

<https://doi.org/10.1038/s43247-025-02917-7>

# Balloon seismology enables subsurface inversion without ground stations

Check for updates

Marouchka Froment <sup>1</sup> ✉, Quentin Brissaud <sup>1,4</sup>, Sven Peter Näsholm <sup>1,2,4</sup> & Johannes Schweitzer <sup>1,3</sup>

Earth's interior structure is derived primarily from data recorded by ground-based seismometers. Such instruments would provide essential insight into the composition and evolution of Venus, but harsh surface conditions hinder their deployment. Balloon-borne seismology offers an alternative by capturing upper-atmospheric infrasonic signatures of seismic waves. We show that subsurface seismic velocities and earthquake source location can be jointly inverted from balloon observations. We demonstrate this method using infrasound signals recorded by four stratospheric balloons following a major earthquake in the Flores Sea, Indonesia. A Bayesian inversion using Markov chain Monte Carlo sampling is implemented to account for trade-offs in the joint location and subsurface velocity estimation. The inverted seismic parameters are consistent with results obtained using ground-based seismometers. This demonstration of the ability to estimate source and velocity parameters without ground deployments motivates further development of seismo-acoustic mission concepts to Venus, and provides new opportunities for seismic exploration in remote Earth regions.

Exploring the interior of Venus could yield crucial insight into its evolution and current geodynamic regime, which remain unknown<sup>1</sup>. The global network of seismometers on Earth's surface was instrumental to developing 1D models of its interior<sup>2–4</sup>, and now contributes to revealing 3D heterogeneities in the mantle and crust<sup>5,6</sup>. Beyond Earth, successful seismometer deployments on Mars and the Moon have provided invaluable information about their structure<sup>7,8</sup> and new seismology missions are now planned to explore Titan<sup>9–11</sup>. However, surface deployment remains challenging on Venus due to the short lifespan of electronics at its high-temperature surface (~460 K)<sup>12,13</sup>.

In recent years, key observations have demonstrated the potential of balloon-borne microbarometers to detect the acoustic signature of seismic waves<sup>14–16</sup>. These signals emerge from the mechanical coupling of seismic ground motion into infrasound—acoustic waves below ~20 Hz—resulting from stress and vertical displacement continuity at the surface<sup>17,18</sup>. Owing to the large velocity contrast between a planet and its atmosphere and low attenuation at low frequencies, seismic waves generate vertically-propagating acoustic waves with dispersion characteristics similar to those of their seismic counterparts<sup>19</sup>. Importantly, this seismo-acoustic coupling is expected to be two orders of magnitude stronger on Venus due to its dense atmosphere<sup>20,21</sup>, enabling the detection of converted seismic waves across a wide range of altitudes. Balloon platforms are therefore considered a realistic alternative to ground deployments to explore Venus' interior<sup>12,13,22,23</sup>. They offer several advantages for subsurface monitoring, such as their mobility and ability to cover large areas. On Venus, balloons

operate under acceptable pressure and temperature conditions between 50 and 60 km altitude and were successfully deployed during the Soviet VEGA missions<sup>24</sup>. They are also relatively inexpensive and benefit from recent advances that enable long-duration flights. Technology for controlled flights is also available<sup>25–27</sup> and offers greater flexibility in mission design and planning; however, it may introduce unwanted noise into infrasound signals.

The recent recordings of earthquake infrasound on Earth, therefore, represent a unique opportunity to assess the use of balloon infrasound for seismic source localization and subsurface exploration. Krishnamoorthy et al.<sup>15</sup> and Bowman and Krishnamoorthy<sup>28</sup> detected and analyzed epicentral infrasound produced by artificial seismic sources and recorded by balloons at close range (<100 km). Although these epicentral signals can provide information on the source, such as its location or mechanism, they mainly propagate in a direct path in the atmosphere from the shallow source to the balloon, and hence carry little information on subsurface seismic properties. Brissaud et al.<sup>19</sup> detected a magnitude 4.2 earthquake using free-floating balloons in Southern California. However, signals were recorded at only one balloon, did not show body wave arrivals, and the surface waves had a low Signal-to-Noise Ratio (SNR), which prevented the joint inversion of source and subsurface properties. A year later, Garcia et al.<sup>29</sup> reported the detection of an  $M_w$  7.5 earthquake in Peru and an  $M_w$  7.3 earthquake in the Flores Sea using freely floating stratospheric balloons from the Strateole-2 campaign<sup>30</sup>. In particular, infrasound from the Flores Sea earthquake was recorded by four balloons at high SNR. Garcia et al.<sup>29</sup> showed general

<sup>1</sup>NORSAR, Kjeller, Norway. <sup>2</sup>Department of Informatics, University of Oslo, Oslo, Norway. <sup>3</sup>Centre for Planetary Habitability (PHAB), University of Oslo, Oslo, Norway. <sup>4</sup>These authors jointly supervised this work: Quentin Brissaud, Sven Peter Näsholm. ✉e-mail: [marouchka.froment@norsar.no](mailto:marouchka.froment@norsar.no)

agreement between balloon pressure signals and ground-based vertical velocity records. Recently, Gerier et al.<sup>31</sup> modeled this event numerically, including the seismo-acoustic coupling, and demonstrated that major seismic phases—P and S body waves and Rayleigh surface waves (LR)—are identifiable in the balloon data. However, it is still largely unknown to what extent such waveforms can provide insight into seismic sources and seismic velocity models.

In the present contribution, we demonstrate that body- and Rayleigh-wave arrival times at various frequencies provide sufficient information to constrain these seismic parameters through a Bayesian inversion approach, even with a small number of balloon stations. We apply this inversion to jointly retrieve the hypocenter of the 2021  $M_w$  7.3 Flores Sea earthquake and a 1D model of subsurface seismic velocities in the region. The inversion is first tested using P, S, and LR arrivals identified at ground stations, and then using arrivals identified in Strateole-2 balloon recordings. We finally quantify the uncertainty in the retrieved source location and seismic velocities and discuss the sensitivity of the inversion method.

## Results

### A joint approach for the inversion of source and subsurface parameters

Earthquake-induced infrasound signals are scaled images of the vertical velocity of the ground surface below the balloon<sup>29,32</sup>. As low-frequency coupled waves propagate mostly vertically and up to relatively low stratospheric altitudes, they remain unaffected by dispersion or non-linear distortion through the Earth's atmosphere. A forward model of arrival times at balloon platforms can thus be readily derived from classical seismological methods. Consequently, we use both body- and surface-wave arrival times in several frequency bands to retrieve the source and subsurface parameters. Relying on arrival times instead of full waveform modeling eliminates the need for an accurate source model and the reliance on low-frequency waveforms, which are typically contaminated by buoyancy oscillations and turbulence in balloon data<sup>29,33</sup>. For planetary exploration, a joint subsurface and source inversion is required due to our lack of prior knowledge of subsurface structures and source locations. Additionally, the sparsity of balloon networks on Earth, and possibly on Venus, calls for careful assessment of uncertainty in hypocenter coordinates<sup>34</sup>. To solve the ill-posed hypocenter-velocity problem<sup>35</sup>, we employ a Bayesian approach, which performs a global search through model space using a Markov chain Monte Carlo (MCMC) method (see, e.g., the monograph by Tarantola<sup>36</sup>). This approach combines the misfit between predicted and observed arrival times (likelihood) with the provided a priori information for each inverted parameter (prior) to infer the probability distribution for these parameters (posterior). The present MCMC inversion is adapted from the Ensemble Sampler<sup>37,38</sup>.

The inverted source variables are origin time  $t_s$ , source latitude and longitude, and depth ( $lon_s$ ,  $lat_s$ ,  $h_s$ ). The subsurface is modeled as six homogeneous layers over a halfspace, and the shear wave velocity  $v_{s,i}$ , Poisson's ratio  $\nu_i$  and thickness  $H_i$  of each layer  $i$  are inverted for. Prior ranges for these variables are described in the Methods and Supplementary Table S3.

### Concept validation on Earth data: the 2021 Flores earthquake

The Flores Sea earthquake occurred on December 14, 2021, with a magnitude of  $M_w = 7.3$ . Following relocation, Supendi et al.<sup>39</sup> associated the event with the Kalaotoa fault system, identifying a strike-slip mechanism at a depth of 12.2 km. This aligns with USGS estimates of 14.2 depth, based on source location, and 17.5 km depth from moment tensor inversions<sup>40</sup>.

Few subsurface velocity models are available near the Flores Sea, a region characterized by high heterogeneity due to the presence of several subduction zones. To provide a meaningful reference for the interpretation of our inverted subsurface models, we define a Median Model based on the median of CRUST1.0 models<sup>41</sup> and the LLNL-G3D-JPS tomographic models in the mantle<sup>42</sup> below our stations (see Supplementary Fig. S4). In these comparisons, we consider 15 km depth, latitude  $-7.6^\circ\text{N}$  and longitude

$122.2^\circ\text{E}$  as the reference hypocenter (USGS CMT solution<sup>40</sup>), and 03:20:23 UTC as the reference origin time.

At the time of the event, four Strateole-2 balloons, identified as TTL4-07, TTL4-15, TTL5-16, and TTL3-17, were located between 680 and 2800 km to the northeast of the event. The balloon inversion uses body and surface wave arrival times extracted from their pressure traces. The Strateole-2 pressure sensor has a flat frequency response below about 0.3 Hz<sup>29,43</sup>, however, the SNR is low at long periods due to the presence of the balloon buoyancy resonance<sup>33</sup>. These oscillations were partly corrected using a method inspired from Podglajen et al.<sup>43</sup> (see Methods and Supplementary Fig. S3 for details). P-wave arrival times were picked for the four balloons with an estimated uncertainty between 7 and 35 s, and S wave arrivals with uncertainties between 8 and 49 s. Due to the low-frequency noise, the LR arrival could only be identified with confidence for TTL3-17 and TTL5-16 between 0.005 and 0.1 Hz, with a mean uncertainty of around 50 s. The picks are shown in more details in the Supplementary Figs. S1–S2, and S33–S36.

