Table S1 Location results of all eight testing events

Event No.	Location method	Location result			Location error/m
		x/m	*y*/m	*z*/m	
1	P_SBL	381689.2	2997740.4	1108.4	20.6
	P_BL	381653.6	2997788.0	1112.8	41.1
	S_BL	381797.1	2997724.1	1061.3	128.0
2	P_SBL	381634.6	2997399.1	1073.9	31.6
	P_BL	381621.4	2997403.3	1073.9	40.4
	S_BL	381685.7	2997380.2	1111.1	42.8
3	P_SBL	381198.1	2996233.5	1035.2	23.6
	P_BL	381171.9	2996237.0	1016.4	25.7
	S_BL	381205.1	2996333.0	999.3	110.6
4	P_SBL	381709.5	2997792.1	1116.4	31.1
	P_BL	381704.1	2997790.8	1117.2	26.4
	S_BL	381737.5	2997781.4	1073.3	63.4
5	P_SBL	381681.0	2997778.6	1097.4	21.0
	P_BL	381731.2	2997710.8	1129.2	72.4
	S_BL	381698.5	2997691.5	1123.3	72.1
6	P_SBL	381592.3	2997241.0	1054.8	37.1
	P_BL	381561.1	2997243.5	1059.9	45.6
	S_BL	381651.9	2997254.6	1051.1	66.1
7	P_SBL	381520.5	2997624.5	1071.8	49.5
	P_BL	381532.4	2997625.4	1052.2	42.7
	S_BL	381423.9	2997544.9	1005.9	115.7
8	P_SBL	381446.5	2998037.6	1001.2	18.5
	P_BL	381421.1	2998062.1	1017.4	39.2
	S_BL	381433.4	2998060.7	1004.4	35.2

Note: The location results correspond to the fifth-ranked location error among the ten time locations.

