaced horizontal positions on each side of each axis. The four quadrants are timing groups; they are not physically isolated compact patches in the sparse layouts. Figure 1 shows the dense reference layout on each terrain. Sparse layouts are used to compare timing information at fixed outer aperture, rather than to certify grating-lobe-free beamforming.

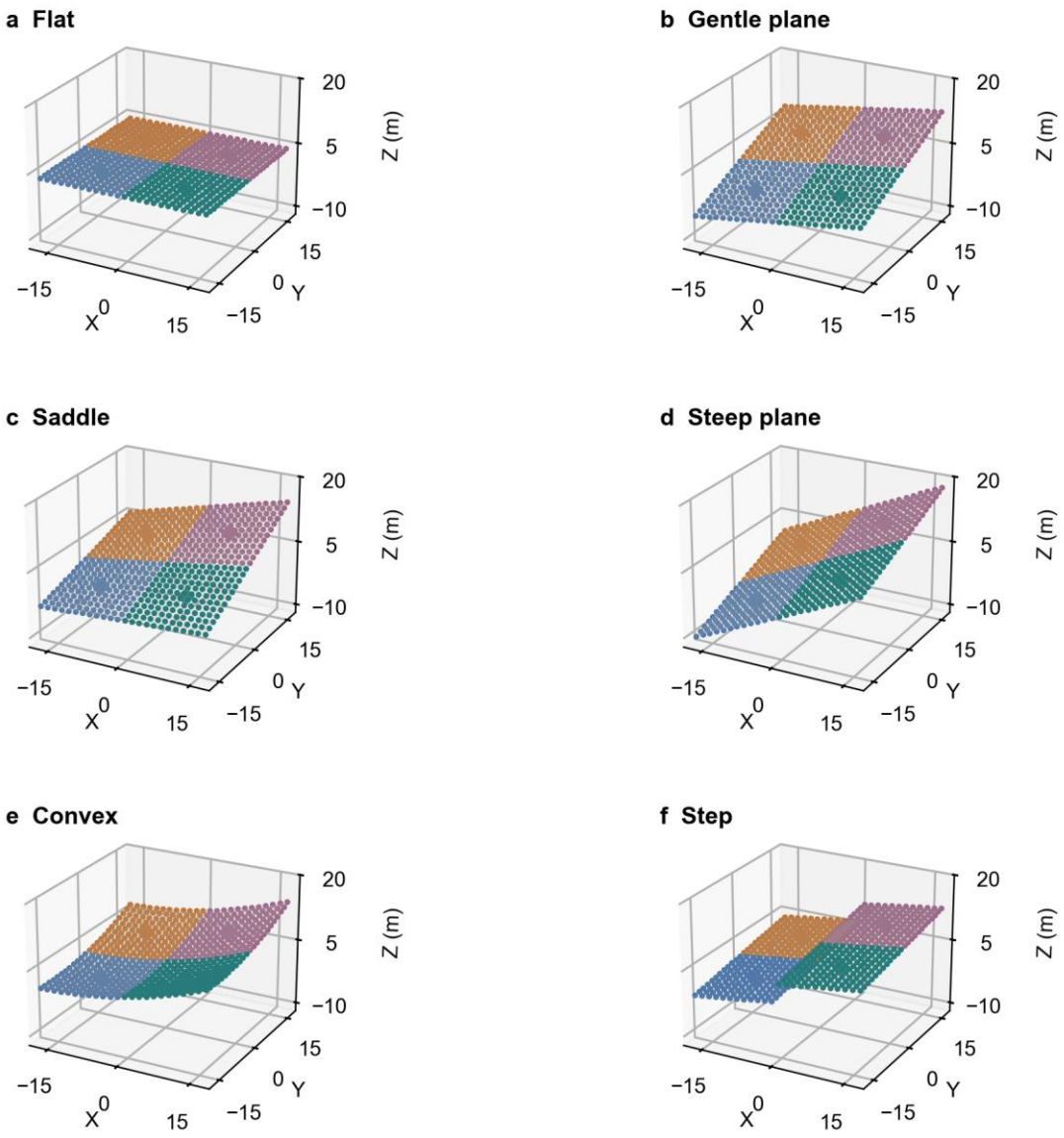


Fig. 1 Dense reference receivers on six controlled terrains. Quadrant colours identify the four acquisition timing groups; each dense group has 144 receivers. All axes use metres and common limits. Gentle and steep surfaces are inclined planes, whereas the saddle, convex and stepped surfaces are nonplanar. Sparse comparison layouts sample the same horizontal envelope; they are not shown as compact dense patches

The timing experiment draws 200 source identities. Their horizontal coordinates are uniform over [−10, 10] m, their depth below the local terrain is uniform over 32–65 m, and origin times are uniform over 300–600 ms. The same identity retains its horizontal coordinates, local depth and random timing innovations across configurations; its absolute Z coordinate changes with terrain height. These are paired numerical trials, not 200 observations at each setting that can be pooled as independent physical events. The chosen source and receiver domains keep sources below, and receivers above, the model interface.

2.2 Independently implemented propagation and calibration

The reference Earth model has two isotropic layers separated by a horizontal interface at Z = −15 m. P velocities are 4000 and 4800 m/s above and below the interface; S velocities are 2300 and 2800 m/s. For each phase, the travel time is the minimum of the path-integrated slowness. With h1 and h2 the receiver and source vertical distances to the interface and R their horizontal separation, the forward generator solves for horizontal slowness p using Snell's law:

$$h_1 \frac{pv_1}{\sqrt{1-p^2v_1^2}} + h_2 \frac{pv_2}{\sqrt{1-p^2v_2^2}} = R, \qquad T = \sum_{l=1}^{2} \frac{h_l}{v_l\sqrt{1-p^2v_l^2}}. \tag{1}$$

The inverse solver does not call the forward generator. It minimises travel time from source s to receiver r over the interface crossing coordinate a using a vectorised Newton iteration and evaluates the source derivative through the envelope theorem:

$$T(\mathbf{s},\mathbf{r}) = \min_{\mathbf{a}:a_z=-15} \left(\frac{\|\mathbf{r}-\mathbf{a}\|}{v_1} + \frac{\|\mathbf{s}-\mathbf{a}\|}{v_2} \right). \tag{2}$$

Independent implementations reduce the risk of accepting a solver merely because its own numerical approximation generated the test arrivals. The two solvers nevertheless share the same physical layer assumptions. Their agreement is an implementation check, not validation of a two-layer approximation to a real slope. Fifteen separately sampled sources are used to compare travel times and analytic gradients with central finite differences. A noise-free, three-start inversion checks recovery without starting at the true source.

Layer velocities are estimated from 12 separate calibration sources with known coordinates and origin times. These sources cover horizontal coordinates within ±14 m and Z between −64 and −24 m. Independent Gaussian phase errors are added to their generated arrivals. Four log-velocity parameters are fitted with bounded least squares. In the principal experiment the calibration fit uses the phase variances plus a conservative 0.03 ms shared-clock covariance, although no shared calibration-clock error is injected. Test locations do not enter calibration. The calibration set is held fixed within each terrain while the test timing innovations vary. Known calibration origin times are a favourable assumption: field calibration would require controlled sources or another independently timed source reference.