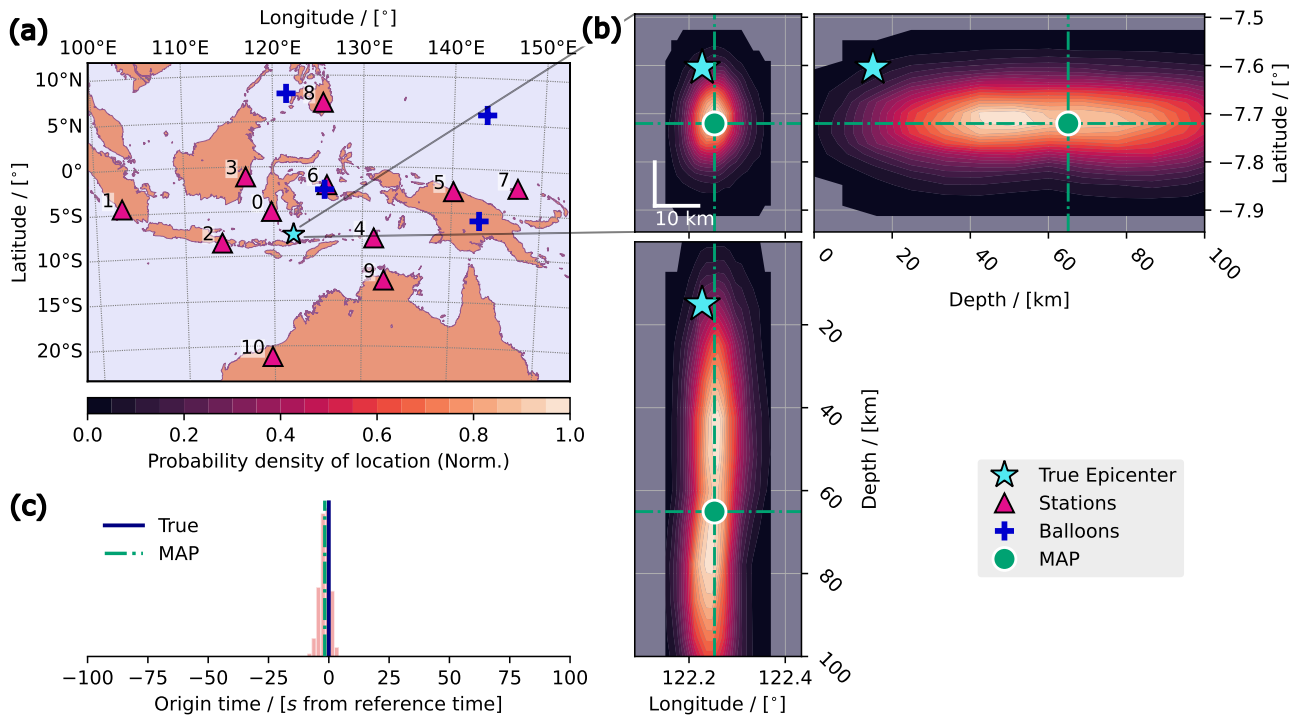
In order to assess the robustness of this infrasound-based inversion, we also construct a reference source and subsurface model through the inversion of data recorded at 11 seismic stations, selected among the Global Seismograph Network, the Australian National Seismograph Network, and the German GEOFON seismic network. This separate inversion allows us to build confidence in the joint inversion technique, and to compare the resolution obtained from a small number of receivers at low SNR—the balloon case—to the one obtained from a typical dense ground network of receivers at high SNR—the seismic case. For consistency, we pick the seismic arrivals using vertical velocity signals from seismic stations in the Flores region, emulating a single-component infrasound signal. The 11 chosen stations are illustrated in Fig. 1a and detailed in Supplementary Table S1 and Supplementary Figs. S22–S32. For these stations, the uncertainty of the P wave arrival times ranged between 1 and 2 s, 10 to 40 s for S waves, and 20 to 100 s for LRs between 0.002 and 0.2 Hz.

### Source and subsurface as seen from seismic data

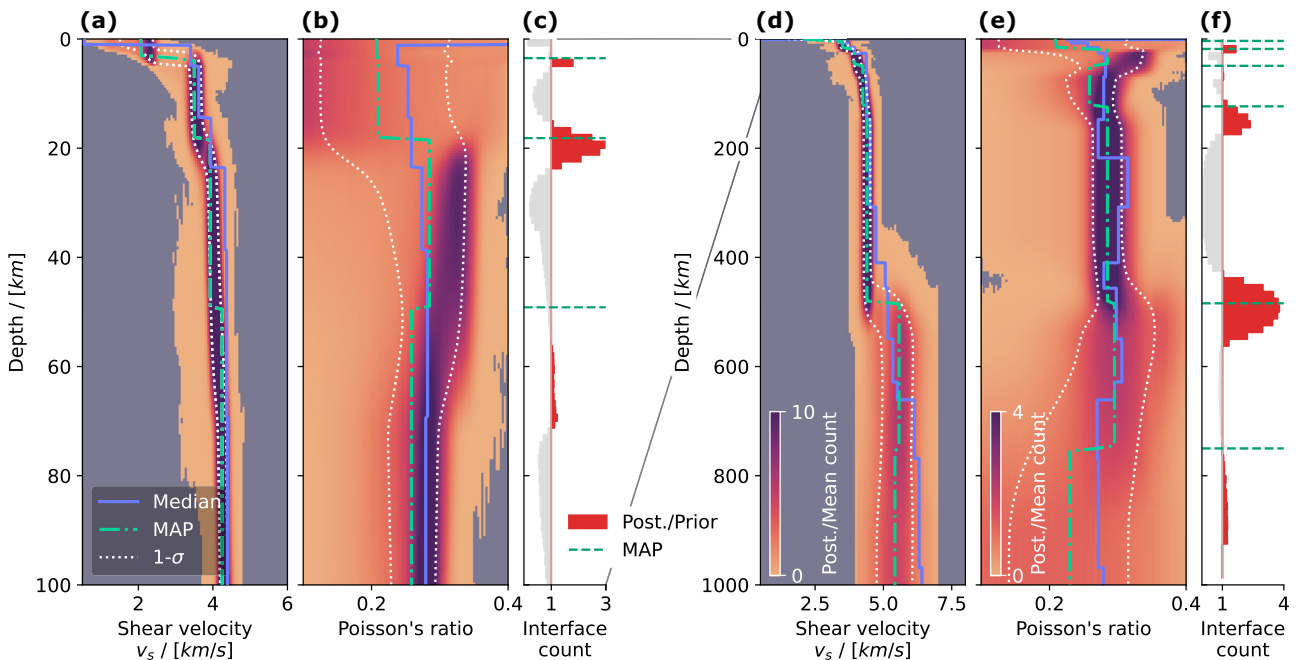
The reference joint inversion is performed using only seismic data picks, obtained from 11 stations. The MCMC simulations return an ensemble of source and subsurface parameters forming the posterior probability distribution. To interpret these results, we reduce the dimensions of the posterior by calculating marginal distributions, and by estimating the Maximum A Posteriori (MAP) parameters, i.e., the solution maximizing the posterior distribution function (see Methods and Supplementary Table S4).

The marginal distribution of the source parameters inverted from these arrival times is shown in Fig. 1b. The MAP source location is shifted 13 km south of the true epicenter, at a slightly greater distance to the majority of the stations, which are to the north. This longer travel time is accommodated by a slightly earlier source origin time, with the MAP value  $\sim 2 \pm 2$  s earlier than the reference time (1c). The inversion also favors a source about 50 km deeper than the reference solution, with  $\sim 100$  km uncertainty. The marginal distributions of source parameters follow Gaussian distributions with little trade-offs between variables.

Figure 2 displays the marginal posterior distributions for the shear wave velocity  $v_s$  and the Poisson's ratio  $\nu$  as function of depth. Both  $\nu$  and  $v_s$  appear constrained down to  $\sim 500$  km depth. The MAP models are in good agreement with the Median profile, constructed from global tomographic Earth models, especially for the shear wave velocity. Posterior values of  $v_s$  have a  $1-\sigma$  uncertainty of  $\pm 0.1$  km s<sup>-1</sup> in the crust and upper mantle layers, and  $\pm 0.6$  km s<sup>-1</sup> in the top sediment layer. This thin, top sediment region has the least well defined velocities, likely due to the high variability of LR dispersion above 0.1 Hz (Fig. 2a) and its limited sensitivity to such narrow features. The  $1-\sigma$  uncertainty becomes 0.5 to 0.6 km s<sup>-1</sup> in the lowermost layer and halfspace (Fig. 2c). The Poisson's ratio takes values between 0.21 and 0.29, within the range expected for most minerals<sup>44</sup>. It is constrained with a large uncertainty of  $\pm 0.02$  between 100 and 400 km depth, and is otherwise undefined.



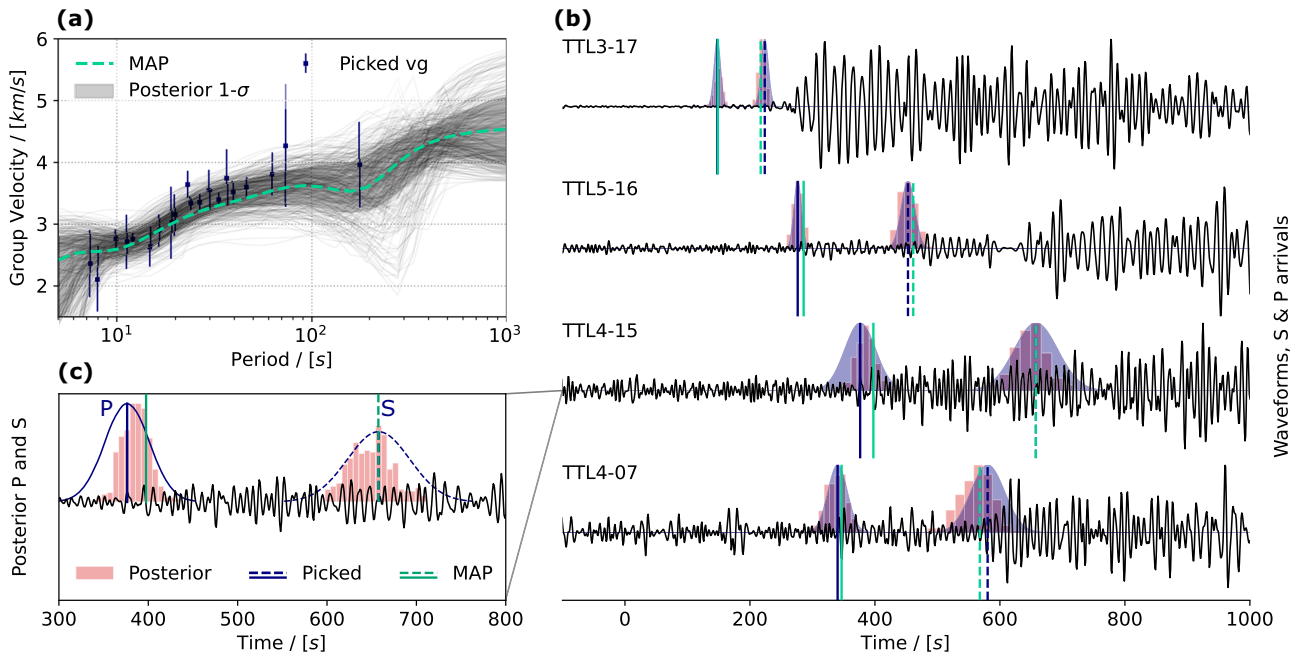
**Fig. 1 | Source origin inverted using 11 seismic stations.** **a** Map of chosen ground seismic stations for the inversion of the 2021  $M_w$  7.3 Flores earthquake. The four Strateole-2 balloons are marked with blue crosses for comparison. Plot **b** shows the marginal distribution of the source hypocenter, up to scale between horizontal and vertical slices. **c** The histogram of source origin time, centered around the true value of zero, with the MAP solution in green.



