High-speed train elastic full waveform inversion over viaducts: new source implementation and resulting near-surface multiparameter resolution

Chen He, Thomas Bohlen, Jingbo Chen, Shoudong Huo, Hao Wang

Geophysical Journal International

Geophys. J. Int. (2026) **246,** 1–19 https://doi.org/10.1093/gji/ggag220
Advance Access publication 2026 June 17
Research Paper

High-speed train elastic full waveform inversion over viaducts: new source implementation and resulting near-surface multiparameter resolution

Chen He,[1,2] Thomas Bohlen,[1] Jingbo Chen,[2] Shoudong Huo[2] and Hao Wang[2]

[1] *Geophysics Institute, Karlsruhe Institute of Technology, 76187 Karlsruhe, Germany. E-mail: chen.he@kit.edu*
[2] *Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China*

Accepted 2026 May 28. Received 2026 May 11; in original form 2025 December 8

SUMMARY

High-speed train (HST) signals offer an abundant and eco-friendly seismic source along the railway, but their application to full waveform inversion (FWI) presents unique challenges due to complex source characteristics and field constraints. We develop an HST-based elastic FWI framework that models the train as a distributed pier-force: a chain of time-delayed excitations applied continuously along deeply embedded bridge piers. This source mechanism generates predominantly *S*- and Rayleigh-wave energy, providing enhanced shallow subsurface illumination. The synthetic anomaly experiment demonstrates that the distributed pier-force enables superior multiparameter reconstruction, particularly achieving higher density accuracy than the explosive and single-force sources. Additional sparse acquisition tests validate the robustness of the method under limited receiver coverage. The field data application successfully reconstructs subsurface layering through simultaneous source and parameter inversion, supporting the physical plausibility of the distributed source model. Overall, this study establishes the distributed pier-force as an effective and computationally efficient HST source for shallow subsurface characterization, while also highlighting limitations under realistic field conditions.

Key words: Elasticity and anelasticity; Numerical modelling; Controlled source seismology; Waveform inversion.

1 INTRODUCTION

Traffic vehicles, including high-speed trains (HSTs), automobiles and metro systems, operate continuously along railway networks and highways and generate repeatable ground vibrations along fixed transportation corridors (A.M. Kaynia *et al.* 2000; Y. Liu *et al.* 2022). These moving anthropogenic sources are now recognized as important contributors to seismic noise and as potential illuminators for subsurface imaging, particularly in urban environments (J.C. Groos & J.R.R. Ritter 2009; D.A. Quiros *et al.* 2016; K. Zhao *et al.* 2025). Among these moving sources, HSTs running on viaducts generate particularly stable and repeatable seismic wavefields with broadband spectral signatures controlled by train speed, axle geometry and bridge response, as demonstrated by analytical theory and field observations (F. Lavoué *et al.* 2020; Y. Liu *et al.* 2021a, b). Although railway networks have limited spatial distribution, HST-generated signals are abundant, highly repeatable, stable and environmentally friendly, making them an attractive and sustainable seismic source for exploration and monitoring (J. Forrest & H. Hunt 2006; G. Kouroussis *et al.* 2014; K. Sager *et al.* 2021; M. Rezaeifar *et al.* 2022; Y. Sheng 2023; T. Rebert *et al.* 2024; S. Yuan *et al.* 2024).

Several recent studies have further characterized the source radiation of HSTs and their propagation in complex near-surface structures using analytical, numerical and field approaches (J. Cao & J. Chen 2019; F. Lavoué *et al.* 2020; Y. Liu *et al.* 2021b). Building on these observations, two main classes of methods have been developed for subsurface exploration with HST signals: seismic interferometry and active-source inversion. Seismic interferometry treats HST vibrations as background noise to perform ambient noise tomography (A. Inbal *et al.* 2018; J. Shao *et al.* 2022), while active-source inversion, such as full waveform inversion (FWI), treats HSTs as controlled seismic sources and directly fit the recorded waveforms.

The application of FWI to HST signals represents a frontier research area for subsurface imaging. Recent investigations have explored the HST excitations for elastic FWI, reverse-time migration, traveltime inversion and Rayleigh-wave FWI, enabling the reconstruction of near-surface structure (G. Hu *et al.* 2019; J. Luo *et al.* 2022, 2023; N. Zhou *et al.* 2024; H. Wang 2025; L. Wang *et al.* 2025). Despite the promising developments, current HST-based seismic inversion faces a fundamental limitation in its source representation. Most existing approaches model the HST passage over a bridge as a single source located at the surface (J. Luo *et al.* 2022; H. Wang 2025), neglecting the critical physical reality that bridge piers penetrate deeply into the subsurface. To address this issue, Z. Wang *et al.* (2020) proposed a bridge–pier coupling mechanism to characterize how HST-induced vibrations are transferred into the ground. Fol-

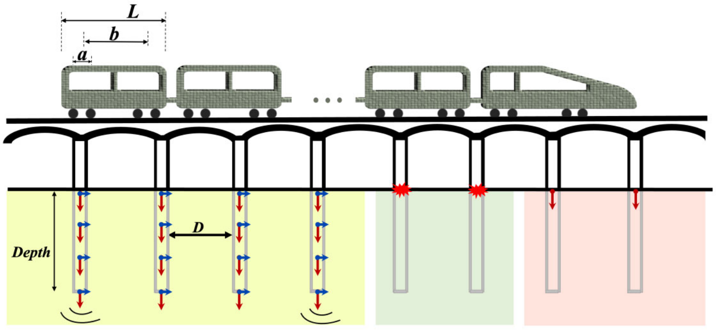

Figure 1. Schematic illustration of an HST moving on a viaduct comparing three source models: (a) Distributed point forces along bridge piers (yellow area) with vertical (red arrows) and horizontal (blue arrows) force components distributed along the embedded pier depth with depth-dependent attenuation; (b) Explosive source at surface excitation (green area) and (c) Single force at surface excitation (pink area). Train geometry parameters: L = carriage length, a and b = axle spacing, D = pier interval.

lowing this physical concept, we represent the transfer of train-induced loads into the subsurface through the bridge piers. However, unlike Z. Wang *et al.* (2020), who model the load transfer using a finite number of discrete source points, we implement the source on the staggered grid as a depth-distributed force along the embedded pier. This distributed loading mechanism fundamentally changes the source wavelet characteristics, wavefield excitation patterns and subsurface illumination relative to a simple surface source.

Source implementation fundamentally affects wavefield generation and FWI performance. Explosive source produces omnidirectional *P*-wave radiation, single surface force generates strong Rayleigh waves with limited subsurface coupling, while distributed pier source preferentially excites *S* waves and efficiently couple energy into surface modes throughout the pier depth, providing superior shallow illumination. Moreover, previous studies have been predominantly constrained to synthetic experiments, lacking validation through field data applications. Additionally, a further constraint of HST-based imaging lies in the fixed railway infrastructure, which results in strongly limited coverage compared to conventional controlled-source surveys.

The key innovations of this study are: (a) implementation of a realistic distributed pier-force source, with comprehensive wavefield analysis demonstrating that it predominantly generates Rayleigh and *S* wave; (b) synthetic reconstruction tests showing that this distributed source configuration achieves higher density resolution compared to a single source; (c) initial field data application showing that the effective HST source signature can be successfully estimated from observed data, supporting the practical feasibility of the distributed source model.

This paper is organized as follows: Section 2 presents the methodology, including HST source function and its characteristics analysis, distributed pier-force implementation and wave propagation theory, source wavelet correction filter and the FWI framework. Section 3 demonstrates synthetic data experiments using anomaly models and sparse acquisition geometries. Section 4 presents field data applications with source wavelet estimation and inversion results. Section 5 discusses the limitations and future research directions of the HST-based FWI method. While Section 6 summarizes the key conclusions.

2 METHODOLOGY

2.1 HST source function

This study considers the seismic sources induced by HSTs on viaducts (Fig. 1) (H.H. Hung & Y.B. Yang 2001; D.P. Connolly *et al.* 2015), incorporating three source implementations: (1) distributed pier-force (yellow area, Fig. 1), applied continuously along the embedded pier depth (0–60 m) following Z. Wang *et al.* (2020); (2) explosive source (green area, Fig. 1), modelled as isotropic stress excitation concentrated at the surface and (3) surface force (pink area, Fig. 1), a single force applied at the surface following J. Luo *et al.* (2022).

Accurate simulation of HST wavefields relies on precise source–time function, derived from the load history at bridge piers. Based on the observation by T. Rebert *et al.* (2024), the wavelets generated by different sleepers exhibit strong similarities and high repeatability in field measurements, we assume the source wavelets for each bridge pier excitation are identical.

Mathematically, the source time function for an HST passing each bridge pier can be expressed as (Z. Wang *et al.* 2020):

$$f_{\mathrm{hst}}(t) = \sum_{q=1}^{N_c} \left[FS_w\left(t - \frac{L}{v}(q-1)\right) + FS_w\left(t - \frac{L}{v}(q-1) - \frac{a}{v}\right) + FS_w\left(t - \frac{L}{v}(q-1) - \frac{a+b}{v}\right) + FS_w\left(t - \frac{L}{v}(q-1) - \frac{2a+b}{v}\right)\right]. \quad (1)$$

Table 1. The parameters of the HST moving source.

	Parameters
HST	$D = 32\,\mathrm{m}$, $L = 25\,\mathrm{m}$, $a = 2.5\,\mathrm{m}$, $b = 17.5\,\mathrm{m}$, $N_c = 8$, $F = 170\,\mathrm{kN}$

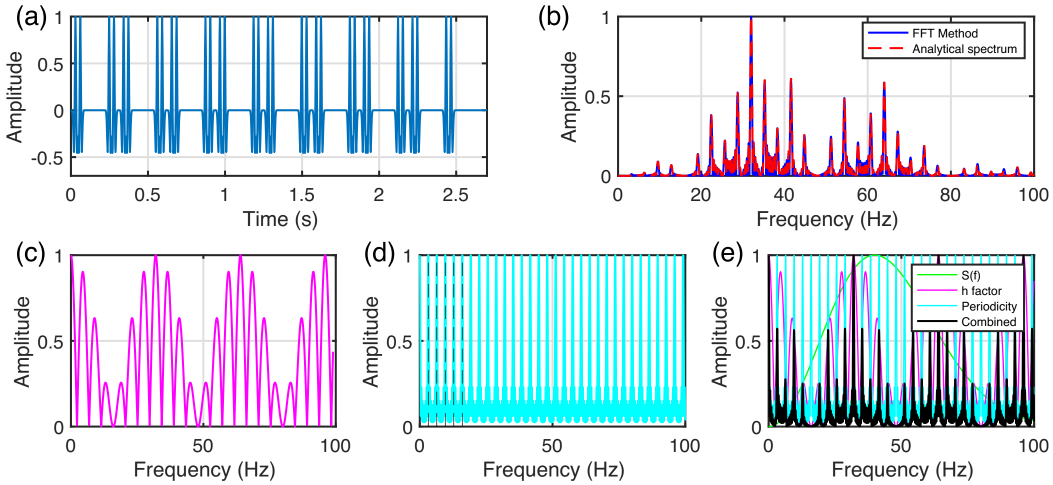

Figure 2. The HST distributed moving source signatures: (a) source wavelet, (b) source spectrum, (c) $\hat{h}$ factor (d) periodicity factor and (e) factor decomposition.