2.3 Shared timing errors and covariance-aware inversion

For phase p and receiver i in timing group g(i), the observation comprises origin time, propagation time, an independent picking perturbation and an array-shared clock perturbation:

$$t_{pi} = t_0 + T_p(\mathbf{s},\mathbf{r}_i;\mathbf{v}) + \epsilon_{pi} + c_{g(i)}, \quad \epsilon_{pi} \sim \mathcal{N}(0,\sigma_p^2), \quad c_g \sim \mathcal{N}(0,\sigma_c^2). \tag{3}$$

The standard deviations of the independent P and S errors are 0.12 and 0.18 ms. Shared clock standard deviations are 0, 0.03, 0.15 and 0.50 ms. A clock draw is common to both phases and every receiver within a group, and independent between groups. The 0.15 ms case is the principal comparison. It is a prescribed synthetic error scale, not a measurement of an installed acquisition system.

Stacking the 2N arrivals produces covariance $C = D + \sigma c^2 BB^T$, where D contains the phase-dependent independent variances and B maps phase-receiver observations to four timing groups. Generalised least squares (GLS) estimates the hypocentre and unknown origin time by minimising the full covariance-weighted residual. The predicted arrival vector F(q) contains origin time plus phase travel time at every receiver:

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \{\mathbf{t}-\mathbf{F}(\mathbf{q})\}^T \mathbf{C}^{-1} \{\mathbf{t}-\mathbf{F}(\mathbf{q})\}, \quad \mathbf{q} = (s_x, s_y, s_z, t_0)^T. \tag{4}$$

A low-rank symmetric whitener implements the covariance efficiently. Its identity $W\,C\,W^T = I$ is checked numerically. The diagonal baseline uses the same marginal variances, $\sigma p^2 + \sigma c^2$, but discards covariance between observations. Thus the two methods share generated data, geometry, calibration and optimisation settings; their estimates can differ because their weighting differs. No waveform quality or true-location information is used to discard observations.

Bounded nonlinear least squares starts from X = Y = 0, Z = −40 m and an origin-time estimate based on the median observed arrival. Spatial bounds are ±30 m horizontally and −90 to −15.01 m vertically; origin time is constrained to 150–800 ms. All optimisation outcomes are retained, including solutions within 0.01 of a bound.

The same solver and stopping settings are used for the homogeneous, coordinate-flattened and velocity-perturbed baselines.

The elementary group-mean variance illustrates the channel-count issue. For n observations with common independent variance σ^2 and shared clock variance σc^2,

$$\mathrm{Var}(\overline{e}_g) = \frac{\sigma^2}{n} + \sigma_c^2, \qquad \mathrm{Var}_{\mathrm{diagonal}}(\overline{e}_g) = \frac{\sigma^2+\sigma_c^2}{n}. \tag{5}$$

The second term does not average away within the group. A diagonal model instead treats the marginal variance as independent and divides both terms by n. This illustrates a potential information floor for common timing modes. It does not imply a universal location-error floor: within-group differential arrivals and P–S differences can still constrain location while cancelling common clock errors.

2.4 Location uncertainty and structural stress tests

The local covariance of the four estimated parameters is obtained from the inverse weighted information matrix, with J the Jacobian of F with respect to source coordinates and origin time at the fitted solution. The spatial block accounts for the jointly estimated origin time. A nominal 95% location ellipsoid uses the three-degree-of-freedom chi-squared threshold:

$$\mathbf{V}_q = (\mathbf{J}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{J})^{-1}, \qquad (\mathbf{s}-\hat{\mathbf{s}})^{\mathsf{T}}\mathbf{V}_s^{-1}(\mathbf{s}-\hat{\mathbf{s}}) \le \chi^2_{3,0.95}. \tag{6}$$

Empirical coverage is the fraction of true test hypocentres inside that ellipsoid. The reported radius is its longest semi-axis, rather than a scalar error guarantee. The primary ellipsoid conditions on fitted velocities. A separate first-order calculation propagates the calibration log-velocity covariance through location sensitivity. Gv is the derivative of F with respect to the four log velocities, Vlog v their fitted covariance, and Sv the sensitivity of the estimated location and origin time to these parameters:

$$\mathbf{S}_v = -\mathbf{V}_q\mathbf{J}^{\mathsf{T}}\mathbf{C}^{-1}\mathbf{G}_v, \qquad \mathbf{V}_{q,\mathrm{total}} \approx \mathbf{V}_q + \mathbf{S}_v\mathbf{V}_{\log v}\mathbf{S}_v^{\mathsf{T}}. \tag{7}$$

This addition concerns uncertainty in four calibrated velocities. It does not cover unknown interface shape, lateral heterogeneity or correlated phase-picking bias. Coverage in the principal experiment is conditional on one calibration dataset per terrain. A supplementary joint-repetition experiment instead refits velocities for 24 independent calibration-noise realisations, each followed by 30 independently drawn test sources and arrival innovations. The 12 calibration positions remain fixed and distinct from test sources; both calibration generation and weighting use independent phase errors without a shared-clock term. Tests use 64 and 576 receivers, four timing groups and a 0.15 ms shared clock deviation on all six terrains. Results and uncertainty intervals are reported by calibration repetition in Online Resource 1.

Three propagation/coordinate baselines use the 64-receiver observations. Homogeneous inversion uses constant P/S velocities of 4500/2600 m/s. Coordinate flattening replaces each receiver's elevation by the terrain mean while retaining its horizontal position. A velocity perturbation multiplies all fitted layer velocities by 1.02. These deliberately misspecified alternatives define how good point accuracy can fail when model assumptions are wrong; they are not equal-complexity claims about all homogeneous or three-dimensional methods.

Four further stress tests use a separate set of timing innovations but the same 200 source identities: assumed clock deviations of 0.03 and 0.50 ms when the actual value is 0.15 ms; phase-specific group biases of 0.15 ms in addition to the common clock; and a lateral slowness anomaly. The anomaly is a Gaussian spatial perturbation with a peak fractional amplitude of 4%, integrated along the straight source–receiver segment using an 80-point midpoint quadrature. This is a controlled first-order path perturbation, not exact ray tracing through a heterogeneous medium. The inversion retains the unperturbed layered model and full clock

covariance. These tests examine whether a correctly modelled common clock protects inference against other errors.

2.5 Waveform generation and automatic arrival picking

A separate waveform experiment uses 100 source identities, reused across six terrains and six noise conditions. Records last 0.5 s at 8000 samples/s. Sources lie within the same spatial distributions as the timing experiment, have unknown origin times between 0.20 and 0.31 s, and pulse frequencies uniformly distributed between 310 and 370 Hz. Causal, smoothly started Ricker pulses represent P arrivals; S pulse frequencies are 0.88 times the P frequency. Receiver gains are independently uniform between 0.7 and 1.3. Shared clock jitter is injected before sampling.

Inputs are two simplified pre-projected phase components, with 12% cross-phase leakage. They do not represent an elastodynamic solution, three-component sensor response or an automatically separated field P/S wavefield. Radiation patterns, scattering, dispersive surface waves, attenuation and nonlinear soil behaviour are outside this generator. This controlled signal model allows automatic timing to be evaluated without suggesting that phase identification has been solved for real slope data.

Noise combines independent Gaussian samples, low-frequency random motion and coherent 180/300 Hz mechanical components. Ordinary input event-window signal-to-noise ratios (SNRs) are −12, −6, 0 and +6 dB. A second −6 dB condition adds an in-band transient coherent within each group and absent from the reference channel. A third contaminates the environmental reference with three times the array-mean clean component. The latter two conditions deliberately challenge spatial coherence and reference subtraction. They should not be interpreted as measured vegetation or vehicle spectra. Signal enhancement techniques address several kinds of microseismic noise (Abbasi et al. 2023), while classification is a distinct task (Xin et al. 2024).