**Fig. 2 | Subsurface models for shear wave velocity  $v_s$  and Poisson's ratio  $\nu$  below the Flores Sea, inverted using 11 local seismic stations.** The marginal distributions of models are shown as normalized 2D histograms, down to 100 km in (a) and (b) and to 1000 km in (d) and (e). The Median literature model is shown in blue, the MAP in green and the 1- $\sigma$  probability region in dashed white lines. Red histograms in panels of (c) and (f) show regions with a high probability of presenting an interface, or strong gradient in subsurface properties, together with the MAP interfaces in green. See the Methods section for details on normalization and the Interface count metric.

The inversion method also returns a distribution of layer thicknesses, which can be converted to a more easily interpretable distribution of interface depths through cumulative summation. In Fig. 2c and f, we compare the posterior distribution of interfaces to the prior, thereby highlighting depth ranges with a higher probability of hosting a change in

subsurface properties, independently from the prior model distribution. Three interfaces, or regions of strong velocity gradients, are strongly suggested in this inverted model: at  $20 \pm 4$  km depth in the crust, and  $150 \pm 30$  km, and  $500 \pm 70$  km depth in the mantle. A very shallow interface is also suggested at 4 km depth.



**Fig. 3 | Infrasound signals and arrivals at balloons following the  $M_w$  7.3 Flores earthquake, compared to arrivals predicted from the inverted models. a** Picked Rayleigh wave group velocities, derived from picked arrival times assuming the true location and time of the  $M_w$  7.3 Flores earthquake, shown in blue. These measurements are compared to 400 group velocity curves constructed from a random

selection of posterior models. The MAP model is shown in green. **b** Pressure waveforms used to pick arrivals, bandpass-filtered between 0.06 and 0.2 Hz, with picked times shown in blue and arrival times predicted from the MAP in green. **c** Zoom on TTL4-15 signal, showing the posterior distribution of arrival times for P and S waves compared to the picked value and its uncertainty in blue.

**Source and subsurface inverted from a network of four balloons**

The balloon inversion result fits the arrival times adequately, as evidenced by the strong match between the observed and posterior distribution of arrivals in Fig. 3. The low number of arrival-times picked from the balloon data, combined with their large uncertainty, limits the precision of the source location. Indeed, Fig. 4a, b shows a larger uncertainty in epicenter using the Strateole-2 balloons, compared to using a subset of four regional seismic stations at similar locations with more precise P, S, and LR picks. The Strateole-2 data inversion returns a MAP epicenter 35 km away from the true epicenter, at coordinates  $-7.5 \pm 1.0^\circ$  latitude and  $122.5 \pm 0.7^\circ$  longitude, against 32 km distance with an uncertainty of  $\pm 0.6\text{--}0.8^\circ$  in latitude and longitude using four seismic stations. This corresponds to an uncertainty of 200 km around the true epicenter. Still, despite the low SNR of balloon infrasound signals, the inversion framework enables an accurate characterization of the source location—a critical task when the network is sparse or poorly distributed in terms of source-station azimuth<sup>34</sup>.

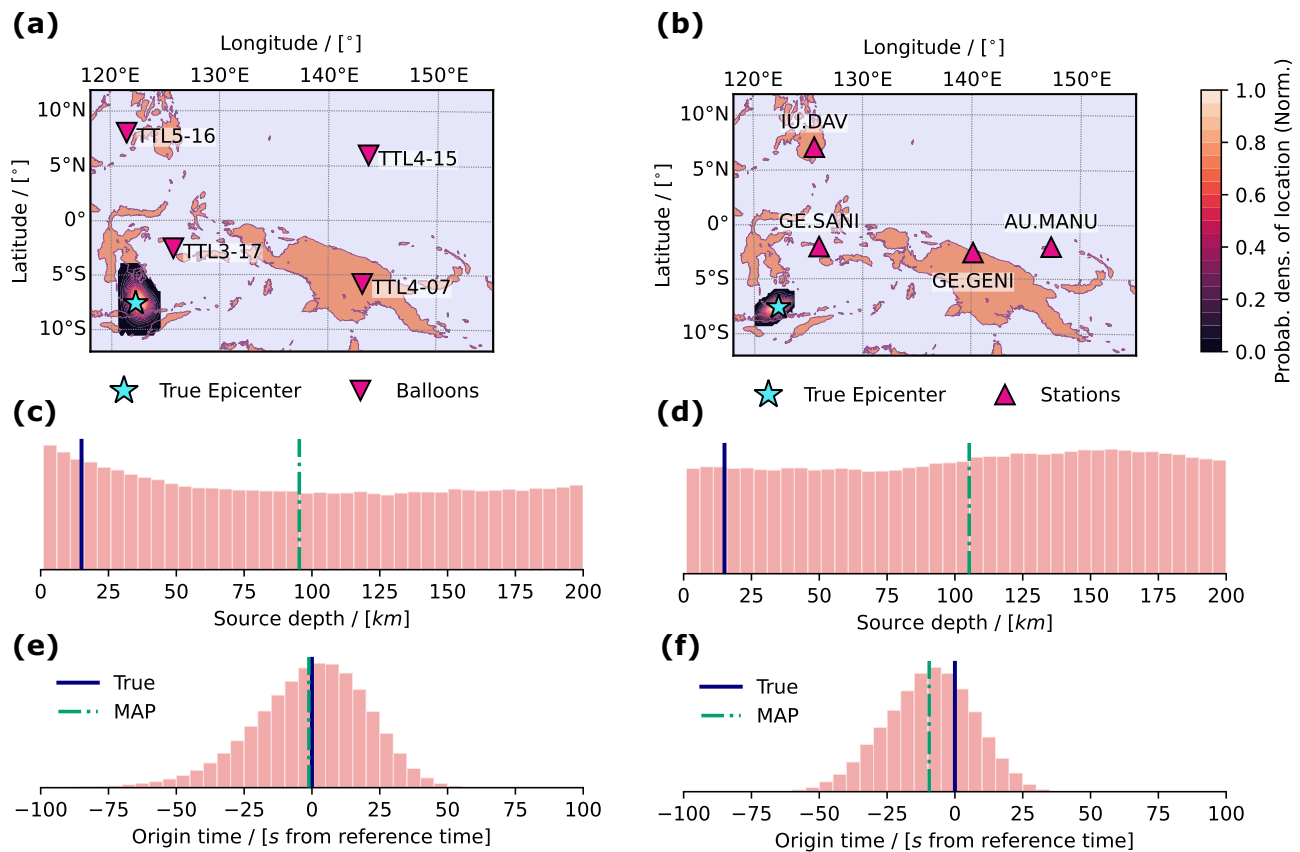
For the balloon inversion result, the origin time is about  $1 \pm 22$  s earlier than the published value (Fig. 4e), while it is predicted to be  $9 \pm 16$  s earlier using four local seismic stations (Fig. 4f). The published value is comprised within the posterior uncertainty, and the difference in MAP value could be due to the slight difference in distribution of balloon stations over azimuth and distance compared to the local ground stations, or to biased picks among the ground stations. In both cases, the source depth displays an almost uniform posterior distribution down to 200 km depth and cannot be constrained (Fig. 4c, d). Similarly, the 11-station inversion returned a MAP depth of  $\sim 40\text{--}50$  km (Fig. 1b) rather than the 12.2 to 17.5 km previously published<sup>39,40</sup>. Inverting for the source depth without stations close to the source (less than a few source depths away) or without identified depth phases is notoriously difficult, hence making this result unsurprising<sup>45</sup>.

With only four P-wave picks, the Strateole-2 data insufficiently constrains the Poisson’s ratio in the subsurface, where both posterior distributions of  $\nu$  or  $\nu_p$  are hardly distinguishable from uniform priors (shown

in Fig. S17 of the Supplementary Information). However, P, S, and LR picks provide constraints on the posterior distribution of  $\nu_s$ , which is shown on Fig. 5a, c. The MAP and posterior models match the Median Model within one standard deviation down to around 600 km depth, and shear wave velocities are constrained with a  $1\text{--}\sigma$  uncertainty of  $\pm 0.3$  to  $\pm 0.6$  km  $s^{-1}$  between 10 and 400 km depth.

Once again, the interface count metric evaluated from the posterior distribution favors changes in subsurface properties at specific depth ranges, in particular at  $19 \pm 6$  km depth in the crust (see Fig. 5b), a value similar to the 11-station inversion. The CRUST1.0 model predicts high variability of crustal thickness in the Malay Archipelago, with values ranging from 10 km to about 40 km (see Supplementary Figs. S5 and S6). However, the median crustal thickness calculated using CRUST1.0 along the great circle path between the source and the seismic stations is 23 km, and respectively 15 km for the balloons, which is consistent with the inverted interface depth (details in Supplementary Table S2). Thus, the distribution of inverted interfaces likely represents the typical Moho depth between the Flores Sea event and the receivers.