Here, $S_w(t)$ denotes the wheel–rail contact pulse associated with a single wheel passing the bridge pier. The resulting $f_{\mathrm{hst}}(t)$ is the composite load history generated by the entire train passage over a pier. N_c and L represent the number of carriages and the length of each carriage, respectively. D is the distance between adjacent bridge piers. F is the load on a single wheel, assuming equal load on all wheels. v is the velocity of the HST. a and b represent the axle-spacing parameters. The parameter settings of our HST source are outlined in Table 1.

Applying the Fourier transform, the corresponding frequency-domain expression of eq. (1) becomes

$$\hat{f}_{\mathrm{hst}}(\omega) = F \cdot \hat{S}_w(\omega) \cdot \frac{1 - e^{-i\omega \frac{L}{v} N_c}}{1 - e^{-i\omega \frac{L}{v}}} \cdot \hat{h}. \tag{2}$$

with the $\hat{h}$ factor given by

$$\hat{h} = 1 + e^{-i\omega \frac{a}{v}} + e^{-i\omega \frac{a+b}{v}} + e^{-i\omega \frac{2a+b}{v}}. \tag{3}$$

The periodicity factor $(1 - e^{-i\omega \frac{L}{v} N_c})/(1 - e^{-i\omega \frac{L}{v}})$ accounts for the periodic contributions of the HST carriages. It reaches a maximum at the carriage-frequency harmonics $f_c = \ell v/L$, where $\ell = 0, 1, 2, \ldots$ (H. Wang & J. Chen 2023).

2.2 HST source characteristics

Having established the formulation of the HST moving source, we now analyse its fundamental characteristics. Eqs (1) and (2) demonstrate the source's strong velocity dependence, revealing how train speed governs key aspects of wavefield generation. For this study, we analyse an 8-carriage train configuration moving at 288 km hr^{-1} (80 m s^{-1}) as a representative case. Each wheel's contact force is modelled using a Ricker wavelet with a dominant frequency of 40 Hz.

Fig. 2 displays this new HST source signature. Fig. 2(a) shows the time domain source wavelet with periodic characteristic pattern. Each HST carriage generates four distinct pulses corresponding to its four wheels. The regular spacing between carriage groups is clearly visible ($\Delta\tau_c = L/v = 0.3125\,\mathrm{s}$). Fig. 2(b) shows the frequency domain source spectrum calculated by applying FFT method to eq. (1), together with the analytical spectrum of eq. (2). The agreement between these two methods validates that the direct implementation of eq. (2) is mathematically correct. Discrete spectral peaks appear at regular intervals, confirming the periodic nature. Although the spectral amplitude below 10 Hz is weaker than that of the dominant harmonic peaks, the spectrum still contains non-zero low-frequency components generated by the finite carriage sequence and train periodicity.

Fig. 2(c) shows the wheel-configuration factor $\hat{h}$, which describes how the physical arrangement of the wheels affects the frequency spectrum. Different wheel-spacing parameters a and b lead to different $\hat{h}$ factors and therefore produce different spectral signatures.

Fig. 2(d) displays the periodicity factor of eq. (2), which represents a finite geometric series,

$$\frac{1-e^{-i\omega\frac{L}{v}N_c}}{1-e^{-i\omega\frac{L}{v}}}=\sum_{q=0}^{N_c-1}e^{-i\omega\frac{L}{v}q}. \tag{4}$$

Physically, this corresponds to coherent summation of phase-shifted contributions from each carriage. In the low-frequency limit ($f\to 0$), the exponential terms $e^{-i\omega\frac{L}{v}q}\approx 1$, causing all N_c terms to constructively interfere. This produces spectral peaks at harmonics of the fundamental carriage frequency $f_c=v/L$ and contributes to the non-zero low-frequency components through coherent stacking. Fig. 2(e) shows the spectral factor decomposition. $S(f)$ is the Ricker wavelet spectrum with peak 40 Hz. The combined factor is the $\hat{h}$ factor multiplied by the periodicity factor.

2.3 Distributed pier-force implementation

The time-domain numerical simulation employs the classical staggered grid finite difference method (J. Virieux 1986; T. Bohlen 2002; S. Dong *et al.* 2021). Additionally, we apply the generalized standard linear solid in the time-domain attenuation modelling (H.P. Liu *et al.* 1976). We consider the subsurface structure as an isotropic viscoelastic medium. The wave propagation in volume V can be expressed as (A. Tarantola 1988; J.O. Robertsson *et al.* 1994; L. Groos *et al.* 2014)

$$\frac{\partial\sigma_{ij}}{\partial t}=\left[\pi\left(1+\tau^{p}\right)-2\mu\left(1+\tau^{s}\right)\right]\frac{\partial v_k}{\partial x_k}\delta_{ij}+\mu\left(1+\tau^{s}\right)\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right)+\sum_{l=1}^{L}r_{ij}^{l}, \tag{5}$$

$$\rho\frac{\partial v_i}{\partial t}=\frac{\partial\sigma_{ij}}{\partial x_j}+f_i, \tag{6}$$

$$\frac{\partial r_{ij}^{l}}{\partial t}=\frac{-1}{\tau_{\sigma l}}\left\{r_{ij}^{l}+\left(\pi\tau^{p}-2\mu\tau^{s}\right)\frac{\partial v_k}{\partial x_k}\delta_{ij}+\mu\tau^{s}\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right)\right\}, \tag{7}$$

where σ_{ij} is the ij-th component of stress tensor ($i,j=1,2$), v_i denotes the i-th component of the velocity and r_{ij}^{l} are the memory variables ($l=1,\ldots,L$) corresponding to the stress tensor σ_{ij}. τ^{p} and τ^{s} are the level of attenuation for P and S waves, respectively. ρ is the density. f_i represents the components of the external HST source.

The numerical modelling of the HST source differs from conventional fixed-source modelling in terms of its source-loading mechanism. Instead of applying the force at a single surface point, we represent the bridge pier as an embedded force-transmission structure. The HST-induced load is distributed along the embedded pier depth and attenuates with increasing depth. This implementation is physically consistent with the bridge–pier coupling concept of Z. Wang *et al.* (2020), but differs in that the source is injected on the staggered grid as a depth-distributed force along the embedded pier, rather than being represented by a finite number of discrete source points.

Based on the force transmission characteristics (Z. Wang *et al.* 2020), the magnitude of the horizontal force component $f_x(\mathbf{x},t)$ is set to half that of the vertical component $f_z(\mathbf{x},t)$. The external HST distributed pier-force vector $\mathbf{f}$ in eq. (6) can be expressed as

$$\mathbf{f}(\mathbf{x},t)=\sum_{p=1}^{N_p}\begin{bmatrix}0.5\\1\end{bmatrix}w(z)\,f_{\mathrm{hst}}\left(t-\frac{x_p}{v}\right)\delta(x-x_p)\chi_{[0,H]}(z), \tag{8}$$

with the attenuation factor $w(z)$

$$w(z)=\frac{1}{(1+\beta z)^{n_{\mathrm{att}}}}. \tag{9}$$

Here $\mathbf{x}=(x,z)$ is the position vector, N_p is the number of bridge piers, p is the pier index and $x_p=(p-1)D$ is the horizontal position of the p-th pier. The indicator function $\chi_{[0,H]}(z)$ equals 1 for $0\le z\le H$ and 0 otherwise, so the source is applied only along the embedded pier depth H. The factor $w(z)$ describes the depth-dependent attenuation of the distributed force along the pier, where β controls the attenuation rate with depth, and n_{att} controls the decay strength.

For numerical implementation on the staggered grid, the continuous HST sources are discretized into a series of fixed point sources with different time-delay (H. Wang & J. Chen 2023). The eq. (8) can be discretized as

$$\mathbf{f}_{i,j}^{k}=\sum_{p=1}^{N_p}\begin{bmatrix}0.5\\1\end{bmatrix}w(z_j)f_{\mathrm{hst}}\left(k\Delta t_{\mathrm{num}}-\frac{i_p\Delta x}{v}\right)\delta_{i,i_p}\chi_{[0,H]}(z_j), \tag{10}$$

where

$$z_j=(j-1)\Delta z,\qquad w(z_j)=\frac{1}{(1+\beta z_j)^{n_{\mathrm{att}}}}. \tag{11}$$

Here, Δx, Δz and Δt_{num} denote the horizontal grid interval, vertical grid interval and the time-step, respectively. i, j and k represent the discrete indices for the x-direction, z-direction and time discretization, respectively. The symbol i_p is the horizontal grid index of pier p, and δ_{i,i_p} is the Kronecker delta, which equals 1 for $i=i_p$ and 0 otherwise.

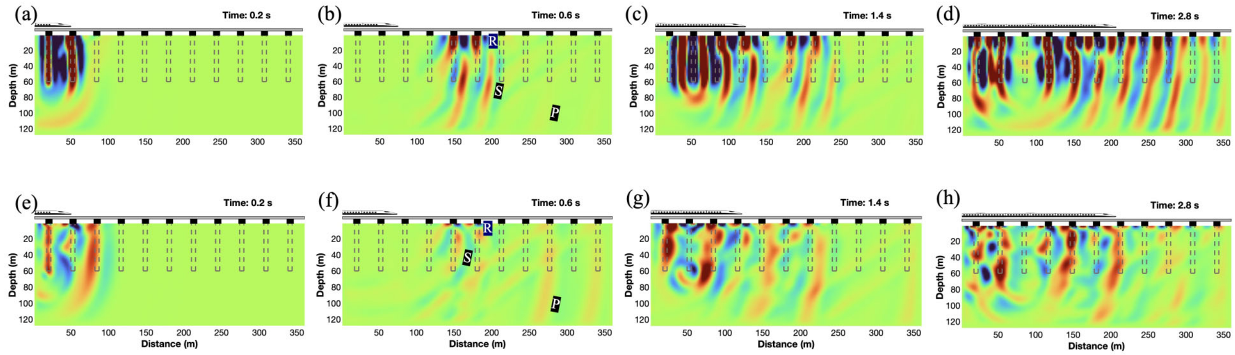

Figure 3. Velocity wavefield snapshots generated by HST distributed pier-force: (a–d) vertical component and (e–h) horizontal component at 0.2, 0.6, 1.4 and 2.8 s. Rectangulars in the subplot indicate the Rayleigh wave (R), *S* wave (S) and *P* wave (P).

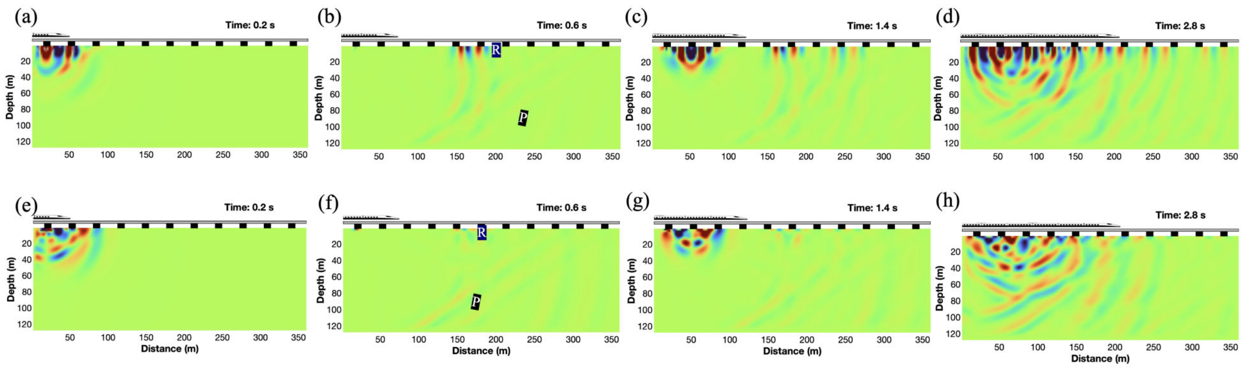

Figure 4. Velocity wavefield snapshots generated by HST explosive source: (a–d) vertical component and (e–h) horizontal component at 0.2, 0.6, 1.4 and 2.8 s. Rectangulars in the subplot indicate the Rayleigh wave (R) and *P* wave (P).