Processing applies a fourth-order 120–850 Hz Butterworth bandpass using forward–backward filtering, followed by regression against an environmental reference. Regression coefficients are estimated only from the fixed 0.025–0.100 s event-free interval. This is an offline procedure; its noncausal filtering and availability of a quiet reference interval preclude a real-time latency claim. In the contaminated-reference case, subtraction of the reference's signal contribution is counted as signal distortion in the SNR decomposition.

A five-template bank uses P frequencies of 310, 325, 340, 355 and 370 Hz and corresponding S frequencies. Per-channel matched correlations are normalised by their pre-event standard deviation. A 129-sample running average of squared correlations is averaged over receivers to form a network score for each phase. The template is selected by maximising the smaller of the two phase peak scores over the fixed 0.105–0.445 s search interval. This algorithm does not receive the true event time, source position or frequency. Detection requires both phase evidence above threshold and the network S peak later than the P peak.

The detection threshold is 1.05 times the maximum network score across 80 independently generated ordinary noise-only calibration records. It is empirical and finite-sample; no population false-alarm guarantee is implied. Per-channel phase picks are correlation maxima within ±96 samples of the detected phase epoch, refined by parabolic interpolation. Every receiver is retained. Layered GLS then estimates the source from these picks. Comparison methods use a fixed phase-specific template or homogeneous propagation with the same detected bank picks. They are internal component controls, not an exhaustive comparison with contemporary learned pickers or migration methods. Conventional recognition and timing have a long methodological history (Allen 1978).

After localisation, predicted relative arrivals align the processed traces for equal-weight delay-and-sum stacking:

$$y_p(t) = \frac{1}{N}\sum_{i=1}^{N} x_{pi}\,(t + \Delta\hat{T}_{pi}), \qquad \Delta\hat{T}_{pi} = \hat{T}_{pi} - \frac{1}{N}\sum_{j=1}^{N} \hat{T}_{pj}. \tag{8}$$

SNR is calculated from separately saved clean and residual-noise contributions over a 20 ms window for each phase. Identical physical event-window definitions are used before and after alignment. True arrivals define evaluation windows and noise scaling only; they are not passed to detection, picking or inversion. The reported SNR change includes filtering, reference subtraction and stacking. It is not assigned to beamforming alone. Held-out negative controls comprise 100 ordinary and 100 coherent-transient noise-only records with separate random seeds.

2.6 Evaluation and reproducibility

The principal timing design comprises 52,800 inversions, the structural experiment 4800 inversions and the waveform comparison 10,800 method-condition records. These totals count repeated analyses: independent source-identity counts are 200 for principal timing and 100 for waveform experiments. The joint calibration experiment adds 17,280 inversions using 720 new test identities arranged in 24 calibration batches. Geometry, methods, clocks and SNR conditions are paired by identity. Location accuracy is Euclidean hypocentral error; P90 is its empirical 90th percentile. Threshold success is error ≤2 m. For waveform experiments, the unconditional success fraction requires both detection and error ≤2 m, counting undetected records as failures. Detected-event P90 is reported separately, without removing badly located detected events.

Per-condition timing P90 intervals use 1000 event resamples. Wilson intervals describe empirical coverage proportions. Pooled waveform figures resample the 100 identities as blocks, keeping their six terrain outcomes together. Joint calibration coverage intervals use 2000 bootstrap resamples of the 24 calibration batches, retaining all 30 test outcomes within each selected batch. Descriptive intervals do not make six terrains independent samples or constitute multiple hypothesis tests. No P-value-based selection is used. The full trial records, calibration inputs, waveform example and random seeds are supplied in Online Resource 2. Online Resource 1 gives expanded parameters, supplementary results and validation checks. Python 3.13.9, NumPy 2.3.5, SciPy 1.16.3 and pandas 2.3.3 were used for the numerical runs.

OpenAI Codex assisted with code implementation, manuscript drafting and figure preparation. Numerical outputs were generated by executing the supplied scripts rather than by generating numerical claims as text. Scientific interpretation and the final submitted materials remain the responsibility of the human authors.

Figures were produced using Matplotlib 3.10.6. Vector originals and high-resolution RGB exports are supplied in Online Resource 2.

3 Results

3.1 Covariance changes accuracy and coverage

Independent forward and inverse solvers agreed to 8.88×10^{-14} ms; the maximum analytic-gradient discrepancy was 4.43×10^{-11} ms/m. The covariance-whitening identity error was 2.22×10^{-15}. Three distinct starts recovered a noise-free test hypocentre within 7.42×10^{-13} m. These checks support implementation consistency within the declared two-layer model. All 52,800 timing optimisations reported convergence; 357 boundary solutions, principally from deliberately misspecified alternatives, were retained.

At 64 receivers and 0.15 ms shared clock deviation, full-covariance inversion gave P90 errors between 0.778 and 0.819 m across six terrains, with empirical 95% coverage between 96.5% and 97.5% (Fig. 2). Diagonal weighting gave P90 errors between 2.058 and 2.120 m, despite using the correct marginal observation variances. Its coverage was only 33.0–34.5%. Table 2 reports the saddle comparison and its bootstrap intervals.

Thus the covariance choice affected both the point estimate and the credibility of the uncertainty statement, not just the plotted ellipsoid size.

Table 2 Saddle terrain arrival-time comparison with 200 paired events

Method	P90 (m)	95% P90 interval (m)	Coverage (%)	Boundary n
Layered GLS	0.790	0.732–0.848	97.5	0
Layered diagonal	2.058	1.910–2.364	34.0	0
Homogeneous GLS	3.608	3.451–3.867	0.0	0
Flattened GLS	26.572	25.786–27.301	0.0	11
Velocity +2%	1.760	1.693–1.886	0.0	0