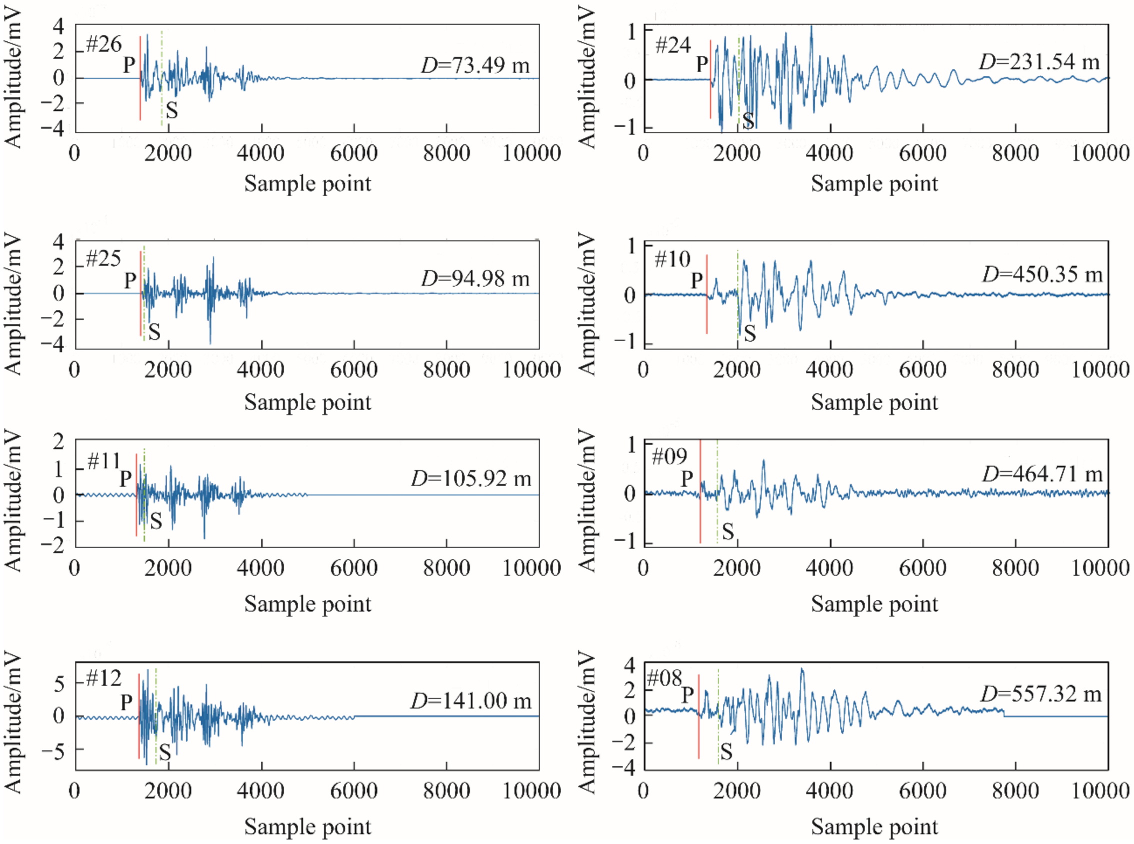

Figure S1 Typical signals and phase arrival time picking results of the blasting event 1 (The vertical red and green lines represent the picked P-wave and S-wave arrival times, respectively. The number indicates the sensor ID and *D* represents the linear distance between the sensor and the blasting event)

Appendix A: Bayesian inversion-based MS source location method

The parameter model $\vec{m}$ is based on prior information of the data, and the Bayesian theorem is mathematically stated as:

$$p(\vec{m}\mid\vec{d})=\frac{p(\vec{d}\mid\vec{m})p(\vec{m})}{p(\vec{d})}=\frac{p(\vec{m})}{p(\vec{d})}\cdot p(\vec{d}\mid\vec{m})\ \propto p(\vec{d}\mid\vec{m}) \tag{A1}$$

where $p(\vec{m}\mid\vec{d})$ is the posterior probability of the model parameter $\vec{m}$ under the constraint of $\vec{d}$. For simplification, an uninformative prior is assumed (i.e., the prior probability $p(\vec{m})$ and the marginal likelihood $p(\vec{d})$ are constants). Then, Eq. (A1) can be converted into $p(\vec{m}\mid\vec{d})\propto p(\vec{d}\mid\vec{m})$, so that $p(\vec{m}\mid\vec{d})$ can be represented by the likelihood function $p(\vec{d}\mid\vec{m})$.

Based on Eq. (A1), the likelihood function can be constructed as follows:

$$p(\vec{d}\mid\vec{m})=\frac{\exp\left[-\frac{1}{2}(\vec{G}-\vec{d})^{\mathrm{T}}\vec{C_{\mathrm{e}}}^{-1}(\vec{G}-\vec{d})\right]}{\sqrt{(2\pi)^{M_1+M_2}\mid\vec{C_{\mathrm{e}}}\mid}} \tag{A2}$$

where $\vec{m}$ is the model parameter comprising the source coordinates (x_0, y_0, z_0) and free weighting w (i.e., $\vec{m}=\vec{m}(x_0,\ y_0,\ z_0,\ w)$); $\vec{G}$ and $\vec{d}$ represent the computed and observed data, respectively, and $\vec{G}-\vec{d}$ contains two parts: the P-wave travel time difference part $[w\cdot\Delta t_{\mathrm{p}}^{1},\ w\cdot\Delta t_{\mathrm{p}}^{2},\ \cdots,\ w\cdot\Delta t_{\mathrm{p}}^{M_1}]$ and the S-wave travel time difference part $[(1-w)\cdot\Delta t_{\mathrm{s}}^{1},\ (1-w)\cdot\Delta t_{\mathrm{s}}^{2},\ \cdots,\ (1-w)\cdot\Delta t_{\mathrm{s}}^{M_2}]$; $\vec{C_{\mathrm{e}}}$ is the uncertainty covariance matrix with diagonal elements $\sigma_1^2,\ \cdots,\ \sigma_1^2,\ \sigma_2^2,\ \cdots,\ \sigma_2^2$.

$$\vec{C_{\mathrm{e}}}=\begin{bmatrix} \sigma_1^2 & \cdots & 0 & & & \\ \vdots & \ddots & \vdots & & 0 & \\ 0 & \cdots & \sigma_1^2 & & & \\ & & & \sigma_2^2 & \cdots & 0 \\ & 0 & & \vdots & \ddots & \vdots \\ & & & 0 & \cdots & \sigma_2^2 \end{bmatrix}_{(M_1+M_2)\times(M_1+M_2)} \tag{A3}$$

where σ_1^2 and σ_2^2 are the covariances of matrix $[w\cdot\Delta t_{\mathrm{P}}^{1},\ w\cdot\Delta t_{\mathrm{P}}^{2},\ \cdots,\ w\cdot\Delta t_{\mathrm{P}}^{M_1}]$ and $[(1-w)\cdot\Delta t_{\mathrm{S}}^{1},\ (1-w)\cdot\Delta t_{\mathrm{S}}^{2},\ \cdots,\ (1-w)\cdot\Delta t_{\mathrm{S}}^{M_2}]$, respectively.