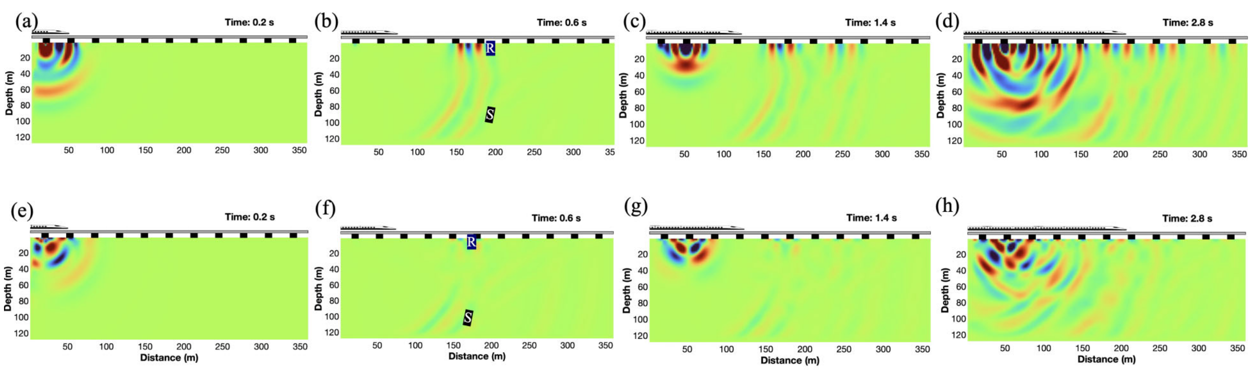

Figure 5. Velocity wavefield snapshots generated by HST single surface force: (a–d) vertical component and (e–h) horizontal component at 0.2, 0.6, 1.4 and 2.8 s. Rectangulars in the subplot indicate the Rayleigh wave (R) and *S* wave (S).

For the HST explosive source, the isotropic stress field excites *P* waves only,

$$\sigma_{xx} = \sigma_{zz} = \sum_{p=1}^{N_p} f_{\text{hst}}\left(t - \frac{x_p}{v}\right) \cdot \delta(\mathbf{x} - \mathbf{x}_p), \tag{12}$$

$\mathbf{x}_p$ is the explosive source location. However, for the explosive source located at the free surface ($\sigma_{zz} = 0$), the Rayleigh waves are generated by mode conversion when *P* waves interact with the free surface.

This fundamental difference in source geometry produces distinctly different radiation patterns and wavefield characteristics. To characterize the excitation mechanism of these HST sources, we visualize the velocity wavefield. Figs 3–5 present the vertical- and horizontal-component velocity wavefield snapshots at four time instances (0.2, 0.6, 1.4 and 2.8 s) for the three source implementations.

The distributed pier-force (Fig. 3), with its strong S- and Rayleigh-wave generation, provides optimal illumination for shear velocity structures throughout the shallow subsurface (0–60 m). The single surface force (Fig. 5), generates strong Rayleigh-wave and moderate S wave, but less efficient subsurface energy coupling compared to the distributed pier-force. The explosive source (Fig. 4), dominated by P-wave radiation, fundamentally misrepresents HST wavefield characteristics and proves unsuitable for HST-based inversion applications. These radiation patterns have important implications for FWI applications, as the parameter sensitivity and resolution capabilities depend critically on the types of waves illuminating the subsurface.

2.4 HST source wavelet correction

We implement the source estimation approach following L. Groos *et al.* (2014). We assume that the HST-induced excitations at different bridge piers have similar source characteristics (T. Rebert *et al.* 2024), reflecting the viaduct's repetitive structure and consistent train-track interactions.

The total modelled seismogram $u(x_r, t)$ is the superposition of the time-delayed contributions from all pier-source regions x_{s_p},

$$u(x_r, t) = \sum_{p=1}^{N_p} G_{m_0,p}(x_r, t; x_{s_p}) * f_{\mathrm{hst}}(t - \tau_p) = \sum_{p=1}^{N_p} \int_0^T G_{m_0,p}(x_r, t - t'; x_{s_p}) f_{\mathrm{hst}}(t' - \tau_p)\, \mathrm{d}t', \tag{13}$$

with the time delay of the p-th pier excitation given by

$$\tau_p = (p-1)\Delta\tau_p = (p-1)\frac{D}{v}, \tag{14}$$

where $u(x_r, t)$ denotes the modelled seismogram at receiver x_r and x_{s_p} denotes the effective source region of the p-th pier. N_p is the number of bridge piers and p is the pier index. $\Delta\tau_p$ is pier-to-pier delay and τ_p is the time delay of the p-th pier excitation. The source–time function $f_{\mathrm{hst}}(t)$ is the effective load history for one full HST passage over one pier, as defined in eq. (1).

The corresponding frequency-domain signal can be written as

$$u(x_r, \omega) = \sum_{p=1}^{N_p} G_{m_0,p}(x_r, x_{s_p}, \omega) \hat{f}_{\mathrm{hst},p}(\omega), \tag{15}$$

where

$$\hat{f}_{\mathrm{hst},p}(\omega) = \hat{f}_{\mathrm{hst}}(\omega) e^{-i\omega\tau_p}. \tag{16}$$

Here, $\hat{f}_{\mathrm{hst}}(\omega)$ is the Fourier transform of the effective source–time function $f_{\mathrm{hst}}(t)$ defined in eq. (1). $\hat{f}_{\mathrm{hst},p}(\omega)$ denotes the same source spectrum after applying the deterministic time delay τ_p for the p-th pier. The factor $e^{-i\omega\tau_p}$ represents the phase shift caused by the delayed excitation of each pier.

Similarly, the observed HST data are expressed as

$$d(x_r, \omega) = \sum_{p=1}^{N_p} G_{m_{\mathrm{true}},p}(x_r, x_{s_p}, \omega) \hat{f}_{\mathrm{hst,true},p}(\omega), \tag{17}$$

where

$$\hat{f}_{\mathrm{hst,true},p}(\omega) = \hat{f}_{\mathrm{hst,true}}(\omega) e^{-i\omega\tau_p}. \tag{18}$$

Here, $\hat{f}_{\mathrm{hst,true}}(\omega)$ denotes the unknown true effective HST source spectrum.

Our objective is to determine a linear frequency-domain filter $c(\omega)$, named the source–wavelet correction filter, that minimizes the discrepancy between the true and corrected effective HST source spectra,

$$\sum_{p=1}^{N_p} \left| \hat{f}_{\mathrm{hst,true},p}(\omega) - c(\omega)\hat{f}_{\mathrm{hst},p}(\omega) \right|^2 = \min. \tag{19}$$

We assume that the Green's functions G_p for the initial model m_0 are sufficiently close to those of the true model m_{true},

$$G_p := G_{m_0,p}(x_r, x_{s_p}, \omega) \approx G_{m_{\mathrm{true}},p}(x_r, x_{s_p}, \omega). \tag{20}$$

Using this approximation and summing over all delayed pier contributions, the source-correction problem can be written in the data domain as

$$\left| \sum_{p=1}^{N_p} \left[\tilde{d}_p(x_r, \omega) - c(\omega)\tilde{u}_p(x_r, \omega) \right] \right|^2 = |d(x_r, \omega) - c(\omega)u(x_r, \omega)|^2 = \min, \tag{21}$$

with the modelled contribution from the p-th pier given by

$$\tilde{u}_p(x_r, \omega) = G_{m_0,p}(x_r, x_{s_p}, \omega)\hat{f}_{\mathrm{hst},p}(\omega), \tag{22}$$

and the corresponding theoretical observed contribution given by

$$\tilde{d}_p(x_r, \omega) = G_{m_{\text{true}}, p}(x_r, x_{s_p}, \omega)\hat{f}_{\text{hst,true},p}(\omega). \tag{23}$$

Here, $u(x_r, \omega)$ and $d(x_r, \omega)$ denote the complete blended modelled and observed HST records, respectively. The individual terms $\tilde{d}_p$ are introduced only to describe the theoretical contribution of each delayed pier excitation. In the implementation, these individual contributions are not separated from the observed data. Instead, the correction filter is computed directly from the complete blended records $d(x_r, \omega)$ and $u(x_r, \omega)$. Therefore, the source correction is estimated from the full interfered HST wavefield rather than from independent pier gathers.

The Fourier coefficients $\tilde{c}_l$ of the correction filter are obtained by solving a damped least-squares problem,

$$F(\tilde{c}_l; \epsilon) = \sum_{l=0}^{N_f-1} \sum_{k=1}^{M} g(r_k) \left| \tilde{d}_{lk} - \tilde{c}_l \tilde{u}_{lk} \right|^2 + M\bar{E}\epsilon^2 \sum_{l=0}^{N_f-1} |\tilde{c}_l|^2, \tag{24}$$

where N_f and M are the number of Fourier coefficients and receivers, respectively. The variable r_k denotes the source–receiver offset of receiver k, and $g(r_k)$ is an offset-dependent weighting function. Following L. Groos *et al.* (2014), we define

$$g(r_k) = \left(\frac{r_k}{1\,\text{m}}\right)^{2\kappa}, \tag{25}$$

so that middle- and far-offset traces contribute more strongly to the objective function. In this study, since traces are amplitude-normalized prior to the filter estimation, we set $g(r_k) = 1$ unless stated otherwise. $\bar{E}$ is the average power of the Fourier coefficients $\tilde{u}_{lk}$ scaled by $g(r_k)$,

$$\bar{E} = \frac{1}{MN_f} \sum_{l=0}^{N_f-1} \sum_{k=1}^{M} g(r_k)|\tilde{u}_{lk}|^2. \tag{26}$$

The solution is obtained by setting the gradient of $F(\tilde{c}_l; \epsilon)$ with respect to $\tilde{c}_l$ to zero, yielding

$$\tilde{c}_l = \frac{\sum_{k=1}^{M} g(r_k)\tilde{d}_{lk}\tilde{u}_{lk}^*}{\sum_{k=1}^{M} g(r_k)|\tilde{u}_{lk}|^2 + M\bar{E}\epsilon^2}. \tag{27}$$

Here, $\tilde{d}_{lk}$, $\tilde{u}_{lk}$, and $\tilde{c}_l$ are the discrete Fourier coefficients of the observed data, modelled data, and correction filter, respectively. The modelled data $\tilde{u}_{lk}$ already contain the summed contribution of all time-delayed pier excitations. Thus, a single correction filter is estimated for the complete blended HST record and applied to the common effective HST source signature $f_{\text{hst}}(t)$. The corrected signature is then used for all piers with the deterministic delays τ_p.

In the following FWI workflow, the source correction filter is updated at the beginning of each frequency stage using the current model, and the corrected source signature is then fixed during the subsequent iterations of that stage.