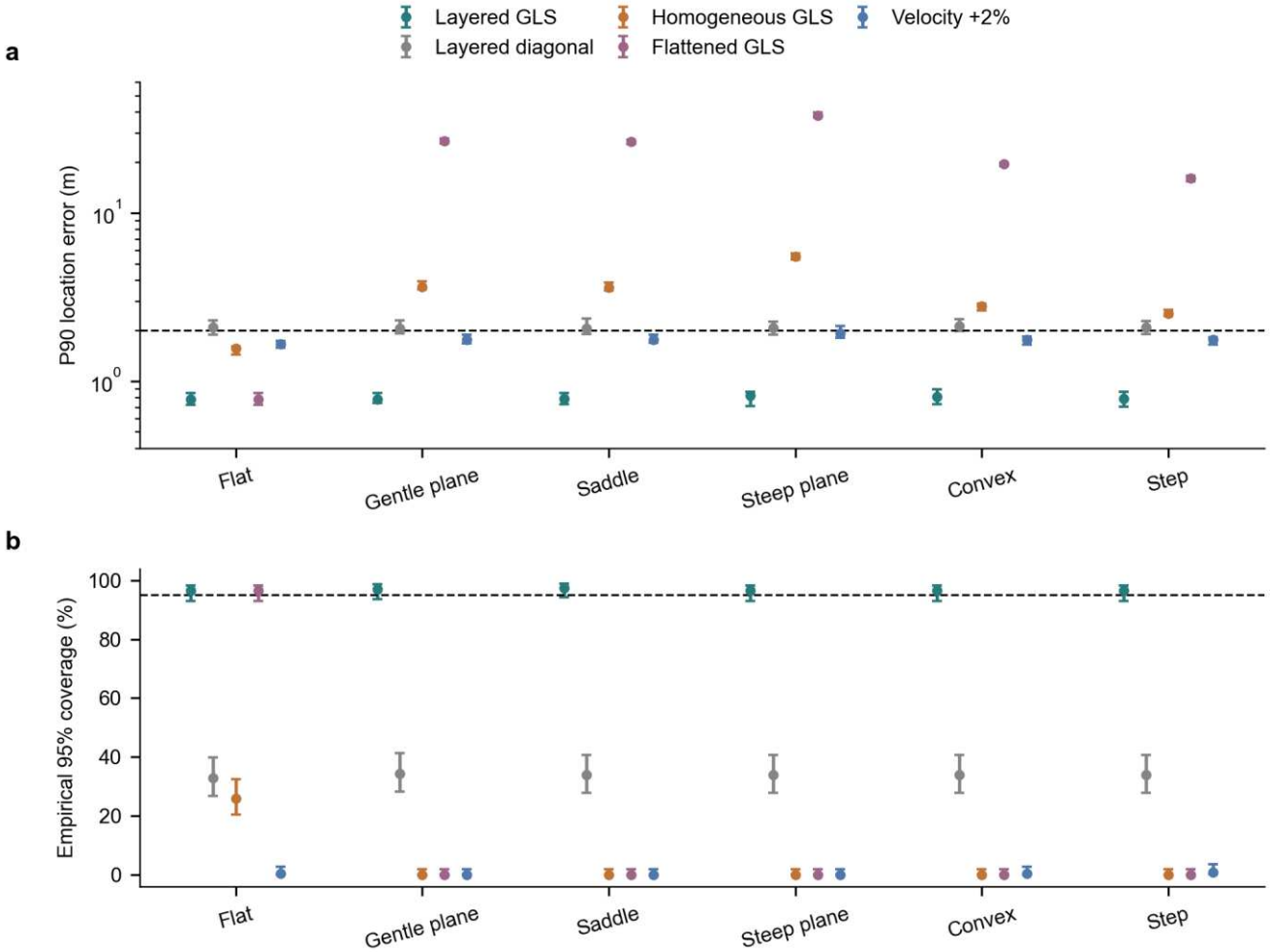


Fig. 2 Accuracy and uncertainty under five inversion assumptions. a P90 Euclidean location error; vertical intervals are 95% event-bootstrap intervals. b Empirical coverage of nominal 95% conditional location ellipsoids; intervals are Wilson 95% intervals. Each point uses 200 paired event identities, 64 receivers and 0.15 ms shared clock deviation. Dashed lines mark 2 m and 95%. The location-error axis is logarithmic

Increasing receiver count sharpened the full-covariance estimates while exposing the diagonal model's false precision (Fig. 3). With 576 receivers, full-covariance P90 errors were 0.320–0.332 m, whereas diagonal P90 errors remained 2.082–2.167 m. Diagonal coverage fell to 3.5–4.0%. Additional arrivals supplied differential timing information to GLS, but could not make array-shared errors independent. The comparison holds the

outer aperture and four timing groups fixed; it does not demonstrate the benefit of enlarging the monitored area or adding independently clocked groups.

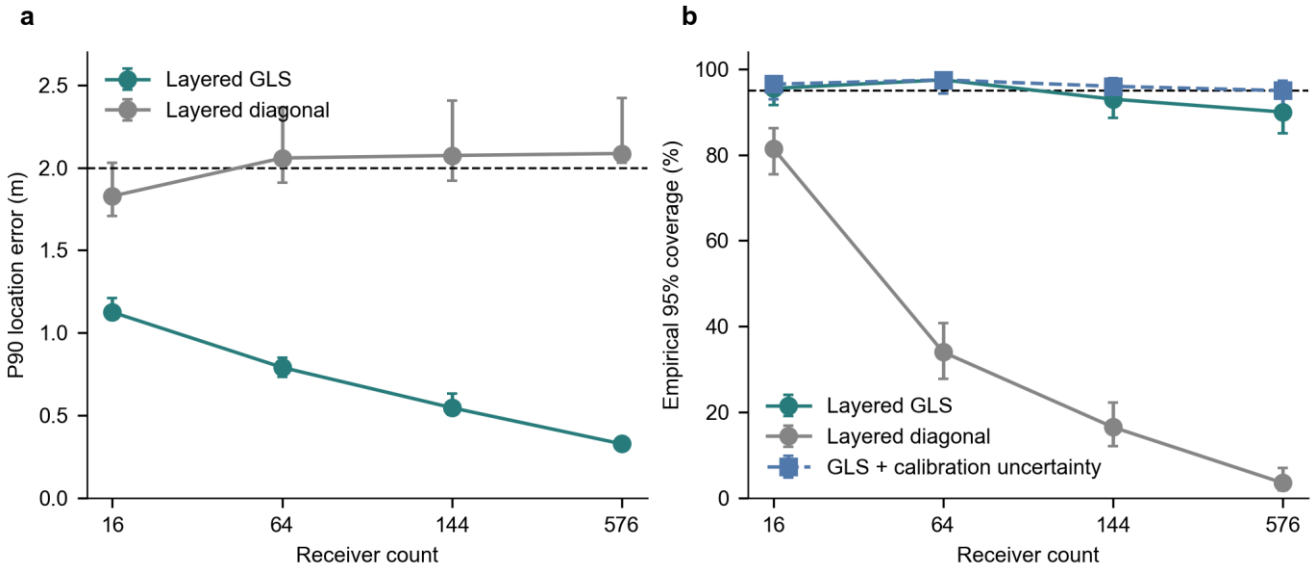


Fig. 3 Receiver count at fixed 34.5 m outer envelope on the saddle terrain. a P90 error with event-bootstrap intervals. b Conditional ellipsoid coverage with Wilson intervals, and coverage after velocity-calibration uncertainty propagation. Each count uses the same 200 source identities and four timing groups at 0.15 ms shared clock deviation. Other terrain results are supplied in Online Resource 2

Conditional full-covariance coverage also fell below nominal at the densest layout, to 90.0–92.0%. Calibration uncertainty becomes relevant as independent observation errors shrink. Adding first-order velocity-calibration covariance increased dense-layout coverage to 94.0–96.0% and modestly enlarged the reported radius. This is evidence for retaining calibration uncertainty in dense-array evaluation. It is not evidence that a first-order ellipsoid will remain calibrated under an unknown velocity structure.

Independent calibration repetitions confirmed that this effect was not confined to the original fitted velocities. Across six terrains and 24 calibration batches, 576-receiver GLS conditional coverage was 84.6–86.9%, increasing to 94.9–95.7% after calibration propagation. Diagonal coverage remained 3.1–3.8%. Each terrain-count-method condition contains 720 test outcomes, but the uncertainty intervals resample 24 calibration batches rather than treating their shared velocity fits as independent observations (Table S7). This extension remains conditional on known calibration positions and the correct two-layer model.

3.2 Accurate terrain coordinates do not remove propagation bias

With calibrated layered propagation and correct coordinates, the six terrains gave similar central-case accuracy. The outcome does not establish that nonplanar layouts universally outperform planar layouts. Gentle and steep surfaces are inclined planes; their elevations vary while their receivers remain coplanar. On the flat terrain, replacing receiver elevations by their mean changed nothing, as expected. On the other terrains, using flattened coordinates produced P90 errors of 16.04–38.08 m, with the largest value on the steep plane. This is primarily coordinate-model inconsistency, not a fair test of a separately designed flat installation.

The fixed homogeneous model gave P90 errors of 1.56 m on the flat terrain and 2.53–5.53 m on the remaining surfaces. Coverage fell to zero on the non-flat surfaces. The flat result illustrates that a seemingly acceptable error threshold can arise from a favourable combination of depth, geometry and an effective constant velocity. It cannot validate constant-velocity propagation for another slope. Independent calibration or a justified site model is needed before assigning performance across terrains.