After constructing the Bayesian posterior probability density function, the parametric model is sampled by the Metropolis-Hastings (M-H) algorithm, where the M-H algorithm is an MCMC method that samples the parameters from the posterior probability distribution. First, a random input that contains four parameters (source coordinates $(x_0,\ y_0,\ z_0)$ and free weighting (w) is generated. Then, one randomly selected parameter is updated in each iteration, where the updated step size is equal to the product of the search step size η and $g(\overrightarrow{m'}\,|\,\overrightarrow{m})$, $\eta_1=\eta_2=\eta_3=10$ and $\eta_4=0.01$ are used in this study. For example, the updating of the randomly selected parameter x from x_0 to x' can be written as:

$$x' = x_0+\eta_1\cdot g(\overrightarrow{m'}\,|\,\overrightarrow{m}) \tag{A4}$$

The probability density function $p(\vec{d}\,|\,\overrightarrow{m'})$ of a new parameter model $\overrightarrow{m'}$ can be calculated after updating the inversion parameters. The acceptance ratio $\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})$ of the new model is given as:

$$\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})=\begin{cases}1, & p(\vec{d}\,|\,\overrightarrow{m'})\ge p(\vec{d}\,|\,\overrightarrow{m})\\ p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m}), & \text{else}\end{cases} \tag{A5}$$

The natural logarithm of Eq. (A5) is equivalent to:

$$\ln\left(\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})\right)=\begin{cases}0, & \ln\left(p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})\right)\ge 0\\ \ln\left(p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})\right), & \text{else}\end{cases} \tag{A6}$$

According to Eq. (A2), $p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})$ can be expressed as:

$$p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})=\frac{\exp[-\frac{1}{2}(\overrightarrow{G'}-\vec{d})^{\mathrm{T}}\overrightarrow{C'_{\mathrm{e}}}^{-1}(\overrightarrow{G'}-\vec{d})]}{\sqrt{(2\pi)^{M_1+M_2}\,|\overrightarrow{C'_{\mathrm{e}}}|}}\Bigg/\frac{\exp[-\frac{1}{2}(\vec{G}-\vec{d})^{\mathrm{T}}\overrightarrow{C_{\mathrm{e}}}^{-1}(\vec{G}-\vec{d})]}{\sqrt{(2\pi)^{M_1+M_2}\,|\overrightarrow{C_{\mathrm{e}}}|}} \tag{A7}$$

Taking the nature logarithm of $p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})$, Eq. (A6) may be simplified as:

$$\ln\left(p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})\right)=-\frac{1}{2}\ln|\overrightarrow{C'_{\mathrm{e}}}|+\frac{1}{2}\ln|\overrightarrow{C_{\mathrm{e}}}|-\frac{1}{2}(\overrightarrow{G'}-\vec{d})^{\mathrm{T}}\overrightarrow{C'_{\mathrm{e}}}^{-1}(\overrightarrow{G'}-\vec{d})+\frac{1}{2}(\vec{G}-\vec{d})^{\mathrm{T}}\overrightarrow{C_{\mathrm{e}}}^{-1}(\vec{G}-\vec{d}) \tag{A8}$$

where $\overrightarrow{C'_{\mathrm{e}}}$ is the updated covariance matrix corresponding to the parameter model $\overrightarrow{m'}$ and $\overrightarrow{G'}$ corresponds to computed data from the updated model $\overrightarrow{m'}$.

In the Bayesian inversion, the $p(\vec{d}\,|\,\overrightarrow{m'})\,/\,p(\vec{d}\,|\,\overrightarrow{m})$ is first calculated by Eq. (A8). The $\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})$ may then be obtained by Eq. (A5). The acceptance of the new model $\overrightarrow{m'}$ depends on the $\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})$ and u, where $u\in[0,\ 1]$ is a uniform random number: the new model will be accepted when $\gamma(\overrightarrow{m'}\,|\,\overrightarrow{m})\ge u$; otherwise, the new model will be rejected. In this study, the number of MCMC iterations is set to 100000 for both synthetic events and mine blasting events, and the average of the last 5000 iterations in the post-burn-in period is taken as the final location result.