2.5 FWI framework

In this section, we present the FWI workflow based on moving source seismic data. Seismic wave propagation, as described by the eqs (5)–(7), can be represented by the following matrix equation

$$\mathcal{A}_{ij}(\mathbf{x})\,\mathbf{u}(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t), \tag{28}$$

with the differential operator given by

$$\mathcal{A}_{ij}(\mathbf{x}) = \rho(\mathbf{x})\,\delta_{ij}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_k}\boldsymbol{\psi}_{ijkm}(\mathbf{x}, t) * \frac{\partial}{\partial x_m}, \tag{29}$$

where $\mathbf{u}$ is displacement vector, $\mathbf{f}$ is the moving source vector, $\mathbf{x} = (x, z)$ represents the position vector. The notation $*$ and δ_{ij} are convolution and Kronecker symbol, respectively. $\boldsymbol{\psi}$ represents the rate of relaxation, the detailed definition shown in Appendix A.

The inversion framework aims to minimize the misfit between the modelled data $\mathbf{u}$ and the observed data $\mathbf{d}$ (D. Köhn 2011; D. Köhn *et al.* 2012). We formulate the optimization problem using a typical normalized L_2 norm (Y. Choi & T. Alkhalifah 2012),

$$E(\mathbf{m}) = \frac{\sum_i^{nr}\sum_j^{nc}\left|\mathbf{u}_{i,j} - \mathbf{d}_{i,j}\right|^2}{\sum_i^{nr}\sum_j^{nc}\left|\mathbf{d}_{i,j}\right|^2}, \tag{30}$$

where $\mathbf{m}$ is the model parameter vector (e.g. density, velocities and attenuation factors). $\mathbf{u}$ and $\mathbf{d}$ are normalized to the maximum amplitude of each trace. The $\sum_i^{nr}$ is the sum over the receivers and the $\sum_j^{nc}$ is the sum over the components used in the inversion. A lower residual energy $E(\mathbf{m})$ indicates better agreement between the modelled and observed seismic data. A key advantage of our HST-based FWI is its computational efficiency: only one source term (representing the entire train's multiple time-delayed excitations) is required per iteration. In contrast, conventional fixed-source FWI requires summation over all the sources, as each source position

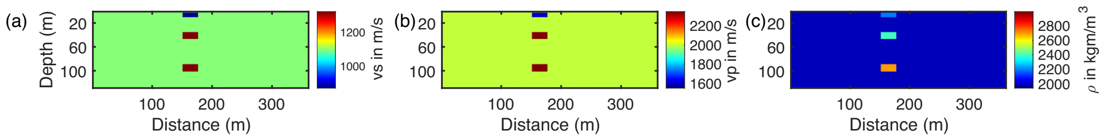

Figure 6. True anomaly models for: (a) S-wave velocity V_s, (b) P-wave velocity V_p and (c) density ρ.

contributes partial gradient information. By eliminating the source summation requirement, our HST source approach significantly reduces computational costs compared to fixed-source approach.

Although this no source-summation is conceptually related to encoded-source ideas, it differs from classical source-encoding FWI in an important way. In encoded FWI, multiple independent shots are artificially blended by random phase or time-shifts. So the gradient from blended data approximates the sum of single-shot gradients and cross-talk arises if encoding is insufficient or changes between iterations (J.R. Krebs *et al.* 2009; A. Guitton & E. Diaz 2012; P.P. Moghaddam *et al.* 2013; C. Castellanos *et al.* 2014). In our case, the HST excitation is a deterministic superposition of time-delayed pier excitations governed by known kinematics (train speed, pier spacing and carriage geometry). The cross-talk mechanism associated with approximate or random encoding does not apply to our formulation.

The model misfits $\delta\mathbf{M}$ between inverted model $\mathbf{m}_i$ and true model $\mathbf{m}_{\text{true}}$ are defined as

$$\delta\mathbf{M} = \sum_x \sum_z \frac{(\mathbf{m}_i(x,z) - \mathbf{m}_{\text{true}}(x,z))^2}{\mathbf{m}^2_{\text{true}}(x,z)}. \tag{31}$$

Here, the model parameters $\mathbf{m}$ includes S-wave velocity (V_s), P-wave velocity (V_p) and density (ρ) models.

The detailed gradient calculation of eq. (30) is provided in the Appendix B (L. Groos *et al.* 2014). The inversion method in this study employs the pre-conditioned conjugate gradient (PCG) optimization algorithm with appropriate regularization to maintain model smoothness and prevent overfitting. Our inversion incorporates attenuation effects to enhance subsurface characterization. For all the tests, we use the viscoelastic media with priori known Q-values. Both P- and S-wave quality factors are set to $Qs = Qp = 20$ with three relaxation mechanisms. Additionally, we impose a free surface boundary condition at the top of the models (R. Mittet 2002; T. Duretz *et al.* 2016), while the other three sides are treated with convolution perfectly matched layer (C-PML) boundary conditions to prevent artificial reflections (R. Martin & D. Komatitsch 2009).

3 SYNTHETIC DATA EXPERIMENTS

Having established the comprehensive theoretical framework, we now validate the methodology through synthetic experiments.

3.1 Anomaly model inversion

In this section, we conduct synthetic FWI in heterogeneous anomaly models (Fig. 6) to evaluate parameters recovery under different source implementations.

3.1.1 Inversion setup

The model dimensions are 180 × 64 grid points with uniform spacing $\Delta x = \Delta z = 2$ m. The initial models, we assume homogeneity, with S-wave velocity $V_s = 1100\ \text{m s}^{-1}$, P-wave velocity $V_p = 2000\ \text{m s}^{-1}$ and density $\rho = 2000\ \text{kg m}^{-3}$. The true models (Fig. 6) incorporate three distinct anomalies at depth range of (4, 10) m, (36, 56) m, (90, 100) m, respectively. We set the following per centage contrasts for the anomalies: 20 per cent for V_p and V_s, and 10 per cent, 20 per cent and 35 per cent for ρ. The sources and receivers are positioned along the bridge piers locations with 32 m spacing. The HST source characteristics follow the configuration in Fig. 2.

Fig. 7(a) presents HST seismic data simulated using the distributed pier-force. The distributed sources create time-delayed superposition effects, producing complex coupled waveforms. Fig. 7(b) displays the extracted third single-trace seismogram with its corresponding amplitude spectrum (Fig. 7c). The time-domain waveform exhibits highly periodic characteristics with clear amplitude modulation, reflecting consistent energy radiation from the moving train. The amplitude spectrum reveals broad-band frequency content with well-defined harmonic peaks at regular intervals, directly reflecting the train's periodic carriage structure. The spectrum contains usable low-frequency components, although their amplitudes are weaker than those of the dominant higher frequency harmonics. These low-frequency components provide the lowest frequency starting information for the multiscale inversion.

Based on the spectrum analysis, a frequency bandwidth of 1–80 Hz with 5 Hz increments is selected for the multiscale inversion. A minimum of five iterations per frequency is enforced to achieve adequate model updates before progressing to higher frequencies. To stabilize the inversion, smoothness is imposed through two procedures. The first is a spatial filtering of the adjoint-state gradients

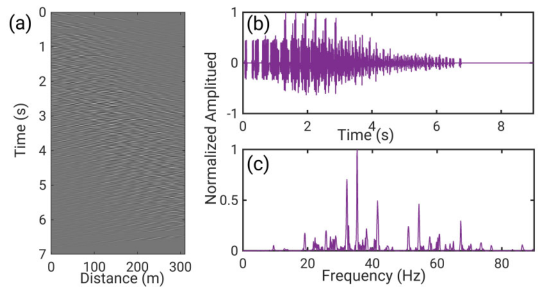

Figure 7. HST seismic data simulated using distributed pier-force: (a) Trace-normalized vertical-component synthetic seismograms, (b) extracted third single-trace seismogram and (c) corresponding amplitude spectrum.

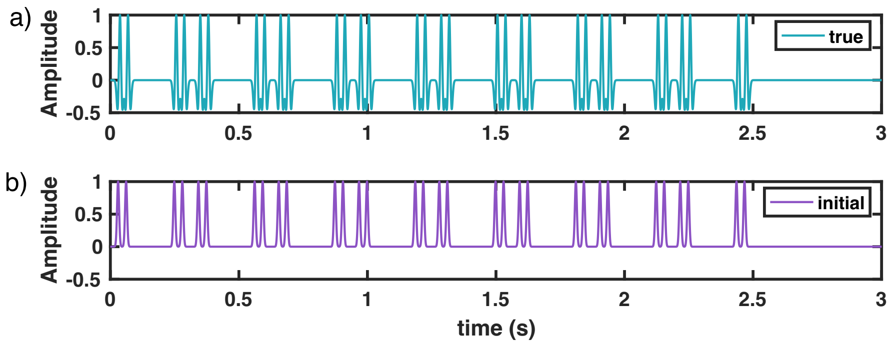

Figure 8. Source wavelets used in the synthetic STFI test: (a) true Ricker wavelet used to generate the data, and (b) mismatched initial Gaussian wavelet used to start the inversion.

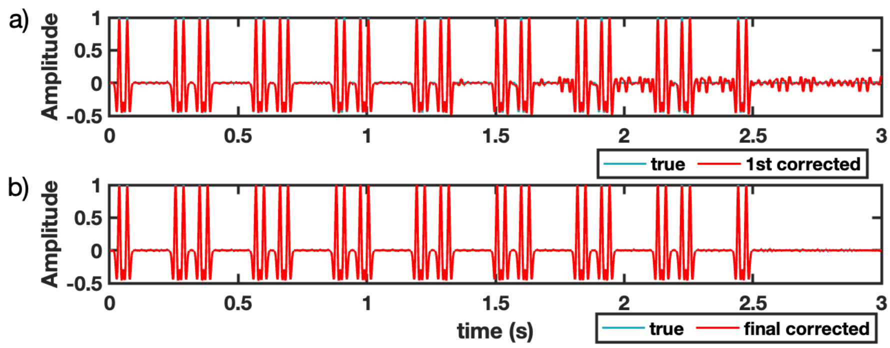

Figure 9. Estimated HST pier-source wavelet after source–wavelet correction: (a) corrected wavelet at the first frequency stage, and (b) corrected wavelet at the final frequency stage.

applied after each iteration to suppress grid-scale oscillations and stabilize the multiparameter updates. In our implementation, this is a 1-D Gaussian smoothing in the horizontal direction. The second is a mild model-smoothing applied after each model update.

3.1.2 Synthetic test with unknown source wavelet

Before comparing the performance of the different sources, we first evaluate the robustness of the inversion when the source wavelet is unknown. We performed an additional synthetic experiment in which the observed data were generated using a true Ricker wavelet (Fig. 8a), whereas the inversion was initialized with an intentionally mismatched Gaussian wavelet (Fig. 8b). During the inversion, we applied the frequency-domain source correction filter described in Section 2.4 and updated the HST pier-source signature at the beginning of each frequency stage.