A uniform +2% velocity perturbation gave P90 errors of 1.66–1.93 m, below the selected 2 m P90 reference for all surfaces, but coverage of only 0–1%. A point-error criterion therefore missed a serious uncertainty failure. This distinction is central to reporting: a synthetic result can satisfy one distance threshold while its nominal confidence regions almost never contain the true source.

3.3 Covariance misspecification and lateral structure

Underestimating the shared clock deviation as 0.03 ms increased P90 errors to 1.55–1.63 m and reduced coverage to 36.0–37.5% (Fig. 4). Assuming 0.50 ms instead retained P90 errors of 0.78–0.81 m and coverage of 95.0–96.0%. This asymmetry applies to the present geometry and error scales. It supports checking sensitivity to the assumed covariance rather than choosing a timing deviation simply to tighten the reported ellipsoid.

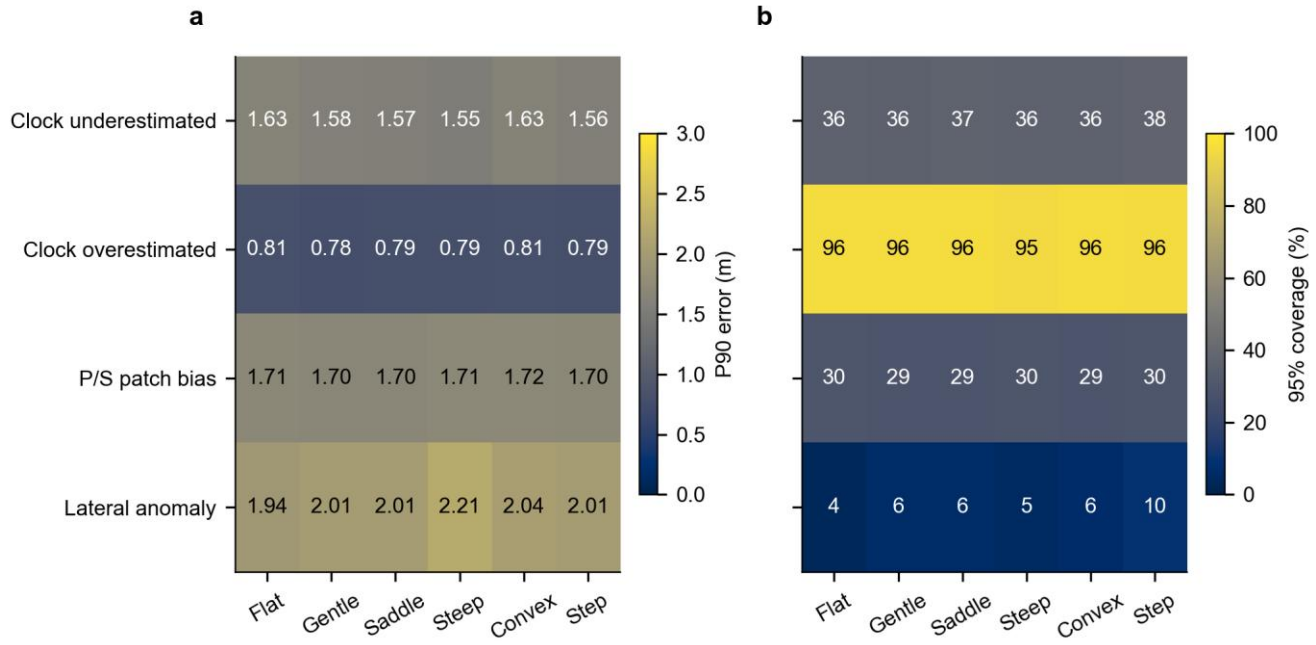


Fig. 4 Failure of conditional location inference under covariance and propagation stress. a P90 error. b Empirical 95% ellipsoid coverage. Each cell uses 200 paired sources and 64 receivers. Actual shared clock deviation is 0.15 ms; the first two rows assume 0.03 and 0.50 ms. The additional phase-specific bias is 0.15 ms. The lateral anomaly is a 4% peak slowness perturbation integrated along straight paths, not an exact heterogeneous ray solution

An extra phase-specific group bias reduced coverage to 29.0–29.5%, although P90 errors were approximately 1.70–1.72 m. The common-clock model cannot absorb a perturbation that differs between P and S within the same group. The lateral anomaly was more damaging to confidence calibration: P90 errors were 1.94–2.21 m, while coverage was 3.5–9.5%. These failures remain visible even though the generator used only a small first-order slowness perturbation. Correct common-clock covariance is therefore necessary for the central comparison, but insufficient for dependable inference in an inadequately modelled slope.

3.4 Waveform-derived performance and negative controls

Figure 5 traces the processing of a predetermined saddle event at 0 dB input SNR: observed and clean phase components, the detection score, arrival-time errors, and the estimated-delay stack. The example was selected by terrain, SNR and event index before ranking performance. Pick offsets contain injected clock errors and template-related timing shifts. A common offset can be absorbed into estimated origin time; differences between phases or receivers affect the hypocentre. Thus an absolute picking bias is not interchangeable with location error.