Three deep regions of velocity change are suggested at  $420 \pm 50$  km depth, between 60 and 200 km and below 800 km depth, although with little confidence (Fig. 5d). These depths are similar to those found in the 11-station inversion, at  $150 \pm 30$  km and  $500 \pm 70$  km. No global mantle interface is known between 60 and 200 km depth. The wide uncertainty in inverted depth suggests that the high interface probability may not indicate an abrupt change in thermochemical properties, but rather a smooth increase in velocity, as expected at the top of the mantle. Meanwhile, important velocity changes are known to occur in the mantle transition zone, such as at 410 km depth where the olivine-Wadsleyite phase transition is found<sup>46</sup>. However, both inverted interface distributions show large uncertainties below 400 km, due to the low sensitivity of P, S, and LR travel times to changes at these depths (see Supplementary Figs. S7 and S8). Thus, this concentration of interfaces could rather indicate that a gradual increase in seismic velocity is necessary in this mantle region to fit LR arrival times in the 0.002–0.005 Hz range.



**Fig. 4 | Source origin inverted from four balloons (left) or four seismic stations (right).** Posterior distributions of source location (a), depth (c) and origin time (e) inverted from four Strateole-2 balloons and (resp. (b), (d), (f)) four seismic stations at similar locations. The true (resp. MAP) values are shown with blue (resp. green) vertical lines.

### Sensitivity of the inversion method

We further examine the sensitivity of the joint inversion method to subsurface and source parameters, and the resolving power of the body wave and surface wave data employed.

There is no well-established method to pick body wave arrivals from single-component balloon recordings. In particular, the S arrivals were measured with large uncertainties due to the lack of clear impulsive arrivals in the balloon pressure data. We assess the sensitivity of our results to the picking of these arrivals by performing an inversion using only P and LR arrivals. The results, provided in Supplementary Figs. S12 and S13, show little difference compared to Figs. 5 and 4, with a ~4% difference in MAP  $v_s$ . The predominant effect is a larger uncertainties in the posterior distribution of  $v_s$  and source location (up to +50%). P and LR picks are thus found to be sufficient to constrain the source location and the main interfaces in the subsurface.

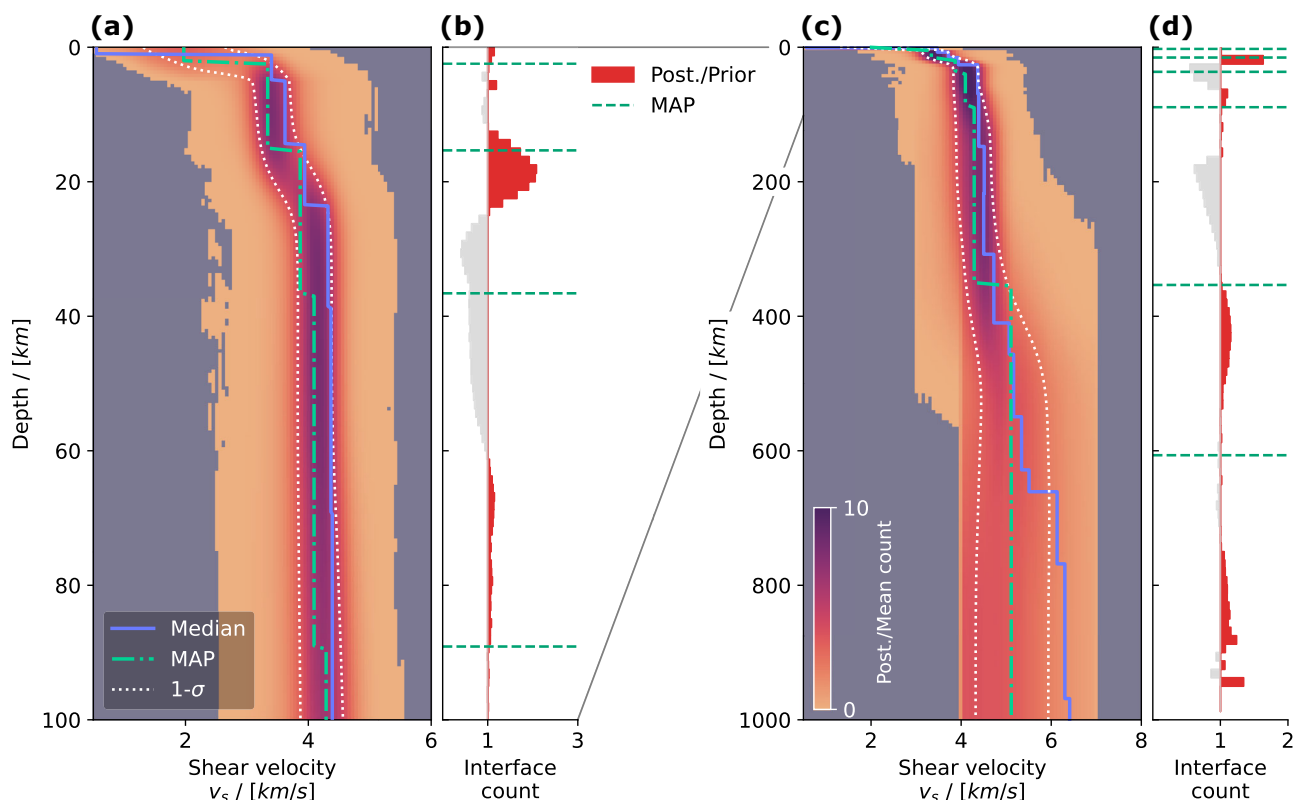
To further evaluate the sensitivity of our inversion results to the different types of seismic waves, we perform inversions using only LR arrivals, and respectively only P and S arrivals. This assessment uses the ground seismic data to ensure a higher resolution in the comparison. The results are summarized in Fig. 6. Body waves are shown to have the strongest sensitivity to source location, with an uncertainty ellipse almost identical to our complete inversion (Fig. 6a) and a slight sensitivity to depth (Fig. 6b). Both datasets have similar sensitivities to origin time (Fig. 6c), with the LR-based inversion yielding a delayed source. Indeed, the LR-based source location is shifted to the north, and is therefore closer to the majority of the seismic stations, thereby shortening the travel times and requiring a later origin time. Finally, Fig. 6d illustrates the subsurface resolution provided by each dataset, by comparing the  $1 - \sigma$  uncertainty of the inverted  $v_s$ , the first 100 in the crust are poorly resolved using the body wave information, contrary to the Rayleigh waves, which display high sensitivity in the frequency range considered (see Supplementary Figs. S7–S8). Both datasets are sensitive to

shear wave velocities between 100 and 400 km depth, and lose resolution at greater depths. Individual plots of  $v_s$  for the P, S-based, and LR-based inversions can be found in Supplementary Fig. S14.

The MCMC framework also allows for an analysis of trade-offs between inverted parameter in the context of the hypocenter-velocity problem. Figure 7 shows the marginal posterior probability densities of several inverted variables along one and two dimensions. Trade-offs are observed between the origin time  $t_s$  and the source epicenter defined by ( $lon_s, lat_s$ ) (Fig. 7a, b). Regarding the subsurface, there are complex, non-linear trade-offs between the thickness of layers and their seismic velocities (Fig. 7c, d). This is a known phenomenon, due to the fact that Rayleigh wave group velocities are sensitive to seismic velocities over a range of depths (see Supplementary Fig. S8). Finally, there are also trade-offs between the Poisson's ratio, which for the balloon inversion is weakly resolved between 0.1 and 0.4, and the shear wave velocity in the same layer (Fig. 7e), and between shear wave velocities in adjacent layers (Fig. 7f). These trade-offs mean that a large number of solutions exist for the non-linear, ill-defined system of equations defining arrival times (Methods, Eq. (1)). Yet, our probabilistic inversion framework still highlights regions of higher probability for source location and subsurface properties.

There are ways to mitigate the non-uniqueness of the solution space and improve the inversion convergence. One possibility is to fix some of the least resolved parameters to reduce the dimensionality of the inverse problem. In the balloon inversion case, fixing the Poisson's ratio appears particularly relevant. Supplementary Figs. S15 and S16 show results of a test adopting  $\nu \approx 0.28$ , a value representative of the Median model. Location and subsurface solutions are within the uncertainties of the general problem, and posterior uncertainties differ only by a few percent.

Alternatively, the misfit function itself can be reshaped to eliminate some of the undetermined parameters. The *time-difference of arrival* (TDOA) method is a typical application of this concept in seismology,



**Fig. 5 | Subsurface velocity model inverted using Strateole-2 infrasound data.** Shear wave velocities inverted from balloon data, down to 100 km (a) and 1000 km depth (c), along with the associated ratio of posterior and prior interface depth

distributions ((b) and (d), resp.). The MAP model and interface depths are shown in green, and the Median literature model in blue, along with the 1-σ probability region in dashed white lines.

adapted from the seminal method of hyperbolas of Milne<sup>47</sup>, whereby the inversion seeks to minimize the misfit between observed and predicted travel time differences, instead of absolute travel times<sup>48</sup>. This formulation of the misfit effectively eliminates the origin time variable. We tested this inversion method for the balloon inversion, using the earliest P-wave as a reference for all travel-time differences. Supplementary Figs. S18 and S19 show results identical to the original formulation, confirming that  $t_s$  is effectively ill-defined in this joint inversion problem. Finally, the misfit, or likelihood function, could alternatively be formulated using an L1-norm instead of an L2-norm of time differences (see “Methods” section for its definition), which has the advantage of being less sensitive to data outliers<sup>49</sup>. In the balloon case, this formulation yields nevertheless similar to the original L2-norm, as show in Supplementary Figs. S20 and S21, with posterior uncertainties differing by at most 30%.