Fig. 9 shows the corrected wavelet estimated at the first and final frequency stages, and Fig. 10 compares the corresponding amplitude spectra of the true, initial and final corrected sources. These results demonstrate that the corrected wavelet converges toward the true signature in both the time and frequency domains. Finally, Fig. 11 presents the inversion results obtained using the source–

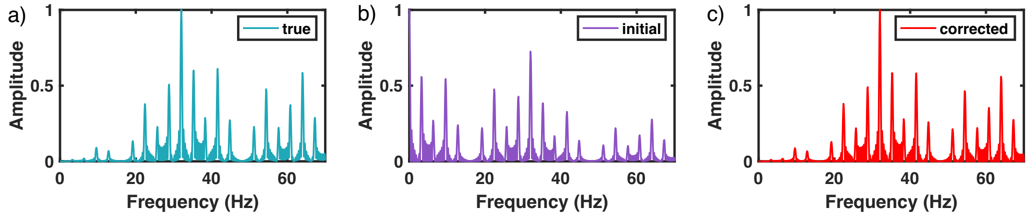

Figure 10. Corresponding amplitude spectra of the HST source: (a) true source, (b) initial (mismatched) source and (c) corrected source.

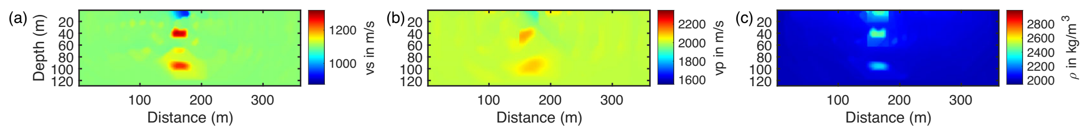

Figure 11. Inversion results for the synthetic test using the source–wavelet correction: (a) *S*-wave velocity V_s, (b) *P*-wave velocity V_p and (c) density ρ.

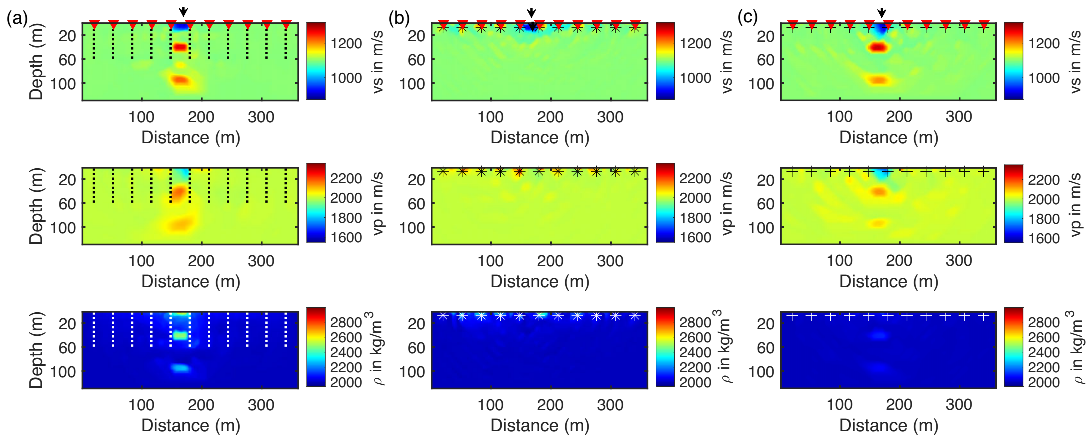

Figure 12. Inversion results using (a) HST distributed pier-force, (b) HST explosive source and (c) HST single surface force. Rows show *S*-wave velocity (V_s), *P*-wave velocity (V_p) and density (ρ) from top to bottom. Dotted lines mark the bridge piers extending from the surface down to the pier depth (60 m). Stars and crosses mark the piers (sources) positions. Inverted triangles mark receiver positions at 32 m intervals. The arrow at $x = 170$ m marks the profile location for Fig. 13.

wavelet correction, indicating that the proposed strategy yields stable and accurate multiparameter recovery despite the initially incorrect wavelet. This test confirms that the source–wavelet correction implemented in this study can compensate for an unknown or mismatched wavelet and is therefore suitable for the field-data application.

3.1.3 Comparison of source implementations

The following tests compare FWI results from three sources implementations: (a) distributed pier-force (this study), (b) explosive sources excited at the surface and (c) single surface forces excited at the surface (J. Luo *et al.* 2022). After reaching the maximum frequency, we obtained the final inversion results displayed in Fig. 12. The distributed pier-force approach (Fig. 12a) successfully reconstructs both shallow and deep anomalies across all three parameters, demonstrating the effectiveness of this source implementation. In contrast, the explosive source results (Fig. 12b) fail to recover subsurface structure for any parameter, confirming that isotropic radiation patterns are fundamentally inappropriate for modelling HST-induced vibrations. The single surface force implementation (Fig. 12c) exhibits mixed performance: while it successfully reconstructs both V_s and V_p anomalies, density recovery is notably degraded compared to the distributed pier-force approach.

Fig. 13 presents vertical velocity profiles at $x = 170$ m, comparing true, initial and inverted models from the three sources. For both V_s (Fig. 13a) and V_p (Fig. 13b), distributed pier-force and single surface force achieve comparable high recovery, whereas the explosive sources only capture the shallow V_s anomaly. The key difference appears in the density reconstruction (Fig. 13c): the distributed source recovers the anomalies, whereas the others remain nearly flat.

The convergence results (Fig. 14) show that the distributed pier-force achieves steady convergence with a 25 per cent reduction, whereas the explosive and single-force sources exhibit pronounced stagnation, with residuals improving by less than 1 per cent.

Model misfit evolution (Fig. 15) reveals that the distributed pier-force shows the most reliable convergence with strong and consis-

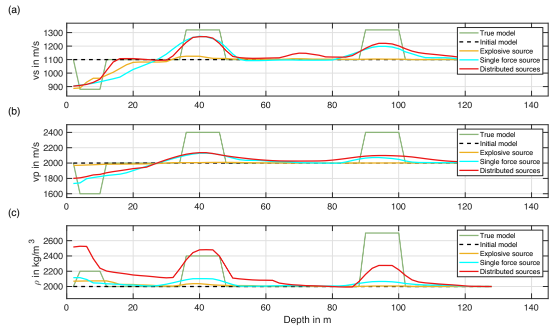

Figure 13. Vertical velocity profiles at x= 170 m for (a) S-wave velocity, (b) P-wave velocity and (c) density, comparing true model, initial model, distributed pier-force, explosive source and single force.

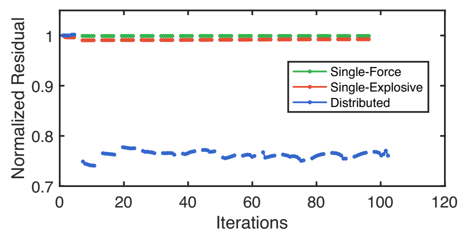

Figure 14. Residual convergence curves for the HST distributed pier-force, HST explosive source and HST single force.

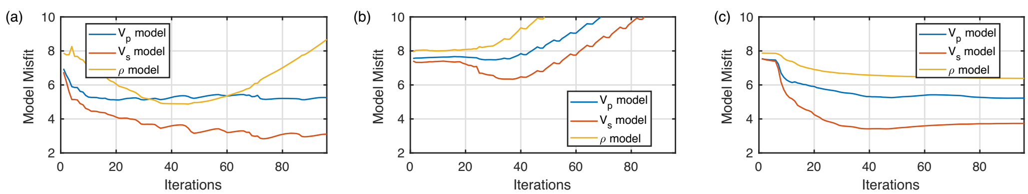

Figure 15. Model misfit curves for (a) the HST distributed pier-force, (b) HST explosive sources and (c) HST single force. Each figure shows S-wave velocity (V_s), P-wave velocity (V_p) and density (ρ) misfit.

tent misfit reduction, especially for V_s. Single force exhibits moderate improvement but remain more challenging for density inversion. Explosive sources show minimal misfit reduction for all parameters, confirming their inability to adequately fit HST seismic data.

3.2 Sparse acquisition geometry test

To evaluate the practical applicability of distributed pier-force under realistic field conditions, we conducted a comprehensive assessment using a sparse acquisition that mirrors the limitations encountered in our field surveys (see the next section for detailed acquisition parameters). The sparse acquisition consists only three receivers. To simulate realistic source conditions, we utilize 11-pier positions with a regular spacing of 32 m. The inversion parameters remain identical to the last synthetic test. Following the geological framework established by H. Wang (2025), we construct a true linear velocity model embedded with anomalies (Figs 16a, c, e) to evaluate the method's sensitivity to realistic subsurface heterogeneity under receiver-limited conditions. The initial model is purely linear, lacking anomalies. The model dimensions are 340 $*$ 120 grid points with uniform spacing of $\Delta x = \Delta z = 1$ m. Three receivers are deployed at offsets of 166, 206 and 238 m, respectively (marked by inverted triangles in subplots a and b of Fig. 16).

After 200 iterations, we obtain the inversion results (Figs 16b, d and f). The S-wave velocity inversion successfully captures the primary anomalies features, particularly in the shallow subsurface. This indicate that Rayleigh wave maintain sufficient sensitivity even with limited receiver coverage, particularly in the shallow subsurface. The P-wave velocity inversion reveals limitations under sparse acquisition conditions, with notable inversion artefacts distributed throughout the model domain. The density parameter shows the highest sensitivity to sparse acquisition geometry, exhibiting the most severe degradation in inversion quality. This poor performance reflects the inherently weak sensitivity of seismic waves to density variations compared to velocity parameters. Nevertheless, certain anomalous signatures remain discernible, indicating that density inversion may still yield valuable geological information despite acquisition constraints. These sparse acquisition inversion results reveal distinct parameter-dependent performance characteristics and provide valuable insights into the following practical application.

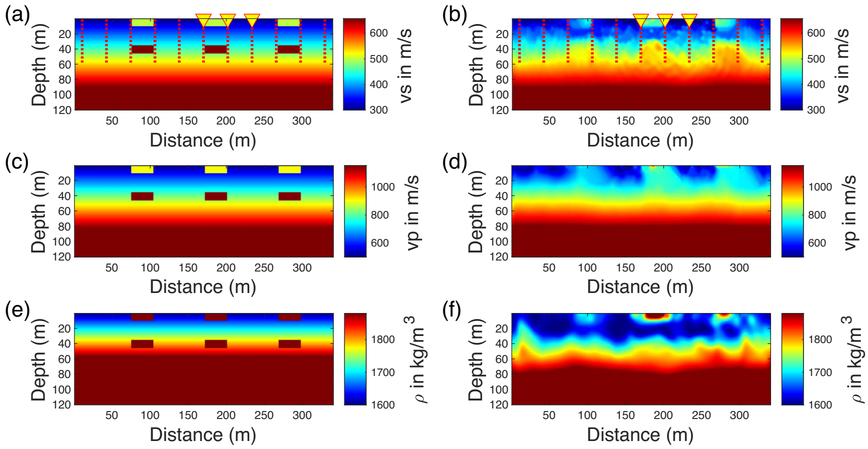

Figure 16. True models showing (a) *S*-wave velocity V_s, (c) *P*-wave velocity V_p and (e) density ρ, with corresponding inversion results for (b) V_s, (d) V_p and (f) ρ. Dotted lines mark the bridge piers extending from the surface down to the pier depth (60 m). Three seismic receivers were deployed at offsets of 166, 206 and 238 m (indicated by inverted triangles in subplots a and b).