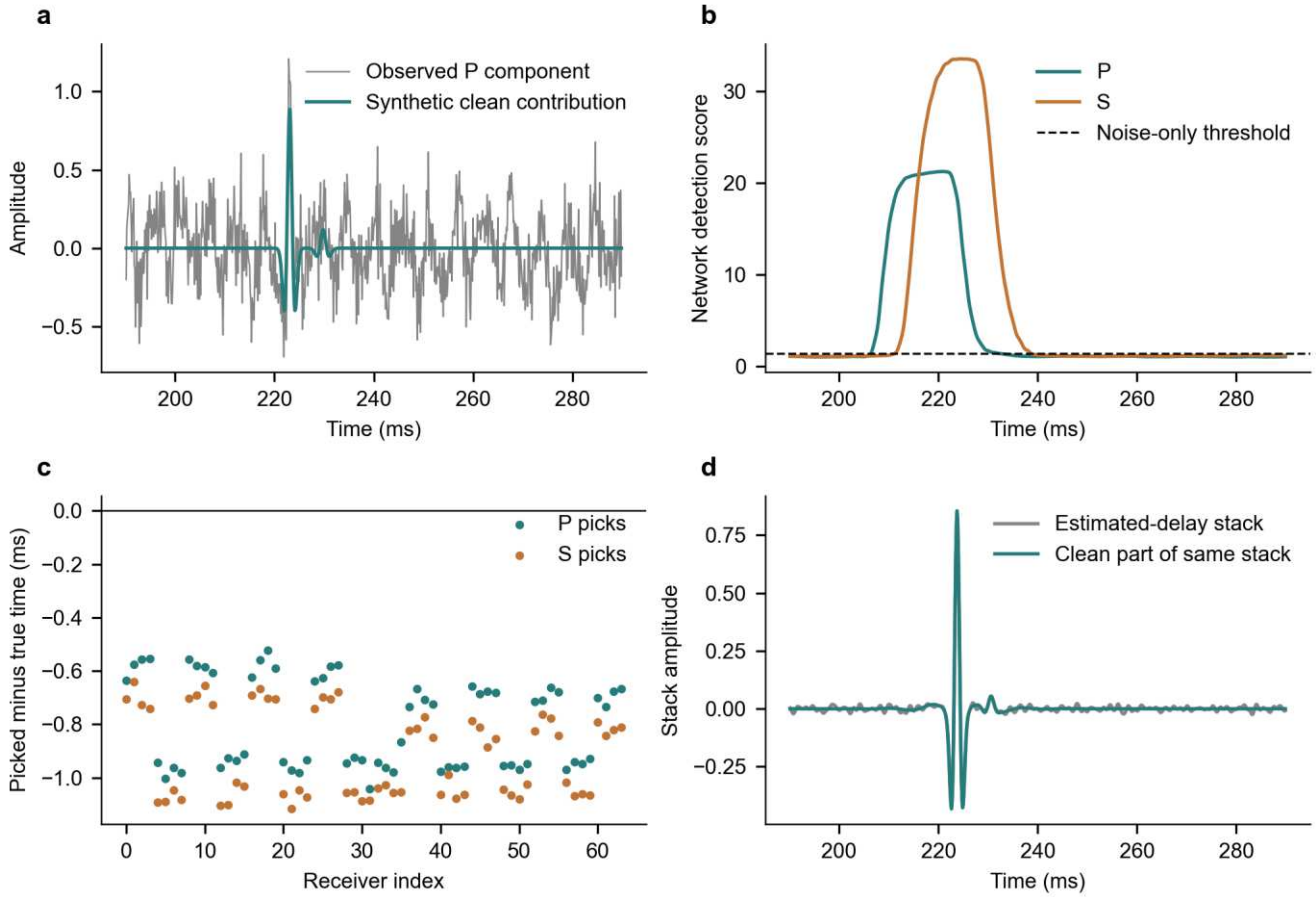


Fig. 5 Predetermined waveform example on the saddle terrain at 0 dB input event-window SNR, event identity 0. a Observed P component and its saved clean contribution at receiver 0. b P/S network scores and calibrated noise-only threshold. c Automatically picked minus true arrivals, without removing the shared clock offsets. d Estimated-delay stack and its clean contribution. Truth is used for plotting and evaluation only

Under ordinary noise at −6 dB, template-bank layered inversion detected 588/600 records (98.0%) and correctly detected and located 96.17% within 2 m. At 0 dB, the unconditional fraction reached 99.83%, with terrain-specific detected-event P90 errors of 0.798–0.868 m. At −12 dB, detection remained 95.17%, but unconditional ≤2 m success fell to 2.83% and detected P90 errors reached 12.46–14.98 m. A high detection fraction therefore did not ensure accurate arrival differences or source locations (Fig. 6; Table 3).

The fixed-template control performed better on point accuracy in this controlled pulse family: its pooled detected-event P90 at 0 dB was 0.485 m, versus 0.822 m for the bank. The bank also had larger absolute pick offsets. Selecting a template by network detection evidence is therefore not equivalent to minimising timing bias. We retain this unfavourable comparison rather than claiming template-bank superiority. Homogeneous inversion continued to show terrain-dependent bias even when detection and picks matched the layered method.

Table 3 Pooled waveform results under ordinary noise

Input SNR (dB)	Records	Detection (%)	Detected ≤2 m (%)	Detected P90 (m)	Median SNR gain (dB)
-12	600	95.17	2.83	14.702	24.34
-6	600	98.00	96.17	0.950	27.86
0	600	99.83	99.83	0.822	27.87
6	600	100.00	100.00	0.811	27.88

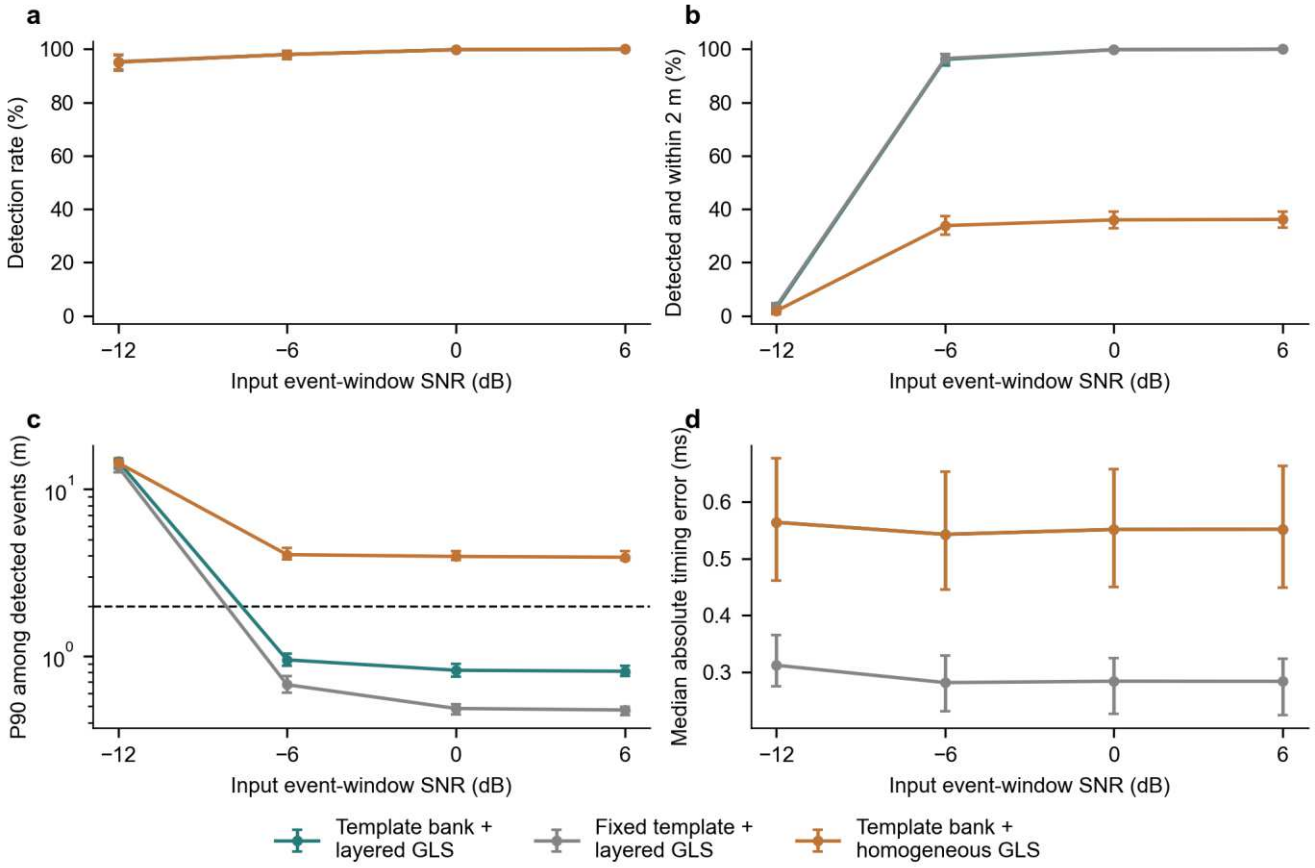


Fig. 6 Waveform-derived evaluation across ordinary noise conditions. a Detection fraction. b Unconditional detected-and-within-2-m fraction. c P90 error among all detected events. d Median event-level absolute pick error after averaging paired terrain summaries. Each SNR uses 100 identities reused on six terrains, giving 600 records per method. Intervals resample identities as blocks; repeated terrain outcomes are retained together. The error axis in c is logarithmic

At −6 dB, coherent-transient interference reduced unconditional detected-and-within-2-m success from 96.17% to 76.50%. Detected-event P90 errors rose to 56.41–63.89 m across terrains, despite comparatively small median errors for many records. Reference contamination reduced pooled unconditional success to 22.17%, with terrain-specific P90 errors of 5.01–13.97 m. These tails are retained, including boundary solutions.