### Perspectives for balloon seismology

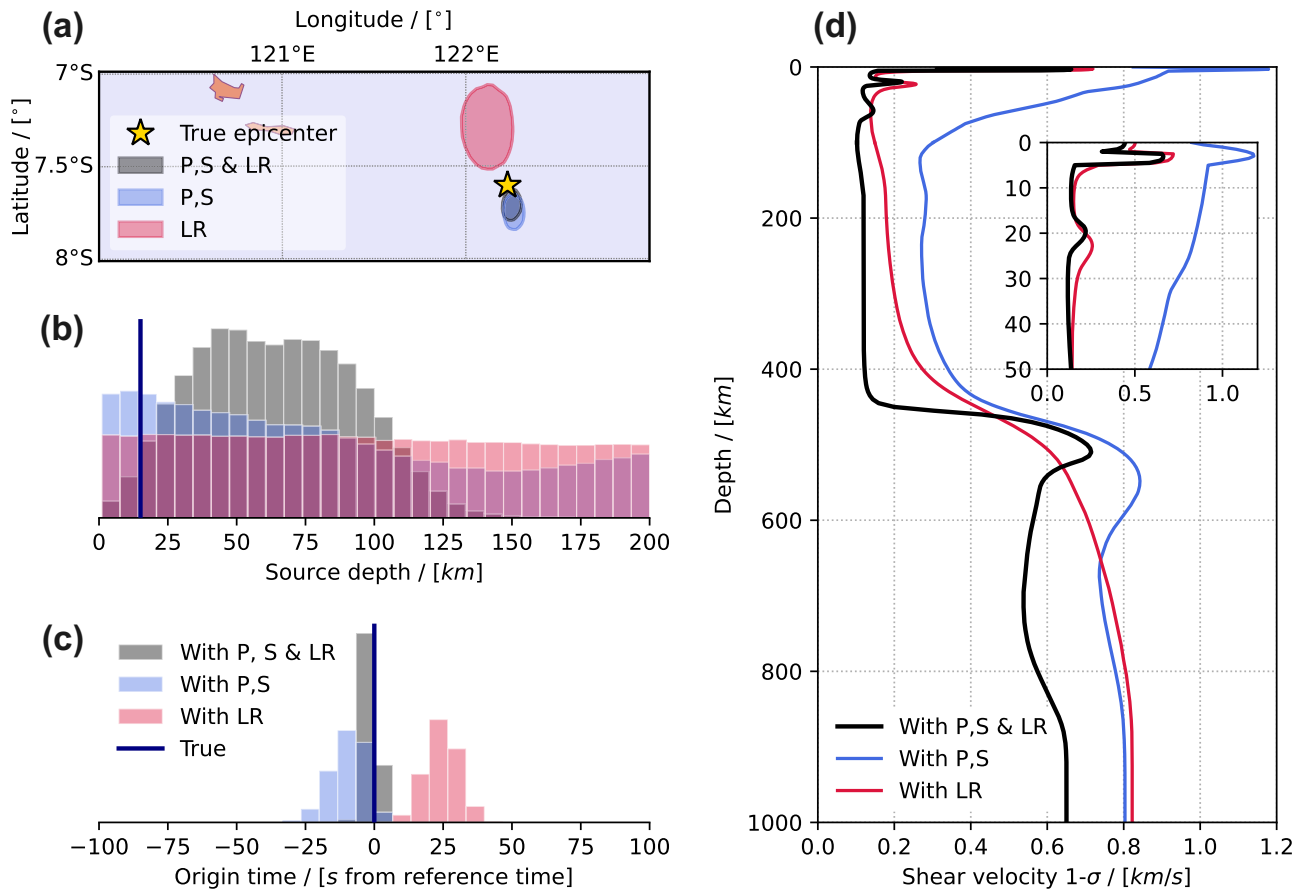
We achieve the inversion of a subsurface seismic velocity profile based on earthquake infrasound signals recorded at airborne balloon platforms. The distributions of subsurface profiles inverted using data from 4 balloons (Fig. 5) are consistent with the Median Model, a median representation of seismic velocities in the Malay Archipelago from the literature. We also capture a crustal interface at 19 km, consistent with the local Moho depth, with ±6 km uncertainty. The Bayesian approach enables an examination of parameter trade-offs and distributions in the simultaneous estimation of source location and subsurface velocity.

We identify the main challenge of this balloon-based inversion as the reliable picking of seismic phases in single-component data. Besides the lack of waveform polarization estimation, balloon signals suffer from lower SNR on Earth at low frequencies due to buoyant motion through atmospheric perturbations, and possibly to local turbulence induced by this motion<sup>31</sup>. Without knowledge of the source location, coda arrivals may also be misidentified: a broadband energy pulse from a wind burst or a secondary

P-phase can be wrongly interpreted as an S-wave arrival, and higher-mode LR energy can obscure the fundamental mode at higher frequencies. These limitations could be mitigated in the future by improved signal processing methods, such as template matching or machine learning-based picking, as well as by additional instrumentation. Recent studies have proposed using Inertial Motion Units (IMUs) onboard balloons to better characterize the polarization of the velocity perturbation associated with a pressure wave and derive its direction of arrival, providing additional information for the interpretation of the signal and the resolution of the location problem<sup>50</sup>.

In conclusion, our findings confirm the viability of using balloons for seismic exploration. Our results strengthen the case for balloon seismology on Venus, as we demonstrated the ability to address challenges related to unknown sources and subsurface properties through a joint inversion using balloon infrasound data. Beyond the arrival time analysis proposed by the present study, seismic analyses based on signal amplitude, or even the full waveforms could be envisioned, potentially allowing to access additional information on the source depth and mechanism, or on the subsurface properties on Earth or Venus. Consequently, balloon seismology could provide valuable insights into the planet’s current tectonic activity and internal structure.

On Earth, the free-floating nature of balloons enables them to reach remote regions, such as the oceans and poles, where seismic sensor deployment remains challenging. Balloons have proven their ability to capture epicentral infrasound at close range<sup>28</sup> and volcanic infrasound at larger distances<sup>43</sup>, providing potential constraints on various types of remote seismo-acoustic sources. Earthquake infrasound data can for instance be used to determine the event magnitude<sup>32</sup>. Under certain conditions, signals recorded by balloons near the source may match the quality of those from distant seismic stations, and they could thus serve as complement to conventional monitoring networks. Further investigations could help identify the regions and typical magnitudes for which balloons fill an observational gap also on Earth.



**Fig. 6 | Comparison of inversions using different datasets.** **a** The 1- $\sigma$  contours of the PDFs for source location inverted using all available wave picks (black, same as in Fig. 1), or only using body waves (blue) and Rayleigh wave picks (red). The location obtained from body wave picks is almost superimposed with the one of Fig. 1. **b, c** The PDFs of source depth and origin time for the respective datasets. **d** The 1- $\sigma$  uncertainty of  $v_s$  retrieved from the posterior distribution of each inversion, with a close-up on the first 50 km.

## Methods

### Markov chain Monte Carlo inversion

Sophisticated Monte Carlo sampling approaches, such as Ensemble Sampling<sup>37</sup>, Hamiltonian Monte Carlo<sup>51</sup>, or Parallel Tempering<sup>52</sup>, allow a thorough search through model space robust to the presence of multiple local minima.

In the current work, we employ the open-source implementation of an Ensemble Sampler written in Python named `emcee`<sup>38</sup> as the basis for our inversion framework. Its use here is motivated by its simplicity of application and its efficiency when sampling highly correlated parameter spaces, which can be encountered in the hypocenter-velocity problem.

Each MCMC simulation is run for  $10^6$  iterations on an ensemble of 50 chains, resulting in a total of about  $50 \times 10^6$  samples. The simulations are run with 32 CPUs on a high-performance computing server.

### Forward model and misfit

The inversion method is based on measurements of arrival times for different seismic phases, namely P, S, and LRs at a network of receivers. Considering a common and arbitrary reference time for the receivers and source of interest, the time of arrival of a wave  $W$  at receiver  $R$  can be written:

$$t_{W,R} = t_s + \Delta t_{W,R} + \Delta t_{air,R}, \quad (1)$$

where the earthquake occurs at time  $t = t_s$  (s) since the reference time,  $W$  is the wave type among seismic or air-coupled P, S, and LR,  $\Delta t_{W,R}$  is the seismic travel time from the source to the piercing point at a ground station or below a floating balloon, and  $\Delta t_{air,R}$  is the additional travel time from the surface to the floating balloon, if applicable. For simplicity, we use the origin time of the

Flores earthquake published by the USGS (03:20:23 UTC on 14 December 2021) as our reference time<sup>40</sup>. In the case of LRs,  $\Delta t_{W,R}$  is frequency-dependent, allowing to model a range of arrival time measurements. At a specific receiver  $R$ , the travel time  $\Delta t_{air,R}$  is independent of the phase type, and can be estimated knowing the balloon altitude and the atmospheric state at the time of the event. The recording of multiple phases  $W$  at several receiver locations provides a system of equations similar to Eq. (1).