Appendix B: Influence of bad phase arrival pickings

The location accuracy will be affected when the picked arrival of the first sensor is not very good and the first sensor is always treated as a basis for subtracting. However, in this study, only 80% randomly selected sensors are used in the Bayesian inversion during each location iteration, where the numbering of the first sensor varies in the iteration. This will reduce the influence of larger picking errors. The detailed interpretation is given as follows:

The theoretical travel time difference T_P^t and T_S^t separately are calculated by subtracting the travel time of the first sensor from T_P^i and T_S^i:

$$T_\mathrm{P}^\mathrm{t} = T_\mathrm{P}^i - T_\mathrm{P}^1 = (\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} - \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2})/v_\mathrm{P} \tag{B1}$$

$$T_\mathrm{S}^\mathrm{t} = T_\mathrm{S}^j - T_\mathrm{S}^1 = (\sqrt{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2} - \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2})/v_\mathrm{S} \tag{B2}$$

The observed travel time difference t_P^a and t_S^a (the picked P-wave and S-wave arrival time) separately are calculated by subtracting the arrival time of the first sensor t_P^1 and t_S^1 from t_P^i and t_S^i:

$$t_\mathrm{P}^\mathrm{a} = t_\mathrm{P}^i - t_\mathrm{P}^1 \tag{B3}$$

$$t_\mathrm{S}^\mathrm{a} = t_\mathrm{S}^j - t_\mathrm{S}^1 \tag{B4}$$

When the first sensor contains a picking error, the observed travel time difference will be written as:

$$t_\mathrm{P}^\mathrm{a} = t_\mathrm{P}^i - t_\mathrm{P}^{1\prime} = t_\mathrm{P}^i - t_\mathrm{P}^1 - \Delta t_\mathrm{P} \tag{B5}$$

$$t_\mathrm{S}^\mathrm{a} = t_\mathrm{S}^j - t_\mathrm{S}^{1\prime} = t_\mathrm{S}^j - t_\mathrm{S}^1 - \Delta t_\mathrm{S} \tag{B6}$$

where $t_\mathrm{P}^{1\prime} = t_\mathrm{P}^1 + \Delta t_\mathrm{P}$ and $t_\mathrm{S}^{1\prime} = t_\mathrm{S}^1 + \Delta t_\mathrm{S}$, and Δt_P and Δt_S are the picking errors.

Then, the location objective function F and modified location objective function F' can be written as:

$$\begin{aligned} F = {} & \sum_{i=1}^{M_1} \left| t_\mathrm{P}^\mathrm{a} - T_\mathrm{P}^t \right| + \sum_{j=1}^{M_2} \left| t_\mathrm{S}^\mathrm{a} - T_\mathrm{S}^t \right| = \sum_{i=1}^{M_1} \Big| t_\mathrm{P}^i - t_\mathrm{P}^1 - \\ & (\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} - \\ & \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2})/v_\mathrm{P} \Big| + \\ & \sum_{j=1}^{M_2} \Big| t_\mathrm{S}^i - t_\mathrm{S}^1 - (\sqrt{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2} - \\ & \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2})/v_\mathrm{S} \Big| \end{aligned} \tag{B7}$$

$$
\begin{aligned}
F' &= \sum_{i=1}^{M_1}\left|t_{\mathrm{P}}^{a\prime} - T_{\mathrm{P}}^{\mathrm{t}}\right| + \sum_{j=1}^{M_2}\left|t_{\mathrm{S}}^{a}{}' - T_{\mathrm{S}}^{\mathrm{t}}\right| \\
&= \sum_{i=1}^{M_1}\left|t_{\mathrm{P}}^{i} - t_{\mathrm{P}}^{1} - \Delta t_{\mathrm{P}} - \left(\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} - \right.\right. \\
&\quad \left.\left.\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}\right)\Big/ v_{\mathrm{P}}\right| + \\
&\quad \sum_{j=1}^{M_2}\left|t_{\mathrm{S}}^{i} - t_{\mathrm{S}}^{1} - \Delta t_{\mathrm{S}} - \left(\sqrt{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2} - \right.\right. \\
&\quad \left.\left.\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}\right)\Big/ v_{\mathrm{S}}\right|
\end{aligned} \tag{B8}
$$

It is easy to see that F' will be influenced by the Δt_{P} and Δt_{S}. In other words, the location result will be affected when the picked arrival of the first sensor is not very good and the first sensor is always treated as the basis for subtracting. However, in this study, only randomly selected sensors are used in the Bayesian inversion during each location iteration, where the numbering of the first sensor varies in the iteration. This strategy takes advantage that even if a few stations contain large picking errors, they will only influence the outcome of a specific iteration, rather than significantly affecting the global location result.