Held-out ordinary noise produced 0/100 false positives. Coherent-transient noise produced 46/100. Zero observed ordinary false positives does not imply a zero population rate, and the contrasting transient result shows sensitivity to an unrepresented noise family (Fig. S1, Online Resource 1). At −6 dB the ordinary median same-window SNR improvement was 27.86 dB; favourable signal enhancement did not prevent the localisation failures in the stress conditions. No waveform confidence coverage is claimed because the timing-error scales used for inversion were not estimated from these waveform residuals.

The low-SNR and interference outcomes distinguish array sensitivity from reliable source identification. Coherent energy unrelated to the injected microseismic source can satisfy a network detector, and the same assumption that enables stacking can strengthen that interference. The waveform benchmark therefore supports only conditional performance for the stated phase components and noise families. It does not establish ≤2 m field accuracy, universal event classification or a 60 dB noise-rejection specification.

4 Discussion

4.1 Information is controlled by error structure

The receiver-count result is not a claim that dense arrays are unhelpful. Correctly weighted dense observations materially reduced error in the benchmark. The problem is assigning independent information to errors shared by an acquisition group. The simple variance expression in Eq. 5 explains one common mode, while the complete inversion also uses spatially varying derivatives and two phase velocities. Shared errors can be partly bypassed by differential constraints; their effect on location cannot be reduced to a single sample-count formula. Reporting both point-error quantiles and empirical confidence coverage preserves this distinction.

This interpretation is compatible with established least-squares and probabilistic location theory (Tarantola and Valette 1982; Lomax et al. 2000). The added value is a controlled, accessible demonstration across terrain, receiver density, calibration uncertainty and waveform-derived arrivals. We do not assert that explicit covariance is a new estimator. Rather, the benchmark isolates a failure that can be hidden when dense synthetic arrays are evaluated with diagonal weights or when an idealised clock error is averaged independently at each channel.

Calibration and test uncertainty should also be distinguished. More receivers can improve the location conditional on a fixed velocity model until uncertainty in that model becomes comparable to the remaining random error. The first-order propagation result captures that transition under the correct layer family. It cannot resolve structural discrepancy, as the lateral anomaly shows. A practical uncertainty budget should separate coordinate surveying, instrument synchronisation, arrival picking, calibrated parameters and unresolved propagation. Combining every source into one arbitrarily enlarged picking variance may obscure systematic bias even if it enlarges ellipsoids.

4.2 Terrain and array design implications

Elevation variation is not synonymous with noncoplanarity. Any set of receivers on a plane satisfies an affine relation between X, Y and Z, including the gentle and steep slopes studied here. In a homogeneous medium, reflection of a source across that plane preserves receiver distances if both sides are allowed. Restricting sources to the subsurface side removes that particular mirror ambiguity, but does not guarantee a well-conditioned depth estimate. Conversely, a planar array can recover three-dimensional location from near-field curvature and P/S timing under suitable velocity knowledge. A blanket claim that coplanar arrays cannot locate depth would be incorrect.

The strongest terrain result is the penalty for analysing surveyed elevations as if they were flat. It does not rank six natural terrain types as universally better or worse. The interface stays horizontal while receiver surface elevation changes, so path lengths through the upper layer also change. Source absolute depths and inverse sensitivities consequently change with terrain. A site-specific comparison would need an actual velocity model, access constraints, coupling quality and monitored volume. The chosen surfaces make those dependencies explicit without replacing them by a visually attractive three-dimensional layout.

Spacing constraints for directional beamforming must be considered separately from timing localisation. A conservative regular-array condition is $d \leq v_{min}/(2f_{max})$. With the upper-layer S velocity of 2300 m/s and f_{max} = 1000 Hz, it gives 1.15 m, so a 1.5 m spacing is not certified alias-free across the full P/S bandwidth. For that spacing the corresponding S-wave limit is about 767 Hz. Projected spacing is also insufficient at steep or stepped surfaces: Euclidean adjacent-receiver distance can be larger. The sparse equal-envelope layouts in this paper intentionally do not meet a regular dense-beamforming criterion at high frequencies.

Additional regular-line aperture/taper calculations are provided as supplementary context. Conventional Dolph–Chebyshev weighting trades mainlobe width for sidelobe reduction (Dolph 1946). It cannot correct a wrong velocity model, array clock correlation or an incorrect phase pick. Far-field angular half-power beamwidth and cut sidelobe level should not be read as near-field localisation error or a full three-dimensional grating-lobe guarantee. The journal comparison therefore uses them only as array-pattern descriptors, while evaluating localisation independently from actual arrival data.

4.3 Waveform discrimination and slope monitoring

The waveform experiment addresses one weakness of an arrival-only evaluation by estimating arrivals without injected-time search windows. Its template bank covers the pulse family used to generate the events, which favours detection compared with unknown natural source functions. Fixed-template controls quantify one aspect of that dependence, but they are not substitutes for real waveform diversity. Phase projection, a known event-free interval and an environmental reference also make this benchmark more tractable than unattended field operation. These assumptions should be tested with independently recorded or independently simulated wavefields before deployment.

Noise rejection should be expressed as a measurable processing result with a declared window and denominator. A factor of 1000 in amplitude corresponds to 60 dB; a factor of 1000 in power corresponds to 30 dB. Neither statement is equivalent to amplifying the useful signal by 1000. For N equal, perfectly aligned receivers with independent noise, the ideal equal-weight SNR gain is 10 log10 N dB: approximately 18.1 dB for 64 receivers. Correlated noise, picking errors and reference leakage alter that gain. The larger combined-processing gains in favourable synthetic cases include filtering and reference regression, so they cannot be attributed to channel summation or used as a universal hardware suppression specification.

Early-warning interpretation requires another level of evidence. Event rate, energy and spatial clustering can be informative, but they also depend on detector sensitivity, changing noise and catalogue completeness. The negative-control failure demonstrates why a rising detected count cannot automatically be called a rise in fracture activity. Microseismic catalogues should be considered alongside displacement, rainfall, geology and operational records, as suggested by complementary slope-monitoring approaches (Faillettaz et al. 2019; Fiolleau et al. 2023). No warning classifier, lead-time estimate or collapse forecast is tested here.

4.4 Limitations and next validation steps

All data are synthetic. The study omits full elastic scattering, surface-wave propagation, attenuation, source radiation, heterogeneous sensor coupling and actual vegetation or machinery recordings. The velocity calibration assumes known source origins and a correct interface. Timing ellipsoids rely on local linearisation and prescribed error variances; waveform confidence calibration is not established. The lateral perturbation is an approximate straight-path integral and only one anomaly family is used. A broad statement about realistic heterogeneous slopes would require independent three-dimensional propagation and a wider structural ensemble.

The receiver-count comparison uses four fixed timing groups and an equal outer envelope, not equal installation cost or alternative independent acquisition architectures. It also does not optimise sensor selection. Joint calibration repetitions vary noise and test sources but retain one calibration geometry and the correct layer family, so they do not establish robustness to unknown calibration locations or geology. The detector threshold is calibrated on a finite ordinary-noise sample, and held-out interference can lie outside that distribution. These limitations define the next tests: independent waveform propagation, calibrated picking covariance, alternative group-clock architectures, surveyed-coordinate perturbations, and ground-coupled measurements using

independently timed calibration sources. The present package supplies failure records as well as favourable cases so that these extensions can be compared transparently.