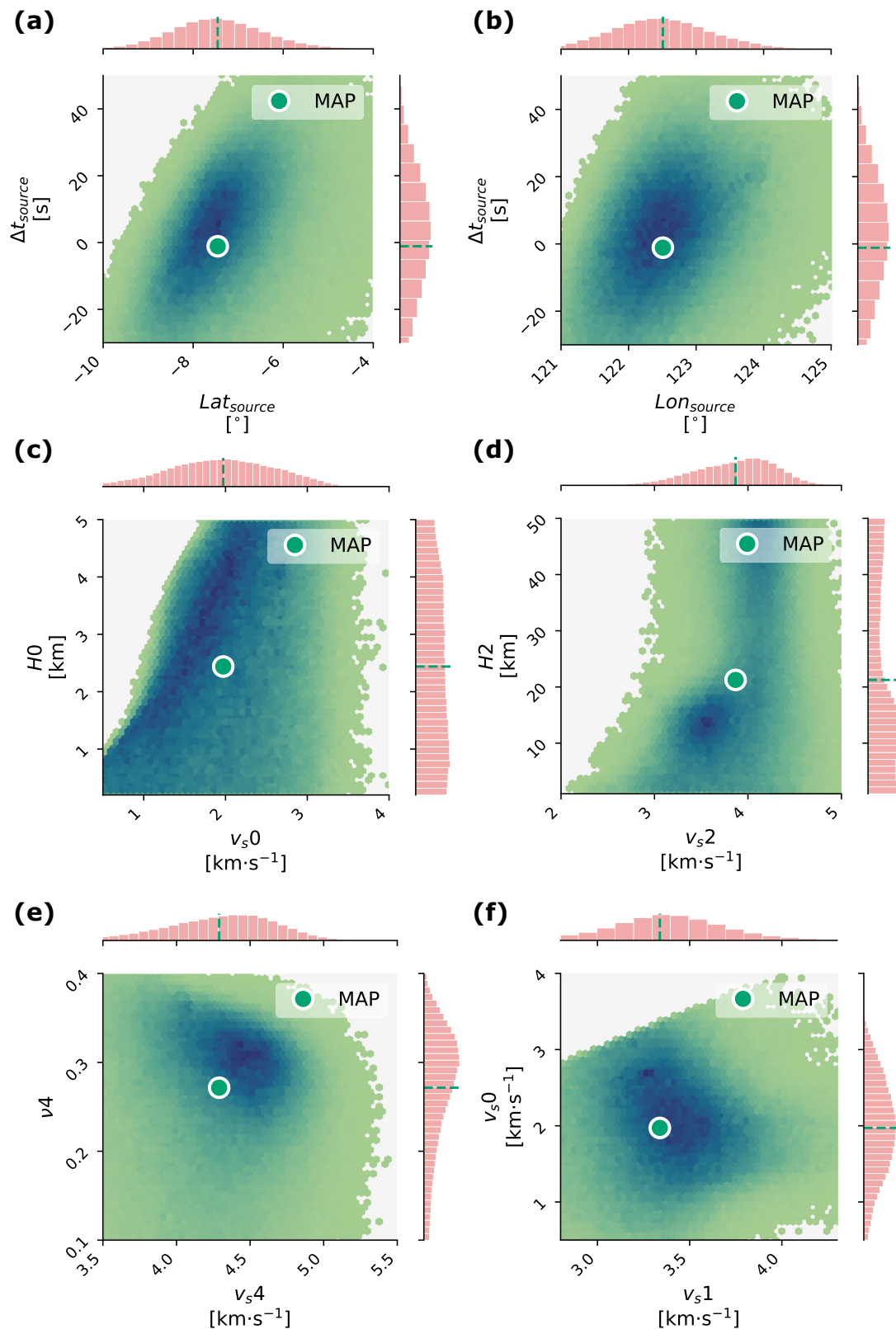
Upon selection of a source location, origin time, and subsurface model by the Monte Carlo sampling algorithm, the forward model is in charge of predicting the arrival time of waves at each of the station/balloons following Eq. (1). The travel times  $\Delta t_P$  and  $\Delta t_S$  of P and S waves are calculated using a ray-tracing method derived from the `LAUFZE` suite<sup>53</sup>. This Fortran routine takes in a source-receiver distance and a layered subsurface model and in return predicts the arrival time of the fastest direct P and S body waves.

The travel times  $\Delta t_{LR}(f)$  of LRs are calculated using a NumPy-accelerated Python implementation of the `surf96` code<sup>54</sup>, called `disba`<sup>55</sup>. The code is given a layered subsurface model input and outputs the group velocity  $v_g(f)$  of the Rayleigh waves at the fundamental or higher modes. We obtain the travel time by the approximation:

$$\Delta t_{LR}(f) = \frac{d_s}{v_g(f)}, \quad (2)$$

where  $d_s$  is the epicentral distance, considering only the fundamental mode.

The travel time  $\Delta t_{air}$  from the ground to the balloon is calculated by integrating the vertical variation of sound speed  $c_{air}(z)$  from  $z = 0$  to the



**Fig. 7 | Marginal probability density distributions resulting from the inversion of balloon data.** Distributions of **a** origin time and latitude, **b** origin time and longitude, **c** first layer thickness and shear wave velocity, **d** third layer thickness and

shear wave velocity, **e** Poisson's ratio and shear wave velocity in the fifth layer, and **f** shear wave velocities in the first and second layer. A darker hue represents a higher density of models.

balloon altitude  $z = z_b$ :

$$\Delta t_{\text{air}} = \int_{z=0}^{z_b} \frac{dz}{c_{\text{air}}(z)}. \quad (3)$$

The uncertainty of  $\Delta t_{\text{air}}$  primarily depends on the uncertainty of the atmospheric profile between 0 and about 20 km altitude. Considering a spatial and temporal grid around the Flores event, the variation of sound speed profile obtained from the G2S atmospheric specification<sup>56</sup>—itself based on the MERRA2<sup>57</sup> reanalysis—leads to variations of  $\Delta t_{\text{air}}$  smaller than 0.2 s (see Supplementary Fig. S9), therefore negligible compared to picking uncertainties.

Considering travel-time picks to have an uncorrelated Gaussian distribution with standard deviation  $\sigma$  (which is debatable, see, e.g., Husen and Hardebeck<sup>45</sup>), the log-likelihood function minimized by the Monte Carlo search is the sum of the following L2-norms:

$$\log P(\mathbf{d}|\mathbf{m}) = -\frac{1}{2} \sum_R \sum_i \frac{(t_{W_i,R}(\mathbf{m}) - t_{W_i,\text{obs}})^2}{\sigma_{W_i,R}^2} - \frac{1}{2} \sum_R \sum_i \log(2\pi\sigma_{W_i,R}^2), \quad (4)$$

where  $W_i$  represents the available phase picks among  $\{P, S, LR(f_i)\}$  and  $R$  are the available receivers.  $2\pi\sigma_{W_i,R}^2$  is the normalisation term for the Gaussian distribution of observations. An alternative formulation of the problem, which eliminates the origin time variable, seeks to minimize the misfit between the observed and predicted time-difference of arrival<sup>48</sup>. In this case, we use a reference P-wave arrival,  $t_{\text{Pref}}$  to calculate the arrival time differences. The likelihood function then becomes:

$$\begin{aligned} \log P_{\text{TDOA}}(\mathbf{d}|\mathbf{m}) &= -\frac{1}{2} \sum_R \sum_i \frac{(|t_{W_i,R}(\mathbf{m}) - t_{\text{Pref}}(\mathbf{m})| - |t_{W_i,\text{obs}} - t_{\text{Pref,obs}}|)^2}{\sigma_{W_i,R}^2 + \sigma_{\text{Pref}}^2} \\ &\quad - \frac{1}{2} \sum_R \sum_i \log(2\pi(\sigma_{W_i,R}^2 + \sigma_{\text{Pref}}^2)). \end{aligned} \quad (5)$$

If an L1-norm is preferred, Eq. (4) becomes:

$$\log P_{\text{L1}}(\mathbf{d}|\mathbf{m}) = -\sum_R \sum_i \frac{|t_{W_i,R}(\mathbf{m}) - t_{W_i,\text{obs}}|}{\sigma_{W_i,R}} - \sum_R \sum_i \log(2\sigma_{W_i,R}). \quad (6)$$

This formulation assumes that observations follows a Laplace distribution.

### Effects of balloon motion on arrival times

Contrary to a seismic station, a balloon is a non-stationary object, animated with a horizontal motion due to jet winds, and a oscillatory vertical motion due to buoyancy. Due to the horizontal balloon motion, the station-balloon distance is not constant over the duration of the earthquake signal and can impact  $\Delta t_W$  (see Eq. (1)). Assuming that the balloon is located at the distance  $d_0$  from the source at time  $t_s$ , and that, in the worst-case scenario, it is moving with velocity  $v_b$  in the radial direction away or towards the source; and considering a homogeneous media of seismic velocity  $v_W$  for simplicity, the expression of  $\Delta t_W$  becomes:

$$\Delta t_W = \frac{d_0 + v_b \Delta t_W}{v_W}, \quad (7)$$

or

$$\Delta t_W = \left( \frac{1}{1 - v_b/v_W} \right) \frac{d_0}{v_W}. \quad (8)$$

Strateole-2 balloons have a horizontal velocity of 5 to 8 m s<sup>-1</sup>. This means a ~0.1% change in travel time for P waves, and ~0.4% for S and LR waves

compared to a stationary receiver. This Doppler shift can thus be neglected compared to other sources of errors in travel time estimations<sup>31</sup>.

In the same way, constant-volume balloons like the Strateole-2 aerostat experience vertical motion, caused by wind perturbation. The stratification of the atmosphere, with density decreasing with altitude, exerts a restoring force through the volume of air displaced by the balloon. This leads to buoyancy oscillations, whose period depends on the Brunt–Väisälä pulsation  $N$  at the balloon equilibrium altitude, namely  $2\pi f_0 = N = \sqrt{-\frac{d\rho}{dz} \frac{g}{\rho_e}}$ , with  $g$  the constant of gravity and  $\rho_e$  the density at the balloon equilibrium altitude<sup>33</sup>. In the case of Strateole-2 balloons, this oscillation has a period between 180 and 240 s and an amplitude of 10–100 m, corresponding to about ~0.5 m s<sup>-1</sup>. This speed is insufficient to produce any notable effect on arrival time or travel time estimations. However, it is responsible for a strong low-frequency noise in the balloon pressure recordings, as a variation of 10–100 Pa is expected at each oscillation. In section “Balloon noise correction”, we describe a method we applied to mitigate this noise.

### Inversion priors

**Source location.** We have assumed no prior information on the epicenter location or source depth. Thus, we set uniform prior bounds of  $[-90^\circ, +90^\circ]$  for  $lat_s$  and  $[-180^\circ, +180^\circ]$  for  $lon_s$ . For the practical examples of this article, for which the epicenter is known a priori from earthquake catalogs, we simply restrict the starting latitudes and longitudes of the MCMC chains to a range  $\pm 20^\circ$  closer to the known epicenter, so as to avoid stuck chains and speed up the computation. The source depth is also considered unknown. We thus chose  $[0, 200]$  km as uniform prior bounds for  $h_s$  (see Supplementary Table S3).

**Source origin time.** The choice of prior bounds for the origin time is strongly dependent on the choice of reference time for arrival time picking. In the practical examples of this article, the chosen reference time is the USGS published origin time<sup>40</sup>, and we set prior bounds for  $t_s$  to  $[-200, +200]$  s. In practice, a rough approximation of the origin time can be calculated by estimating the minimum and maximum possible source distance using prior bounds for seismic velocities, leading to ranges for  $t_s$  closer to thousands of seconds.

**Layers and layer thickness.** We choose to parameterize our subsurface as a succession of homogeneous layers, as this parametrization is best adapted to the numerical methods (`disba`, `LAUFZE`) used in our forward model. The maximum source-receiver distance in our two inversions is about ~3000 km, a distance at which body waves have turning depths of ~600 km on Earth. It is therefore necessary to parameterize subsurface down to mantle depths.