Appendix C: Potential combination with other location objective function

TAN et al [11] proposed a misfit function that combines P-wave and S-wave picks (as employed in our study), as well as the time difference between P-wave and S-wave picks ($\Delta t = t_{\mathrm{P}} - t_{\mathrm{S}}$). This misfit function has been acknowledged for its enhanced reliability in constraining source locations in subsurface monitoring scenarios.

$$
F = \left(\alpha_1 \sum_{i=1}^{N}\left(t_{\mathrm{P}}^{i} - T_{\mathrm{P}}^{i} - t_{\mathrm{P}}^{0}\right)^2 + \alpha_2 \sum_{i=1}^{N}\left(t_{\mathrm{S}}^{i} - T_{\mathrm{S}}^{i} - t_{\mathrm{S}}^{0}\right)^2 + \alpha_3 \sum_{i=1}^{N}\left[\left(T_{\mathrm{P}}^{i} - T_{\mathrm{S}}^{i}\right) - \left(t_{\mathrm{P}}^{i} - t_{\mathrm{S}}^{i}\right)\right]^2\right)^{1/2} \tag{C1}
$$

where t_{P}^{i} and t_{S}^{i} denote the observed P-wave and S-wave arrival times at the i-th receiver; T_{P}^{i} and T_{S}^{i} denote the theoretical travel times; t_{p}^{0} and t_{S}^{0} denote the event occurrence time; α_1–α_3 are weighting factors and they are used to balance the contribution of each part; N is the number of sensors capable of picking up both P- and S-waves arrival times.

However, this misfit function employs the L2-norm to construct the location misfit function, which prevents the independent variance computation of each component in the Bayesian inversion. To handle this, a modified misfit function is proposed as:

$$
\begin{aligned}
F = &\sum_{i=1}^{M_1} w_1 \cdot \left|(t_{\mathrm{P}}^{i} - t_{\mathrm{P}}^{1}) - (T_{\mathrm{P}}^{i} - T_{\mathrm{P}}^{1})\right| + \sum_{j=1}^{M_2} w_2 \cdot \left|(t_{\mathrm{S}}^{j} - t_{\mathrm{S}}^{1}) - (T_{\mathrm{S}}^{j} - T_{\mathrm{S}}^{1})\right| + \\
&\sum_{i=j=1}^{N}\left(1 - w_1 - w_2\right) \cdot \left|(t_{\mathrm{P}}^{i} - t_{\mathrm{S}}^{j}) - (T_{\mathrm{P}}^{i} - T_{\mathrm{S}}^{j})\right|
\end{aligned} \tag{C2}
$$

where w_1 and w_2 are weighting factors, $w_1,\ w_2 \in (0,\ 1)$, and they are used to balance the contribution of each part. Eq. (C2) is the objective function of a P- and S-wave arrival time and P-S arrival time combined Bayesian location (P_S_P-SBL) method.

The location performances of the two misfit functions Eqs. (8) and (C2) have been tested by two synthetic events. The synthetic events 1 and 2 separately are located at (381400, 2997000, 1000) m and (381400, 2997000, 1200) m. The propagation velocities of P-wave and S-wave are set to 5200 and 3000 m/s, respectively. Gaussian noises with a standard deviation of 2 and 4 ms were included in the P-wave and S-wave theoretical arrival times, respectively. For the same synthetic event, 50 time locations are conducted. The location errors of the misfit functions (8) and (C2) are depicted in Figure C1. It can be seen that these two misfit functions have a similar location performance.

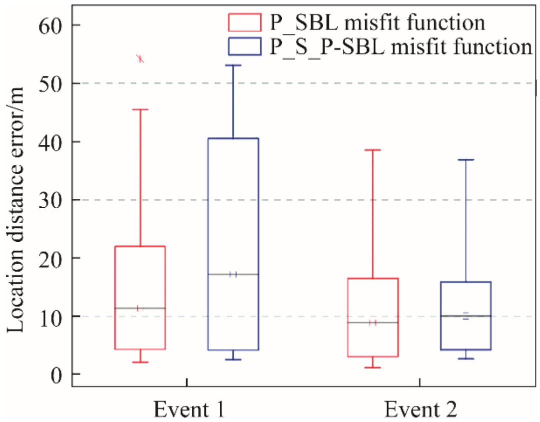

Figure C1 Location errors of Bayesian method based on different misfit functions