5 Conclusions

A controlled synthetic slope-array study shows that correct marginal timing variances are insufficient when acquisition errors are shared. Covariance-aware inversion uses additional differential arrival information, whereas diagonal weighting can report increasingly narrow but poorly covering ellipsoids as receiver density rises. Velocity-calibration uncertainty becomes relevant for dense arrays, and modest point error can conceal severe undercoverage under propagation or phase-bias misspecification. Automatic waveform tests and noise-only controls further show that enhanced coherent energy is not necessarily correctly located microseismic activity. These findings support joint reporting of detection, localisation error and uncertainty coverage under explicit terrain and noise assumptions. They provide a reproducible basis for subsequent field validation rather than a demonstrated early-warning instrument.

Statements and Declarations

Funding: No funding was received for this study.

Competing interests: The authors declare no financial or non-financial competing interests.

Author contributions: Lai Wei led the study and manuscript preparation. Bin Yan provided academic supervision. Xiangli Zhang provided guidance on the research concept. Yu Chang and Zehao Li provided supporting assistance. All five authors have agreed to the authorship and proposed submission.

Data availability: All data are synthetic. Trial-level results, receiver coordinates, calibration data, source inputs, a representative waveform record and analysis summaries supporting this study are supplied in Online Resource 2. No field seismic data were collected or analysed.

Code availability: The Python scripts for generation, inversion, uncertainty analysis, verification and figure production, together with dependency versions and execution instructions, are supplied in Online Resource 2.

Ethics approval: Not applicable. This study did not involve human participants or animals.

References

Abakumov I, Roeser A, Shapiro SA (2020) Arrival-time picking uncertainty: Theoretical estimations and their application to microseismic data. Geophysics 85:U65-U76. https://doi.org/10.1190/geo2019-0589.1

Abbasi S, Jayaram V, Akram J, Alam MI, Yu S (2023) Using geometric mode decomposition for the background noise suppression on microseismic data. Geophysical Prospecting 71:1420-1437. https://doi.org/10.1111/1365-2478.13379

Allen RV (1978) Automatic earthquake recognition and timing from single traces. Bulletin of the Seismological Society of America 68:1521-1532. https://doi.org/10.1785/bssa0680051521

Chmiel M, Roux P, Bardainne T (2019) High-sensitivity microseismic monitoring: Automatic detection and localization of subsurface noise sources using matched-field processing and dense patch arrays. Geophysics 84:KS211-KS223. https://doi.org/10.1190/geo2018-0537.1

Dolph CL (1946) A Current Distribution for Broadside Arrays Which Optimizes the Relationship between Beam Width and Side-Lobe Level. Proceedings of the IRE 34:335-348. https://doi.org/10.1109/jrproc.1946.225956

Faillettaz J, Funk M, Beutel J, Vieli A (2019) Towards early warning of gravitational slope failure with co-detection of microseismic activity: the case of an active rock glacier. Natural Hazards and Earth System Sciences 19:1399-1413. https://doi.org/10.5194/nhess-19-1399-2019

Fiolleau S, Uhlemann S, Wielandt S, Dafflon B (2023) Understanding slow-moving landslide triggering processes using low-cost passive seismic and inclinometer monitoring. Journal of Applied Geophysics 215:105090. https://doi.org/10.1016/j.jappgeo.2023.105090

Hou C, Li X, Chen Y, Li W, Liu K, Dong L, Sun D (2024) Optimization and Numerical Verification of Microseismic Monitoring Sensor Network in Underground Mining: A Case Study. Mathematics 12:3500. https://doi.org/10.3390/math12223500

Huang X, Li S, Dong L, Yang L (2026) Optimization Research on Sensor Network Layout for Microseismic Monitoring Based on Location Error Analysis. Sensors 26:4533. https://doi.org/10.3390/s26144533

Le Breton M, Bontemps N, Guillemot A, Baillet L, Larose É (2021) Landslide monitoring using seismic ambient noise correlation: challenges and applications. Earth-Science Reviews 216:103518. https://doi.org/10.1016/j.earscirev.2021.103518

Li T, Chen BR, Wang Q, Zhu XH, Wang X, Xie MX (2022) Influence Evaluation of Sensor Coordinate Error on Microseismic Source Location. Frontiers in Earth Science 10:873986. https://doi.org/10.3389/feart.2022.873986

Liang X (2024) Uncertainty of microseismic sources identification and probabilistic location in underground excavation. Frontiers in Earth Science 12:1419330. https://doi.org/10.3389/feart.2024.1419330

Lienert BR, Berg E, Frazer LN (1986) HYPOCENTER: An earthquake location method using centered, scaled, and adaptively damped least squares. Bulletin of the Seismological Society of America 76:771-783. https://doi.org/10.1785/bssa0760030771

Lomax A, Virieux J, Volant P, Berge-Thierry C (2000) Probabilistic Earthquake Location in 3D and Layered Models. In: Advances in Seismic Event Location. Springer, pp 101-134. https://doi.org/10.1007/978-94-015-9536-0_5

Mao W, Gubbins D (1995) Simultaneous determination of time delays and stacking weights in seismic array beamforming. Geophysics 60:491-502. https://doi.org/10.1190/1.1443786

Rost S, Thomas C (2002) Array seismology: methods and applications. Reviews of Geophysics 40:1008. https://doi.org/10.1029/2000RG000100

Shen T, Wang S, Jiang X, Peng G, Tuo X (2023) An Anisotropic Velocity Model for Microseismic Events Localization in Tunnels. Sensors 23:4670. https://doi.org/10.3390/s23104670

Tarantola A, Valette B (1982) Generalized nonlinear inverse problems solved using the least squares criterion. Reviews of Geophysics 20:219-232. https://doi.org/10.1029/rg020i002p00219

Tian X, Zhang J, Zhang W (2022) A hybrid workflow for updating 1D velocity model and microseismic event location. Journal of Applied Geophysics 200:104642. https://doi.org/10.1016/j.jappgeo.2022.104642

Wang S, Panning MP, Vance SD, Song W (2021) Underground Microseismic Event Monitoring and Localization within Sensor Networks. Sensors 21:2830. https://doi.org/10.3390/s21082830

Wang K, Tang CA, Ma K, Ma T (2022) Research on the Design of Coal Mine Microseismic Monitoring Network Based on Improved Particle Swarm Optimization. Applied Sciences 12:8439. https://doi.org/10.3390/app12178439

Xin B, Huang Z, Huang S, Feng L (2024) Ensemble Learning Improves the Efficiency of Microseismic Signal Classification in Landslide Seismic Monitoring. Sensors 24:4892. https://doi.org/10.3390/s24154892

Yang P, Gajewski D, Xie Y (2023) Gaussian-weighted crosscorrelation imaging condition for microseismic source localization. Geophysics 88:L65-L78. https://doi.org/10.1190/geo2022-0455.1

Yfantis G, Pytharouli S, Lunn RJ, Carvajal HEM (2021) Microseismic monitoring illuminates phases of slope failure in soft soils. Engineering Geology 280:105940. https://doi.org/10.1016/j.enggeo.2020.105940

Zhang Z, Arosio D, Hojat A, Zanzi L (2020) Tomographic Experiments for Defining the 3D Velocity Model of an Unstable Rock Slope to Support Microseismic Event Interpretation. Geosciences 10:327. https://doi.org/10.3390/geosciences10090327

Zhang X, Zhang W, Zhang J (2021) Detecting and locating microseismic events with stacking velocity analysis for surface monitoring. Journal of Applied Geophysics 195:104470. https://doi.org/10.1016/j.jappgeo.2021.104470

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf
- ESM2.zip