In this study, the number of layers is fixed to six, in addition to an underlying halfspace. Tests of the effect of the number of layers on the achieved misfit showed that the misfit does not decrease notably for a higher number of layers (see Supplementary Fig. S10). The last layer and the halfspace are intended to represent upper-mantle velocities, hence having large prior thickness between 100 and 400 km. The uppermost layers represent a possible sedimentary region with thickness of 0.2 to 5 km. The remaining layers have intermediate prior thicknesses, allowing for variations within the crust, and can be found in Supplementary Table S3.

**Seismic velocities and Poisson’s ratio.** The inversion covers seismic velocities from the upper crust to the upper mantle. The thin top layer allows for possible sedimentary deposits and has prior bounds for  $v_s$  of  $[0.5, 4]$  km s<sup>-1</sup>. The following four layers correspond to crustal or upper mantle materials and have  $v_s$  within  $[1, 6]$  to  $[3, 6]$  km s<sup>-1</sup>. The last layer and the halfspace are mantle layers with  $v_s$  within  $[4, 7]$  km s<sup>-1</sup>.  $v_p$  is calculated using the values of  $v_s$  and of the Poisson’s ratio  $\nu$ . Prior bounds for  $\nu$  are uniform within a range of  $[0.1, 0.4]$  encompassing typical properties for crustal and mantle minerals<sup>44</sup>. In addition to these uniform bounds for Poisson’s ratio and shear wave velocity, we implement

additional rules to restrict the acceptable prior models. Prior models must have no negative velocity gradient in the first 3 layers. Below that, negative changes in velocity are limited to  $\Delta v_s, \Delta v_p < -1 \text{ km s}^{-1}$ . An upper limit of  $v_p < 12 \text{ km s}^{-1}$  is set, which has an influence on the prior distribution of  $v$ . The distribution of prior models can be found in the Supplementary Information (Fig. S11).

The density, to which P, S, and LR travel times are less sensitive than seismic velocities, is not inverted but rather modeled using Birch's empirical law<sup>58</sup>.

### Balloon noise correction

To improve the low-frequency SNR of the balloon pressure trace, the balloon buoyancy resonance<sup>33</sup> is corrected following a method similar to<sup>43</sup>. The GPS altitude trace  $Z$  of each Strateole-2 balloon is upsampled and interpolated to  $Z_{\text{up}}$ , so as to match the sampling rate of 1 Hz and exact timestamps of the pressure trace  $P$ , using a Hann taper in the frequency domain. Over small variations in altitude, the relation between pressure and altitude is quasi-linear. A sliding window of 500 s is run along the  $P$  and  $Z_{\text{up}}$  traces, and a linear regression is applied to determine the coefficients of their relation, valid for the center point of the window. These are then used to produce an auxiliary pressure trace  $P_{\text{mod}}$ , calculated from  $Z_{\text{up}}$ . Finally, the corrected pressure trace  $P_{\text{corr}}$  is obtained from  $P_{\text{corr}} = P - P_{\text{mod}}$ . The different traces and steps of the correction can be found in Supplementary Fig. S3. This processing step helps partially correct the balloon buoyancy oscillations and improves subsequent frequency-time analysis.

### Data processing and arrival picking

Key to the inversion framework is to properly identify and pick seismic arrival times. An infrasound signal is by nature single-component. Hence, classical techniques for separating P, S, and LRs in 3-component seismic signals based on polarity cannot be used. Instead, we leverage other aspects of these arrivals, namely the impulsive nature of body waves and their envelopes and the dispersive nature of Rayleigh waves, distinguishable in the time-frequency domain.

To pick P and S waves, a two-step method is used. First, the signal is filtered in several frequency bands and its envelope is calculated using a Hilbert transform. For balloon signals, we use the envelope of the low-passed signal below 0.1 Hz, the high-passed signal above 0.05 Hz, and an intermediate signal band-passed between 0.03 and 0.1 Hz. For one-component seismic velocity signals, the signal is first low-passed or high-passed at 1 Hz, or band-passed between 0.02 and 0.8 Hz. Part of the scattering in the envelopes is smoothed by calculating a sliding median over the 5, 10, and 20 s preceding each considered point in time. Using multiple sliding window sizes helps rule out picks in the envelope that could be due to a local scattered arrival. Using this envelope method, a first hypothesis on the arrival of P and S waves and their uncertainty can be made by identifying the start of the P and the S energy envelope. Then, in a second step, these picks are assessed in narrower frequency bands. We construct a filter bank by bandpass-filtering the single-component signal in 10 narrow logarithmic intervals from  $\sim 0.001$  Hz to the Nyquist frequency of the signal. This method presents advantages for refining the P wave pick, more clearly visible at high frequency, and for confirming the S-wave pick, by identifying a later impulsive arrival spanning multiple intervals of frequency. Despite these two steps, the S wave arrival is only identified with a very large uncertainty in some cases.

To pick the Rayleigh wave arrivals, we apply a Frequency-Time ANalysis (FTAN) to the single component signal using the Stockwell transform (also named S-transform), an approach analogous to a Morlet wavelet transform. The LR is identified by its dispersion, and arrival times are picked at different frequencies around the maximum of the dispersed signal. Wavelet or S-transforms optimize the trade-off between time and frequency resolution in FTAN, but arrivals retain a frequency-dependent spread in time, which we interpret as our uncertainty in arrival time.

### Post processing of inversion results

MCMC inversions return a large amount of model parameter samples, out of which several statistically meaningful metrics should be extracted. In the Bayesian framework, we are interested in the most probable model given our data and prior, i.e., the model with the maximum posterior probability, referred to as MAP. Although a Monte Carlo inversion returns millions of samples, the curse of dimensionality means that the MAP is not necessarily among them. To determine an estimate of the MAP out of all our samples, we use the Mean-shift method, through which subset of  $2 \times 10^4$  samples are migrated towards one or more regions of high density in the posterior space. The MAP values calculated for the three main inversions in this study are reported in Supplementary Table S4.

The MAP yields the region of high density throughout all dimensions. We are also interested in the behavior of individual or groups of parameters in lower dimensions. This is done by considering marginal distributions of parameters through histograms or density plots, as was done in the majority of figures throughout this article. In some cases, we apply additional processing to the marginal distribution in order to enhance visualization or enable easier interpretation. The marginal density distribution of subsurface parameters of Figs. 2 and 5 corresponds to a 2D histogram of subsurface models, normalized by a Mean value. This Mean value represents the histogram counts expected if subsurface models were uniformly distributed on the plot area. Thus, the Posterior/Mean scale allows to compare how densely packed posterior models are between two distinct inversions, independently from the inversion parameters.

We apply a similar process to interpret the layer thickness distribution in the posterior models. The posterior distribution of layer thickness can be transformed into a posterior distribution of interface depth using a cumulative summation starting from the top layer. However, comparing this distribution of interface depth to a uniform distribution, for example using a simple histogram, can be misleading, as the sum of uniform distribution from which the layer thicknesses were picked is not uniform itself. Here, we instead calculate the ratio of the number layer thicknesses counted in one bin in the posterior distribution, to the number predicted in a cumulative prior distribution. For each layer  $N$ , the cumulative prior is defined as the cumulative posterior distribution of interface depths for  $k < N$ , summed with the prior distribution of thickness for layer  $N$ :

$$\text{Prior}(d_N) = \sum_{k=0}^{N-1} \text{Posterior}(h_k) + \text{Prior}(h_N), \quad (9)$$

where  $d_N$  is the depth of interface  $N$ . Mixing the prior and posterior distribution in this cumulative prior allows to rule out the effect of above layers on the distribution of lower layers. The final interface count ratio is obtained from  $\sim 5 \times 10^5$  samples from the posterior and from each prior. An interface count ratio greater than one thus means that there is a higher probability that an interface is located at this depth than would be expected from a cumulated prior distribution.

### Data availability

Seismic and ground infrasound data is retrieved from the International Federation of Digital Seismograph Networks (FDSN) (GEOFON: <http://geofon.gfz-potsdam.de/doi/network/GE>, Australian National Seismograph Network <http://pid.geoscience.gov.au/dataset/ga/144675>, Global Seismograph Network: <https://www.fdsn.org/networks/detail/IU/>). The Strateole-2 TSEN data are available at <https://doi.org/10.14768/c417e612-015d-4812-9b59-294a6570c7c3>. The times and uncertainties of picks used in the seismic and balloon inversions are available on Zenodo<sup>59</sup> and at <https://github.com/m-froment/balloon-inv/tree/master/Picks/>.

### Code availability

The data analysis and inversion software developed for this study was written in the Python and Fortran languages. It is available on Zenodo<sup>59</sup> and at <https://github.com/m-froment/balloon-inv>.

Received: 10 June 2025; Accepted: 15 October 2025;  
Published online: 21 November 2025