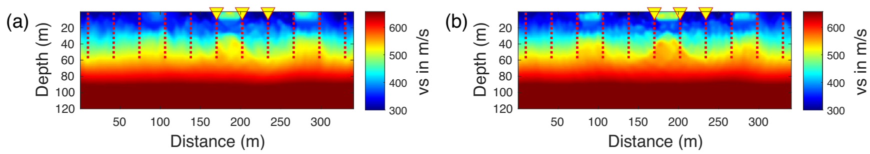

Figure 17. Inverted *S*-wave velocity V_s obtained from (a) multiparameter FWI and (b) monoparametric (V_s-only) FWI. Dotted lines mark the bridge piers extending from the surface down to the pier depth (60 m). Three seismic receivers were deployed at offsets of 166, 206 and 238 m (indicated by inverted triangles).

To further assess whether limited scattering angles in the sparse acquisition geometry lead to multiparameter trade-off, we performed an additional monoparametric inversion. The monoparametric test only updates the V_s model, while keeps V_p and ρ fixed to their true models in Fig. 16(a). For a fair comparison, the multiparameter and monoparametric inversions use the same acquisition geometry and the same frequency schedule up to 26 Hz. Fig. 17 shows the comparison of the recovered V_s models. For the anomaly located farther from the receivers, the V_s-only inversion produces a cleaner reconstruction. In contrast, the multiparameter inversion is more affected by parameter trade-off under the sparse acquisition geometry, since the limited scattering-angle illumination is insufficient to reliably recover V_p and ρ. Overall, this comparison indicates that focusing the inversion on V_s can reduce cross-talk and improve the interpretability of the recovered V_s structure.

4 FIELD DATA APPLICATION

Having rigorously validated the distributed source through synthetic anomaly comparison and sparse acquisition scenarios, we now test its application to real data. This section presents the reconstruction of subsurface structures using field HST signals, providing a field-data assessment of the proposed source model.

4.1 Data acquisition and processing

The field data were collected in Rongcheng County, Hebei Province, China. The survey deployed 12 seismic stations (R1–R12) under and around the viaduct to record HST-induced ground vibrations (H. Wang & J. Chen 2023) (Fig. 18). Because our inversion is performed in 2-D, we selected three representative stations (R1, R2, R3) that are approximately aligned with the railway direction (Fig. 19). The remaining stations are distributed in a 3-D configuration and are not used in the following 2-D inversion.

Fig. 20 presents the recorded three-component seismograms from the 8-carriage train. The data were transformed from geographical coordinates (latitude–longitude) to railway-aligned coordinates, with the *x*-axis oriented along the railway line. This transformation ensures optimal alignment between source positions and receiver coordinates for subsequent inversion processing. The frequency spectra reveal significant energy at frequencies as low as 0.1 Hz, providing crucial long-wavelength information for mitigating cycle-skipping issues. Additionally, the frequency spectra extend from 0.1 Hz to approximately 80 Hz, enabling multiscale inversion approaches.

Regarding the 3-D–2-D spreading correction, our objective function normalizes each trace by its maximum amplitude (eq. 30). Therefore, a standard 3-D–2-D spreading correction that primarily acts as an offset-dependent amplitude scaling is expected to have

Figure 18. Field HST acquisition geometry showing all twelve seismic stations (R1–R12). Stations R1–R3 (red markers) are the receivers selected for the 2-D FWI because they are approximately aligned with the railway, satisfying the 2-D profile assumption. The remaining stations (R4–R12, blue markers) are distributed in a 3-D configuration around the track and are shown here to clarify the overall deployment. The location of R1–R3 is zoomed in Fig. 19.

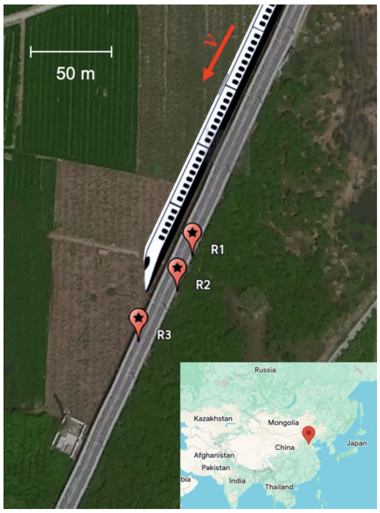

Figure 19. Zoomed view of the selected 2-D receiver line. Three stations (R1–R3) used in the inversion are shown together with an arrow indicating the train running direction along the railway.

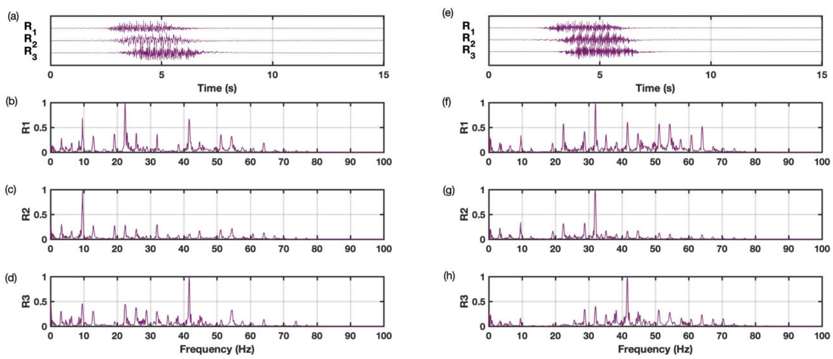

Figure 20. The eight-carriage HST velocity responses of the three stations. Left column (a–d): horizontal-response and its corresponding frequency spectra. Right column (e–h): vertical-response and its corresponding frequency spectra.

only a limited impact after normalization. We verified if this correction has an effect on the FWI result by applying a standard 3-D–2-D amplitude correction to the field data and repeating the inversion. Both the misfit evolution and the inverted models remain essentially unchanged. Consequently, no explicit 3-D–2-D correction is applied in the main workflow.

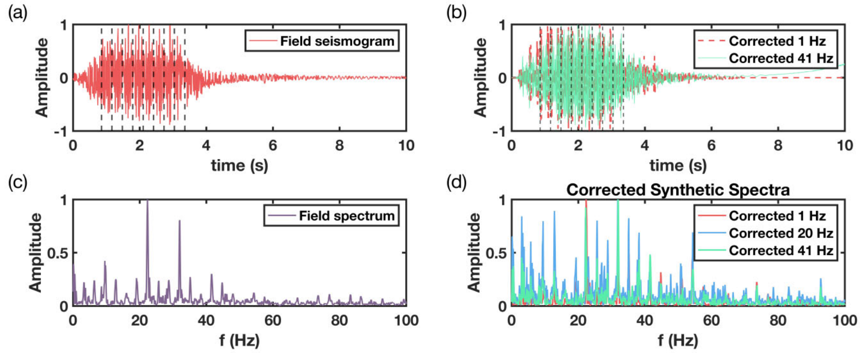

Figure 21. Comparison of field and corrected synthetic data in time and frequency domains. Left column: (a) Field data seismogram and (c) its amplitude spectrum. Right column: (b) Corrected synthetic seismograms and (d) the corresponding amplitude spectra at different correction stages (1, 20 and 41 Hz) during the inversion process.

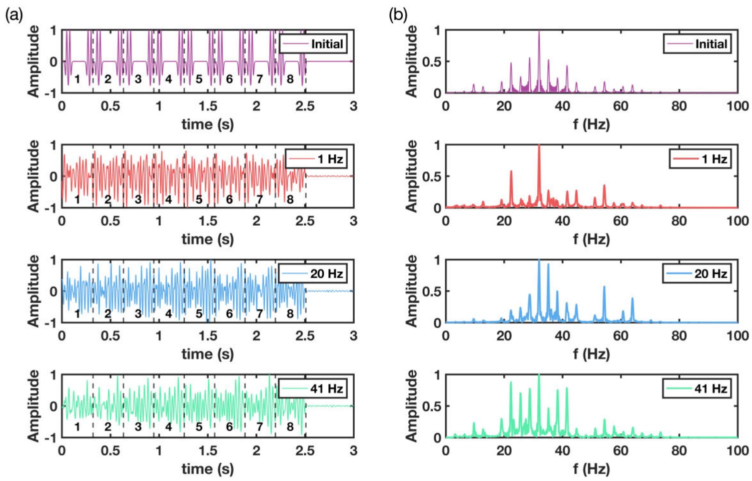

Figure 22. Comparison of the initial and corrected source wavelets (a) and their corresponding amplitude spectra (b) at different correction stages (1, 20 and 41 Hz) during the inversion process. The number 1–8 represents the individual carriages of the 8-carriage HST.

4.2 Real HST-source estimation

The synthetic tests in Section 3 provide a controlled comparison of the three source implementations. In the field data, only the actual HST-induced wavefield is recorded. Therefore, the individual contributions of different idealized source types cannot be isolated. We use the field data as a plausibility test of the proposed distributed pier-force model. Specifically, we evaluate whether this model, combined with source correction, can reproduce the main waveform and spectral characteristics of real HST observations.

A critical challenge in HST field data FWI is the unknown nature of the source wavelet generated by the HST. Unlike conventional seismic surveys, HST-induced seismic signals result from complex interactions between train dynamics, track structure and ground coupling mechanisms.

Fig. 21 presents a comparison between the field records and the corrected synthetic seismograms in both the time and frequency domains. The field data exhibit complex waveforms with clear amplitude variations and broad-band spectral content. After applying the source-wavelet correction during the multiscale inversion, the synthetic data reproduce the main temporal characteristics and spectral distribution of the observed HST records. This agreement indicates that the distributed pier-force model, combined with the corrected effective source signature, can capture the dominant features of the real HST-induced wavefield.

Fig. 22 tracks source wavelet evolution during multiscale inversion. Fig. 22(a) displays the time-domain source wavelets, showing the initial wavelet (top) and its correction at 1, 20 and 41 Hz. Fig. 22(b) displays their amplitude spectra. The comparison reveals how the source wavelet undergoes systematic refinement during the inversion. The initial wavelet exhibits a simple, low-frequency character, while successive corrections at 1, 20 and 41 Hz progressively enhance the wavelet's high-frequency content and temporal resolution.

In addition, the corrected source spectra in Fig. 22(b) reveal carriage-specific variations from differing axle loads, suspensions and track coupling. The corrected wavelets in Fig. 22(a) show impulse-like characteristics with shorter durations and higher peaks, which

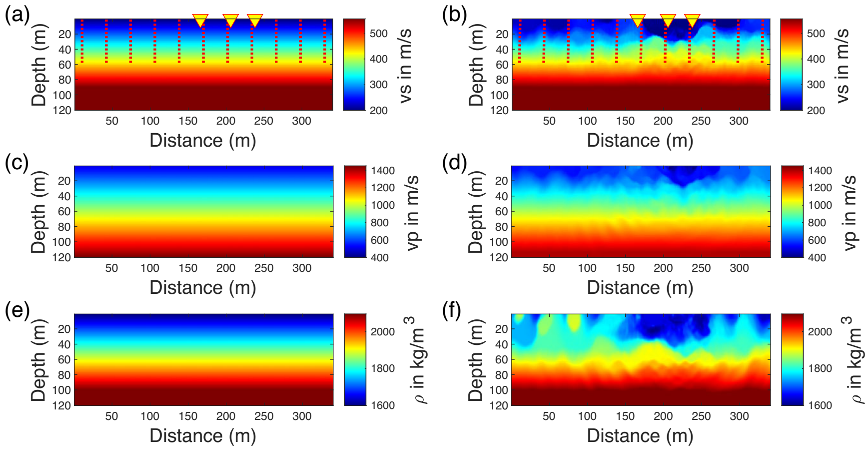

Figure 23. The initial models of *S*-wave velocity V_s (a) *P*-wave velocity V_p (c) and density ρ (e). The inverted models *S*-wave velocity V_s (b), *P*-wave velocity V_p (d) and density ρ (f). Dotted lines mark the bridge piers extending from the surface down to the pier depth (60 m). The three receivers are deployed at offsets of 166, 206 and 238 m, respectively (marked by inverted triangles in subplots a and b).