## References

- Rolf, T. et al. Dynamics and evolution of Venus' mantle through time. *Space Sci. Rev.* **218**, 1–51 (2022).
- Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
- Kennett, B. L. N. & Engdahl, E. R. Traveltimes for global earthquake location and phase identification. *Geophys. J. Int.* **105**, 429–465 (1991).
- Kustowski, B., Ekström, G. & Dziewoński, A. M. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model. *J. Geophys. Res. Solid Earth* **113**, B06306 (2008).
- Tromp, J. Seismic wavefield imaging of Earth's interior across scales. *Nat. Rev. Earth Environ.* **1**, 40–53 (2020).
- Berg, E. M. et al. Shear velocity model of Alaska via joint inversion of Rayleigh wave ellipticity, phase velocities, and receiver functions across the Alaska Transportable Array. *J. Geophys. Res. Solid Earth* **125**, e2019JB018582 (2020).
- Latham, G., Ewing, M., Press, F. & Sutton, G. The Apollo passive seismic experiment. *Science* **165**, 241–250 (1969).
- Banerdt, W. B. et al. Initial results from the InSight mission on Mars. *Nat. Geosci.* **13**, 183–189 (2020).
- Lorenz, R. D. et al. Titan seismology with Dragonfly: Probing the internal structure of the most accessible ocean world. In *Proc. 50th Lunar and Planetary Science Conference*, LPI Contribution No. 2132, 2173 (The Woodlands, Texas, 2019).
- Panning, M. P. et al. Seismology on Titan: a seismic signal and noise budget in preparation for Dragonfly. In *Proc. SEG International Exposition and Annual Meeting (OnePetro)*, 2020.
- Lorenz, R. D., Shiraishi, H., Panning, M. & Sotzen, K. Wind and surface roughness considerations for seismic instrumentation on a relocatable lander for Titan. *Planet. Space Sci.* **206**, 105320 (2021).
- Stevenson, D. J. et al. Probing the interior structure of Venus. <https://resolver.caltech.edu/CaltechAUTHORS:20150727-150921873> (2015).
- Garcia, R. F. et al. Seismic wave detectability on Venus using ground deformation sensors, infrasound sensors on balloons and airglow imagers. *Earth Space Sci.* **11**, e2024EA003670 (2024).
- Krishnamoorthy, S. et al. Detection of artificially generated seismic signals using balloon-borne infrasound sensors. *Geophys. Res. Lett.* **45**, 3393–3403 (2018).
- Krishnamoorthy, S. et al. Aerial seismology using balloon-based barometers. *IEEE Trans. Geosci. Remote Sens.* **57**, 10191–10201 (2019).
- Garcia, R. F. et al. An active source seismo-acoustic experiment using tethered balloons to validate instrument concepts and modelling tools for atmospheric seismology. *Geophys. J. Int.* **225**, 186–199 (2021).
- Mutschlecner, J. P. & Whitaker, R. W. Infrasound from earthquakes. *J. Geophys. Res. Atmos.* **110**, D01108 (2005).
- Brissaud, Q., Martin, R., Garcia, R. F. & Komatitsch, D. Hybrid Galerkin numerical modelling of elastodynamics and compressible Navier–Stokes couplings: applications to seismo-gravito acoustic waves. *Geophys. J. Int.* **210**, 1047–1069 (2017).
- Brissaud, Q. et al. The first detection of an earthquake from a balloon using its acoustic signature. *Geophys. Res. Lett.* **48**, e2021GL093013 (2021).
- Lognonné, P. & Johnson, C. L. 10.03 – Planetary seismology. In *Treatise on Geophysics* 2nd edn, (ed. Schubert, G.) 65–120 (Elsevier, Oxford, 2015).
- Averbuch, G., Houston, R. & Petculescu, A. Seismo-acoustic coupling in the deep atmosphere of Venus. *J. Acoust. Soc. Am.* **153**, 1802–1810 (2023).
- Didion, A. et al. Remote sensing of Venusian seismic activity with a small spacecraft, the VAMOS mission concept. In *Proc. 2018 IEEE Aerospace Conference*, 1–14 (IEEE, 2018) <https://ieeexplore.ieee.org/document/8396447>.
- Sutin, B. M. et al. VAMOS: A SmallSat mission concept for remote sensing of Venusian seismic activity from orbit. In *Proc. Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, 10698, 1651–1670 (SPIE, 2018).
- Linkin, V. M. et al. VEGA balloon dynamics and vertical winds in the Venus middle cloud region. *Science* **231**, 1417–1419 (1986).
- Hall, J. L. et al. Altitude-controlled light gas balloons for Venus and Titan exploration. In *Proc. AIAA Aviation 2019 Forum* (American Institute of Aeronautics and Astronautics, Dallas, Texas, 2019).
- Bellemare, M. G. et al. Autonomous navigation of stratospheric balloons using reinforcement learning. *Nature* **588**, 77–82 (2020).
- Yoder, C. D., Agrawal, S., Motes, A. G. & Mazzoleni, A. P. Aerodynamic tethered sails for scientific balloon trajectory control: Small-scale experimental demonstration. *J. Aircraft* **58**, 1010–1021 (2021).
- Bowman, D. C. & Krishnamoorthy, S. Infrasound from a buried chemical explosion recorded on a balloon in the lower stratosphere. *Geophys. Res. Lett.* **48**, e2021GL094861 (2021).
- Garcia, R. F. et al. Infrasound from large earthquakes recorded on a network of balloons in the stratosphere. *Geophys. Res. Lett.* **49**, e2022GL098844 (2022).
- Haase, J. et al. *Around the World in 84 Days*. <http://eos.org/science-updates/around-the-world-in-84-days> (Eos, American Geophysical Union, 2018).
- Gerier, S., Garcia, R. F., Martin, R. & Hertzog, A. Forward modeling of quake's infrasound recorded in the stratosphere on board balloon platforms. *Earth Planets Space* **76**, 87 (2024).
- Macpherson, K. A., Fee, D., Coffey, J. R. & Witsil, A. J. Using local infrasound to estimate seismic velocity and earthquake magnitudes. *Bull. Seismol. Soc. Am.* **113**, 1434–1456 (2023).
- Massman, W. J. On the nature of vertical oscillations of constant volume balloons. *J. Appl. Meteorol.* **17**, 1351–1356 (1978).
- Arrowsmith, S., Park, J., Che, I.-Y., Stump, B. & Averbuch, G. Event location with sparse data: When probabilistic global search is important. *Seismol. Res. Lett.* **92**, 976–985 (2020).
- Thurber, C. H. Hypocenter-velocity structure coupling in local earthquake tomography. *Phys. Earth Planet. Inter.* **75**, 55–62 (1992).
- Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005).
- Goodman, J. & Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **5**, 65–80 (2010).
- Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. Emcee: The MCMC hammer. *Publ. Astron. Soc. Pac.* **125**, 306–312 (2013).
- Spendi, P. et al. The Kalaotoa fault: a newly identified fault that generated the Mw 7.3 Flores Sea earthquake. *Seismic Rec.* **2**, 176–185 (2022).
- International Seismological Center. On-Line Bulletin (2025).
- Laske, G., Masters, G., Ma, Z. & Pasyanos, M. Update on CRUST1.0 – A 1-degree global model of Earth's crust. In *Proc. EGU General Assembly*, Vol. 15, EGU2013–2658 (Copernicus Publications, 2013).
- Simmons, N. A., Myers, S. C., Johannesson, G., Matzel, E. & Grand, S. P. Evidence for long-lived subduction of an ancient tectonic plate beneath the southern Indian Ocean. *Geophys. Res. Lett.* **42**, 9270–9278 (2015).
- Podglajen, A. et al. Stratospheric balloon observations of infrasound waves from the 15 January 2022 Hunga eruption, Tonga. *Geophys. Res. Lett.* **49**, e2022GL100833 (2022).
- Christensen, N. I. Poisson's ratio and crustal seismology. *J. Geophys. Res. Solid Earth* **101**, 3139–3156 (1996).

45. Husen, S. & Hardebeck, J. *Earthquake Location Accuracy* (Community Online Resource for Statistical Seismicity Analysis, 2010).
46. Helffrich, G. Topography of the transition zone seismic discontinuities. *Rev. Geophys.* **38**, 141–158 (2000).
47. Milne, J. Determination of earthquake origins. in *Earthquakes and Other Earth Movements* (The International Scientific Series, New York, 1886).
48. Zhou, H.-w Rapid three-dimensional hypocentral determination using a master station method. *J. Geophys. Res. Solid Earth* **99**, 15439–15455 (1994).
49. Sambridge, M. & Kennett, B. Seismic event location: nonlinear inversion using a neighbourhood algorithm. *Pure Appl. Geophys.* **158**, 241–257 (2001).
50. Bowman, D. C., Rouse, J. W., Krishnamoorthy, S. & Silber, E. A. Infrasonic direction of arrival determination using a balloon-borne aeroseismometer. *JASA Express Lett.* **2**, 054001 (2022).
51. Neal, R. M. MCMC Using Hamiltonian Dynamics. in *Handbook of Markov Chain Monte Carlo* (Chapman and Hall/CRC, 2011).
52. Sambridge, M. A parallel tempering algorithm for probabilistic sampling and multimodal optimization. *Geophys. J. Int.* **196**, 357–374 (2014).
53. Schweitzer, J. LAUFZE / LAUFPS. In *Proc. New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, (ed. Bormann, P.) 1–14 (Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012).
54. Herrmann, R. B. Computer programs in seismology: an evolving tool for instruction and research. *Seismol. Res. Lett.* **84**, 1081–1088 (2013).
55. Luu, K. Disba: Numba-accelerated computation of surface wave dispersion. Zenodo (2024).
56. Drob, D. P., Picone, J. M. & Garcés, M. Global morphology of infrasonic propagation. *J. Geophys. Res. Atmos.* **108**, 4680 (2003).
57. Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
58. Birch, F. Density and composition of mantle and core. *J. Geophys. Res.* **69**, 4377–4388 (1964).
59. Froment, M. m-froment/balloon-inv: Balloon Inversion code - Initial public release. Zenodo <https://github.com/m-froment/balloon-inv/tree/master/Picks/> (2025).

## Acknowledgements

We thank Daniel Bowman, Voon Hui Lai and an anonymous referee for their review and constructive feedback, which improved the quality of this paper. This work received support from the *Airborne Inversion of Rayleigh waves* (AIR) project, funded by the Research Council of Norway basic research program FRIPRO, Contract 335903, as well as NORSAR institute funding. The codes developed for this study makes use of open-source modules for seismology and MCMC inversions, such as ObsPy (<https://docs.obspy.org/>) and emcee (<https://emcee.readthedocs.io/en/stable/>). We thank their contributors for providing and maintaining them. The computations were performed on resources provided by Sigma2 – the National Infrastructure for High-Performance Computing and Data Storage in Norway (project NN11013K), and by the Educloud service platform at the University of Oslo (Fox HPC cluster). We are grateful to CNES and its partners for developing the Strateole-2 project and making their data available.

## Author contributions

Marouchka Froment contributed to the conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing (original draft and editing), and visualization required for this work. Quentin Brissaud contributed to the conceptualization, methodology, software, investigation, data curation, writing, as well as funding acquisition, and project administration. Sven Peter Näsholm contributed to the conceptualization, methodology, investigation, and writing. Johannes Schweitzer contributed software, writing, validation, and investigation. Quentin Brissaud and Sven Peter Näsholm jointly supervised this work.

## Competing interests

The authors declare no competing interests.

## Consent to publish

All authors have consented to the publication of this manuscript.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-025-02917-7>.

**Correspondence** and requests for materials should be addressed to Marouchka Froment.

**Peer review information** *Communications Earth and Environment* thanks Daniel Bowman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Luca Dal Zilio and Nandita Basu. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025