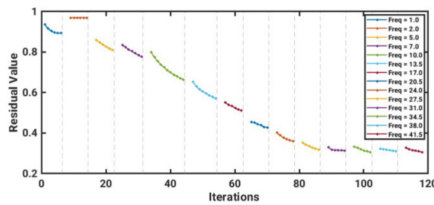

Figure 24. Residual convergence curves with iteration number across multiple frequency bands (1.0–41.5 Hz).

is consistent with wheel-rail contact forces. This behaviour provides field-data support for representing the HST excitation as repeated pier-coupled forces rather than as an isolated surface source.

4.3 Inversion results

The field data inversion employs a multiscale approach, progressively incorporating higher frequencies to resolve subsurface structures at different scales. Beginning with frequency 1 Hz to establish large-scale velocity frameworks, the inversion systematically increased frequency content to resolve finer structural details. The inversion utilized the same viscoelastic formulation and optimization framework validated in the numerical experiments.

The initial model (Figs 23a, c, e) was constructed following the starting model shown in figs 5–20 of H. Wang (2025). Specifically, we constructed the initial model from smooth, laterally homogeneous 1-D background models for V_p, V_s and ρ, with all parameters increasing linearly with depth. During the inversion, the parameters are constrained within the bounds: $V_p \in [400, 1800]$ m s^{-1}, $V_s \in [150, 800]$ m s^{-1} and $\rho \in [1600, 2200]$ kg m^{-3}.

The final inversion results (Figs 23b, d, f) successfully revealed distinct subsurface layering patterns consistent with the previous geophysical investigations (Y. Ma *et al.* 2020; J. Shao *et al.* 2022; G. Wang *et al.* 2022). The V_s and density models exhibit spatially coherent structural patterns with well-defined layer boundaries. The convergence curve (Fig. 24) demonstrates robust optimization performance across all frequency bands, with residual reductions exceeding 70 per cent.

5 DISCUSSION

A new eco-friendly seismic source proposed in this study models the HST as a distributed pier-force, which fundamentally changes how the HST wavefield illuminates the near surface and, consequently, how well elastic FWI can recover multiparameter structure. We treat the deeply embedded piers as extended force chains, and implement it within an elastic FWI framework. Compared with explosive and single-force implementations, the distributed source produces a richer mix of *S*- and Rayleigh-wave (Fig. 3), improving the sensitivity to both velocity and density anomalies. This behaviour reflects our distributed pier-force implementation better represents the mechanical coupling of the train–bridge–ground system. As a result, our method offers a more reliable and physically consistent source model for real HST data acquired along a viaduct.

The anomaly tests highlight an important practical implication that with the distributed pier-force, density becomes a resolvable parameter in addition to Vs and Vp (Figs 6–15). In our case, this improved density recovery results from stronger S- and Rayleigh-wave excitation along the pier depth. The sparse acquisition experiment bridges the gap between idealized numerical tests and realistic field constraints. With only three receivers, the inversion still reconstructs the primary Vs anomalies (Fig. 16), whereas Vp and especially ρ degrade more strongly. This parameter-dependent behaviour indicates what we can expect from HST-based FWI in typical acquisition: robust *S*-wave velocity recovery, but limited resolution for *P*-wave velocity and density unless receiver coverage is improved. Rather than being a weakness, this outcome serves as a guide for field survey design.

A second key aspect of this study is testing the realism of the proposed source concept with field data. The joint source–model inversion drives the estimated wavelets toward short, high-amplitude, carriage-specific pulses consistent with wheel–rail contact forces (Figs 2 and 22), while the final V_s and ρ models show layered structures compatible with independent geophysical information from the area. However, the current data set is limited by sparse acquisition geometry and the absence of borehole data validation. Future work could prioritize collecting HST field data with both improved receiver density and correlated borehole measurements to validate inversion accuracy.

Despite these limitations, the results suggest that, with further development, including 3-D extensions, multiple-train inversion and hybrid strategies that combine FWI with seismic interferometry and other geophysical, geological and engineering data, our approach could become a valuable tool for urban geophysical exploration.

6 CONCLUSION

This study establishes HST-based elastic FWI as an effective approach for high-resolution subsurface characterization through a novel distributed pier-force implementation.

We developed a physically realistic distributed pier-force model that applies time-delayed excitations continuously along embedded bridge piers, fundamentally differing from the conventional single-source. Wavefield analysis demonstrates that this mechanism preferentially generates strong *S*- and Rayleigh-wave energy, providing superior shallow subsurface illumination.

Our comparative synthetic tests demonstrate that distributed pier-force successfully reconstruct all subsurface parameters, while explosive sources fail completely and single surface forces show limited density sensitivity. Sparse acquisition tests further confirm the practical feasibility of the distributed pier-force under field conditions.

Field data application demonstrates the practical feasibility of the proposed workflow through successful source-wavelet estimation and coherent subsurface reconstruction. The inverted source signatures evolve toward physically realistic, high-amplitude impulses matching wheel-rail contact theory, supporting the physical plausibility of the distributed model. Final inversions reveal distinct, consistent subsurface layering with a 70 per cent residual reduction, validating the method's field feasibility.

In conclusion, the HST distributed pier-force provides an effective and innovative source for subsurface exploration.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Joyrles Fernandes de Moraes for his technical assistance and the Applied Geophysics at the Karlsruhe Institute of Technology for the software support.

DATA AVAILABILITY

Due to confidentiality constraints, the raw HST field data cannot be made fully open, but selected processed data are available upon reasonable request. This study uses the IFOS2D software package for elastic full waveform inversion (IFOS2D, 2023; Köhn, 2011).

REFERENCES

Bohlen, T., 2002. Parallel 3-D viscoelastic finite difference seismic modelling, *Comput. Geosci.,* **28**(8), 887–899.

Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.,* **185**(2), 845–870..

Cao, J. & Chen, J., 2019. Solution of green function from a moving line source and the radiation energy analysis: a simplified modeling of seismic signal induced by high-speed train., *Chin. J. Geophys. (in Chinese),* **62**(6), 2303–2312..

Castellanos, C., Métivier, L., Operto, S., Brossier, R. & Virieux, J., 2014. Fast full waveform inversion with source encoding and second-order optimization methods, *Geophys. J. Int.,* **200**(2), 720–744..

Choi, Y. & Alkhalifah, T., 2012. Application of multi-source waveform inversion to marine streamer data using the global correlation norm, *Geophys. Prospect.,* **60**.

Connolly, D.P., Kouroussis, G., Laghrouche, O., Ho, C.L. & Forde, M.C., 2015. Benchmarking railway vibrations–track, vehicle, ground and building effects, *Construct. Build. Mater.,* **92,** 64–81.

Dong, S., Chen, J.B. & Li, Z., 2021. Viscoelastic wave finite-difference modeling in the presence of topography with adaptive free-surface boundary condition, *Acta Geophys.,* **69,** 2205–2217.

Duretz, T., May, D.A. & Yamato, P., 2016. A free surface capturing discretization for the staggered grid finite difference scheme, *Geophys. J. Int.,* **204**(3), 1518–1530..

Forrest, J. & Hunt, H., 2006. Ground vibration generated by trains in underground tunnels, *J. Sound Vibration,* **294**(4-5), 706–736..

Groos, J.C. & Ritter, J.R.R., 2009. Time domain classification and quantification of seismic noise in an urban environment, *Geophys. J. Int.,* **179**(2), 1213–1231..

Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T., 2014. The role of attenuation in 2d full-waveform inversion of shallow-seismic body and rayleigh waves, *Geophysics,* **79,** R247–R261..

Guitton, A. & Diaz, E., 2012. Attenuating crosstalk noise with simultaneous source full waveform inversion, *Geophys. Prospect.,* **60**(4), 759–768..

Hu, G., Zhang, G., Sun, S., Li, Y., Yang, J., Duan, J., He, H. & Zhan, Y., 2019. Fwi of high-speed-train seismic first arrival signal., *Chin. J. Geophys. (in Chinese),* **62**(12), 4817–4824..

Hung, H.H. & Yang, Y.B., 2001. Elastic waves in visco-elastic half-space generated by various vehicle loads, *Soil Dyn. Earthq. Eng.,* **21,** 1–17.

IFOS2D, 2023. 2–D full waveform inversion code, https://gitlab.kit.edu/kit/gpi/ag/software/ifos2d/.

Inbal, A., Cristea-Platon, T., Ampuero, J.P., Hillers, G., Agnew, D.C. & Hough, S.E., 2018. Sources of long-range anthropogenic noise in southern california and implications for tectonic tremor detection, *Bull. seism. Soc. Am.,* **108,** 3511–3527.

Kaynia, A.M., Madshus, C. & Zackrisson, P., 2000. Ground vibration from high-speed trains: prediction and countermeasure, *J. Geotechn. Geoenviron. Eng.,* **126**(6), 531–537..

Köhn, D., 2011. Time Domain 2D Elastic Full Waveform Tomography, Ph.D. thesis, Christian-Albrechts-Universität zu Kiel.

Köhn, D., De Nil, D., Kurzmann, A., Swiatek, A. & Bohlen, T., 2012. On the influence of model parametrization in elastic full waveform tomography, *Geophys. J. Int.,* **191,** 325–345..

Kouroussis, G., Connolly, D. & Verlinden, O., 2014. Railway induced ground vibrations—a review of vehicle effects, *Int. J. Rail Transportation,* **2,** 69–110..

Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. & Lacasse, M.D., 2009. Fast full-wavefield seismic inversion using encoded sources, *Geophysics,* **74**(6), WCC177–WCC188..

Lavoué, F., Coutant, O., Boué, P., Pinzon-Rincon, L., Brossier, R., Dales, P., Rezaeifar, M. & Bean, C.J., 2020. Understanding seismic waves generated by train traffic via modeling: implications for seismic imaging and monitoring, *Seismol. Res. Lett.,* **92**(1), 287–300.

Liu, H.P., Anderson, D.L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. Int.,* **47,** 41–58.

Liu, Y., Yue, Y., Luo, Y. & Li, Y., 2021a. Effects of high-speed train traffic characteristics on seismic interferometry, *Geophys. J. Int.,* **227**(1), 16–32..

Liu, Y., Yue, Y., Luo, Y. & Li, Y., 2021b. The seismic broad-band signature of high-speed trains running on viaducts, *Geophys. J. Int.,* **226,** 884–892..

Liu, Y., Yue, Y., Li, Y. & Luo, Y., 2022. On the retrievability of seismic waves from high-speed-train-induced vibrations using seismic interferometry, *IEEE Geosci. Remote Sens. Lett.,* **19,** 1–5..

Luo, J., Wang, X. & Chen, W., 2022. Subsurface elastic parameter reconstruction based on seismic data from the high-speed trains using full waveform inversion, *IEEE Trans. Geosci. Remote Sens.,* **60,** 1–8..

Luo, J., Wang, X. & Chen, W., 2023. Imaging the subsurface with the high-speed train seismic data-based elastic reverse time migration, *IEEE Trans. Geosci. Remote Sens.,* **61,** 1–9.

Ma, Y., Li, H., Zhang, J., Sun, S., Xia, Y., Feng, J., Long, H. & Zhang, J., 2020. Geophysical technology for underground space exploration in xiongan new area, *Acta Geosci. Sin.,* (4), 535–542..

Martin, R. & Komatitsch, D., 2009. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation, *Geophys. J. Int.,* **179,** 333–344.

Mittet, R., 2002. Free-surface boundary conditions for elastic staggered-grid modeling schemes, *Geophysics,* **67,** 1616–1623.

Moghaddam, P.P., Keers, H., Herrmann, F.J. & Mulder, W.A., 2013. A new optimization approach for source-encoding full-waveform inversion, *Geophysics,* **78**(3), R125–R132..

Quiros, D.A., Brown, L.D. & Kim, D., 2016. Seismic interferometry of railroad induced ground motions: body and surface wave imaging, *Geophys. J. Int.,* **205**(1), 301–313..

Rebert, T., Bardainne, T., Allemand, T., Cai, C. & Chauris, H., 2024. Characterization of train kinematics and source wavelets from near-field seismic data, *Geophys. J. Int.,* **237**(2), 697–715..

Rezaeifar, M., Lavoué, F., Maggio, G., Xu, Y., Bean, C.J., Pinzon-Rincon, L., Lebedev, S. & Brenguier, F., 2022. Imaging shallow structures using interferometry of seismic body waves generated by train traffic, *Geophys. J. Int.,* **233**(2), 964–977..

Robertsson, J.O., Blanch, J.O. & Symes, W.W., 1994. Viscoelastic finite-difference modeling, *Geophysics,* **59**(9), 1444–1456.

Sager, K., Tsai, V.C., Sheng, Y., Brenguier, F., Boué, P., Mordret, A. & Igel, H., 2021. Modelling P waves in seismic noise correlations: advancing fault monitoring using train traffic sources, *Geophys. J. Int.,* **228**(3), 1556–1567..

Shao, J., Wang, Y. & Chen, L.F., 2022. Near-surface characterization using high-speed train seismic data recorded by a distributed acoustic sensing array, *IEEE Trans. Geosci. Remote Sens.,* **60,** 1–11.

Sheng, Y., 2023. Seismic stereometry: an alternative two-station algorithm to seismic interferometry for analysing car-generated seismic signals, *Geophys. J. Int.,* **235**(1), 853–861..

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms including elasticity and attenuation, *Pure appl. Geophys.,* **128**(1-2), 365–399..

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160**(1), 195–216..

Virieux, J., 1986. P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method, *Geophysics,* **51**(4), 889–901..

Wang, G., Wang, H., Li, H., Lu, Z., Li, W. & Xu, T., 2022. Application of active-source surface waves in urban underground space detection: a case study of rongcheng county, hebei, china, *Earth Planet. Phys.,* **6**(4), 385–398..

Wang, H., 2025. Research of forward modeling and inversion for moving sources in the frequency domain and parallel computing, PhD thesis, University of Chinese Academy of Sciences.

Wang, H. & Chen, J., 2023. Wavefield modeling and characteristic analysis for a uniformly moving high-speed-train source under a two-layered model with applications, *IEEE Trans. Geosci. Remote Sens.,* **61,** 1–14..

Wang, L., Ren, Z. & Bao, Q., 2025. Full-waveform inversion of Rayleigh wave from high-speed-train seismic data for shallow-surface velocity building, *Chin. J. Geophys. (in Chinese),* **68**(4), 1444–1456..

Wang, Z., Li, Y. & Bai, W., 2020. Numerical modelling of exciting seismic waves for a simplified bridge pier model under high-speed train passage over the viaduct., *Chin. J. Geophys. (in Chinese),* **63**(12), 4473–4484..

Yuan, S., Liu, J., Noh, H.Y., Clapp, R.G. & Biondi, B., 2024. Using vehicle-induced das signals for near-surface characterization with high spatiotemporal resolution, *J. geophys. Res.: Solid Earth,* **129**.

Zhao, K., Cheng, F., Xia, J., Guan, J. & Li, Z., 2025. Multistage deep clustering of urban ambient noise for seismic imaging—a case study for train-induced seismic noise, *Geophys. J. Int.,* **242**(3), ggaf273..

Zhou, N., Chen, W., Luo, J. & Wang, X., 2024. Elastic wave equation traveltime inversion with dynamic time warping based on high-speed train seismic data, *IEEE Trans. Geosci. Remote Sens.,* **62,** 1–11.

APPENDIX A: THE RATE OF RELAXATION FUNCTION

The rate of relaxation function $\boldsymbol{\psi}$ in eq. (29) is defined as (A. Tarantola 1988):

$$\boldsymbol{\psi}_{ijkm} = \delta_{ij}\delta_{km}\left(\dot{\psi}_p(t) - 2\dot{\psi}_s(t)\right) + \left(\delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}\right)\dot{\psi}_s(t), \tag{A1}$$

where the dot represents time derivative. The relaxation functions are defined as (H.P. Liu *et al.* 1976; J.O. Robertsson *et al.* 1994):

$$\psi_p(t) = \pi\left[1 + \sum_{l=1}^{L} \tau^p e^{-t/\tau_{\sigma l}}\right] H(t), \tag{A2}$$

and

$$\psi_s(t) = \mu\left[1 + \sum_{l=1}^{L} \tau^s e^{-t/\tau_{\sigma l}}\right] H(t), \tag{A3}$$

where L represents relaxation mechanism. π and μ are relaxed P- and S-wave modulus, respectively. τ^p and τ^s are the level of attenuation for P and S wave, respectively. $\tau_{\sigma l}$ is the relaxation times for the stress. The quality factor Q (the attenuation properties of rock) are defined as (T. Bohlen 2002):

$$Q(\omega, \tau_{\sigma l}, \tau) = \frac{1 + \sum_{l=1}^{L} \frac{\omega^2\tau_{\sigma l}^2}{1+\omega^2\tau_{\sigma l}^2}\tau}{\sum_{l=1}^{L} \frac{\omega\tau_{\sigma l}}{1+\omega^2\tau_{\sigma l}^2}\tau}, \tag{A4}$$

with the variable

$$\tau = \frac{\tau_{\varepsilon l}}{\tau_{\sigma l}} - 1, \tag{A5}$$

where $\tau_{\varepsilon l}$ is the strain retardation time.

APPENDIX B: CALCULATION OF THE GRADIENT OF THE MISFIT FUNCTION USING THE ADJOINT APPROACH

We define a general misfit function (E. Bozdağ *et al.* 2011):

$$E(\mathbf{m}) = \sum_{r=1}^{nr} \int_0^T g(\mathbf{x_r}, t, \mathbf{m})\,\mathrm{d}t, \tag{B1}$$

where nr represents the number of receivers and T denotes the recording duration of the seismograms. The functional g is the misfit between observed data and synthetic data at receiver r with position $\mathbf{x}_r$. In this study, we only have one moving source in the derivation. The model parameters vector $\mathbf{m}$ characterizes the subsurface structure through the density $\rho(\mathbf{m})$ and the relaxation function rate in the tensor $\psi(\mathbf{x}, t)$. And $\mathbf{m}$ can be expressed as a combination of the reference model $\mathbf{m}_0$ and a perturbation model $\delta\mathbf{m}$.

$$\mathbf{m} = \mathbf{m_0} + \delta\mathbf{m}. \tag{B2}$$

The gradient of the misfit function with respect to the model parameters $\mathbf{m}$ is:

$$\zeta E(\mathbf{m}) = \sum_{r=1}^{nr} \int_0^T \partial_{\mathbf{u}} g(\mathbf{x_r}, t, \mathbf{m}) \cdot \zeta\mathbf{u}(\mathbf{x_r}, t, \mathbf{m})\,\mathrm{d}t, \tag{B3}$$

where the notation ζ represents $\frac{\partial}{\partial\mathbf{m}}$, and $\partial_{\mathbf{u}} g$ is the Fréchet derivative of functional g with respect to the synthetic displacement wavefield $\mathbf{u}$. Using the Green's function $\mathbf{G}_{jk}(\mathbf{x}, t-t'; \mathbf{x}', 0)$ of a point source at position $\mathbf{x}'$ at time t' $\mathbf{G}_{jk}(\mathbf{x}, t-t'; \mathbf{x}', 0)$, the solution of the wave propagation equation can be expressed as:

$$\mathbf{u}_j(\mathbf{x}, t) = \int_V \int_{-\infty}^{\infty} \mathbf{G}_{jk}\left(\mathbf{x}, t-t'; \mathbf{x}', 0\right) f_k\left(\mathbf{x}', t'\right) \mathrm{d}t'\mathrm{d}^3\mathbf{x}'. \tag{B4}$$

We assume the eq. (B4) in reference model $\mathbf{m}_0$ is solved by $\mathbf{G}_{jk}^{(0)}$ and $\mathcal{A}_{ni}^{(1)}$ denotes the perturbation $\delta\mathbf{m}$ of the differential operator. The jth component of ζu can be calculated by Born approximation (J. Tromp *et al.* 2005; E. Bozdağ *et al.* 2011):

$$\zeta|_{\mathbf{m}=\mathbf{m_0}} u_j(\mathbf{x}_r, t) = \int_V \int_0^t \mathbf{G}_{jk}^{(0)}\left(\mathbf{x}_r, t-t''; \mathbf{x}'', 0\right) \cdot \frac{\partial}{\partial\delta\mathbf{m}}|_{\delta\mathbf{m}=0} \mathcal{A}_{ni}^{(1)}\left(\mathbf{x}'', \delta\mathbf{m}\right) u_{0_i}\left(\mathbf{x}'', t''\right) \mathrm{d}t''\mathrm{d}^3\mathbf{x}'', \tag{B5}$$

where $u_0(\mathbf{x}, t)$ denotes synthetic seismogram for reference model. Using the reciprocity and reversing time, the adjoint wavefield can be defined as (J. Tromp *et al.* 2005; E. Bozdağ *et al.* 2011; L. Groos *et al.* 2014):

$$\mathbf{u}_k^{\dagger}\left(\mathbf{x}'', \tau\right) = \int_0^{\tau} \int_V \mathbf{G}_{kj}\left(\mathbf{x}'', \tau - t; \mathbf{x}, 0\right) f_k^{\dagger}\left(\mathbf{x}, t\right) \mathrm{d}^3\mathbf{x}\mathrm{d}t \tag{B6}$$

with the adjoint sources:

$$f_k^{\dagger}\left(\mathbf{x}, t\right) = \sum_{r=1}^{nr} \partial_{U_k} g\left(\mathbf{x_r}, T - t, \mathbf{m}\right)|_{\mathbf{m}=\mathbf{m_0}} \delta^3\left(\mathbf{x} - \mathbf{x_r}\right) \tag{B7}$$

Inserting eq. (B5) into eq. (B3), and combining eqs (B6) and (B7), the gradient of the misfit function can be expressed as:

$$\zeta E = -\int_V \int_{t''=0}^{T} \mathbf{u}_k^{\dagger}\left(\mathbf{x}'', T - t''\right) \cdot \frac{\partial}{\partial \delta \mathbf{m}}|_{\delta \mathbf{m}=0} \mathcal{A}_{ni}^{(1)}\left(\mathbf{x}'', \delta \mathbf{m}\right) u_{0_i}\left(\mathbf{x}'', t''\right) \mathrm{d}t''\mathrm{d}^3\mathbf{x}'' \tag{B8}$$