AFRL-RV-PS-
TR-2022-0082

AFRL-RV-PS-
TR-2022-0082

TOWARDS IMPROVED UNDERSTANDING OF SEISMIC MOMENT TENSORS: UTILIZING HIGH-RESOLUTION EARTH MODEL WITH ACCOMPANYING UNCERTAINTY WITHIN HIERARCHICAL BAYESIAN INVERSION FRAMEWORK

Hrvoje Tkalčić, et al.

Research School of Earth Sciences
The Australian National University
Canberra, ACT, Australia

10 August 2022

Final Report

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

DTIC COPY

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by AFMC/PA and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RV-PS-TR-2022-0082 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//	//SIGNED//
_____	_____
Dr. Megan P. Flanagan	For: Lt Col. Salvador Aleman, Acting Chief
Program Manager/AFRL/RVB	AFRL Geospace Technologies Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
10-08-2022	Final Report	09 Jan 2020 – 14 Mar 2022

4. TITLE AND SUBTITLE
Towards Improved Understanding of Seismic Moment Tensors: Utilizing High-Resolution Earth Model with Accompanying Uncertainty Within Hierarchical Bayesian Inversion Framework

5a. CONTRACT NUMBER
FA9453-20-C-0072

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
C6601F

6. AUTHOR(S)
Hrvoje Tkalčić, Thanh-Son Phạm, Jinyin Hu, and Seongryong Kim

5d. PROJECT NUMBER
1010

5e. TASK NUMBER

5f. WORK UNIT NUMBER
V1VW

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research School of Earth Sciences
The Australian National University
Canberra, ACT, Australia

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL/RVBN

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Space Vehicles Directorate
3550 Aberdeen Avenue SE
Kirtland AFB, NM 87117-5776

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RV-PS-TR-2022-0082

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited (AFMC-2022-4340 dtd 13 Sep 2022).

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The relative significance of the isotropic component in a moment tensor (MT) solution is critical in discriminating explosive events. For critically shallow sources, such as manmade explosions, it is challenging to resolve the isotropic (ISO) from vertical compensated linear vertical dipole (CLVD) components, because the focal mechanisms generate highly similar waveforms at regional distances. Additionally, rigorous quantification of uncertainty of data noise, mainly due to the observational process, and theoretical error, primarily due to imperfect knowledge of Earth structure, remains at the forefront of MT inversion research. Here, we attempt to progress on these research fronts. First we investigate an inversion scheme where data noise and structural uncertainty, empirically estimated for 1D Earth models, are constructed into covariance matrices defining the likelihood function (Phạm & Tkalčić, 2021). Second, we explore the use of the affine-invariant ensemble samplers, which prove efficient in exploring the mechanism trade-off for shallow source depths. Also in this attempt, to account for Earth's model uncertainty, we treat the time shifts between observed and synthetic waveforms as a proxy for heterogeneity of the Earth along propagation paths. Third, we explore a representation of 3D earth structure uncertainty in the form of empirical covariance matrices. In all cases, we demonstrate the importance of properly treating Earth's structure uncertainty in MT inversion, but the intrinsic trade-off in shallow seismic sources remains to be further investigated.

15. SUBJECT TERMS
moment tensor inversion

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Megan P. Flanagan
Unclassified	Unclassified	Unclassified	Unlimited	130	**19b. TELEPHONE NUMBER** *(include area code)*

This page is intentionally left blank.

Approved for public release; distribution is unlimited.

TABLE OF CONTENTS

Section **Page**

1 Summary 1

2 Toward Improving Point-Source Moment-Tensor Inference by Incorporating 1D Earth Model's Uncertainty: Implications for the Long Valley Caldera Earthquakes 2

Summary 2

Plain Language Summary 2

2.1 Introduction 2

2.2 Theoretical background 4

2.3 Methodology 7
2.3.1 Parameterization of seismic moment tensors 7
2.3.2 Forward simulation of full waveforms 9
2.3.3 Data-noise covariance matrix and noise-magnitude hyper-parameters 10
2.3.4 Structure-uncertainty covariance matrix 12
2.3.5 Iterative seismic inversion scheme 13

2.4 Synthetic Experiments 13
2.4.1 Experiment setup 13
2.4.2 True theory covariance matrices and true data noise levels 15
2.4.3 True theory noise covariance matrices and unknown data noise levels 19
2.4.4 Unknown theory covariance matrices 22

2.5 Implications for a volcanic earthquake in Long Valley Caldera 25
2.5.1 Data Preparation 25
2.5.2 Moment-tensor inversion results 26
2.5.3 Short-period waveform examples 31

2.6 Discussions 31

2.7 Concluding remarks 32

3 Full moment tensors inversion using a Bayesian hierarchical approach with 2D structure uncertainty: Implications for DPRK tests 2009-2017 33

Summary 33

3.1 Introduction 33

3.2 Methodology 34
3.2.1 Bayesian inference of seismic moment tensors 34
3.2.2 Accounting for Earth model uncertainty by station-specific time shifts 36
3.2.3 Exploring the parameter space using affine invariant ensemble samplers 37

3.3 Synthetic Experiments 39
3.3.1 Experiment configuration 40
3.3.2 The inherent ambiguity of non-double-couple source types for shallow seismic sources 41
3.3.3 Inversion results from a shallow synthetic explosion 44

3.4 Application to DPRK nuclear tests 51
3.4.1 Data preparation 51
3.4.2 MT inversion results of DPRK 2009-2017 tests 52
3.4.3 Robustness of the MT inversion 56

TABLE OF CONTENTS (Continued)

Page

3.5 Discussion *58*
3.5.1 The similarity of MT solutions 58
3.5.2 Consistency of time shifts and 2D structure. 59
3.6 The effect of the uncorrelated noise assumption *60*
3.6.1 Uncertainty of MTs for shallow explosions 61
3.7 Conclusions *63*

4 Uncertainties of 3D Earth models in moment tensor inversion of shallow seismic sources 64

4.1 The resolvability of full moment-tensor for shallow nuclear explosions *64*
4.1.1 Motivation 64
4.1.2 Computation of 3D Green's function using the reciprocity principle 65
4.1.3 Resolvability of shallow seismic moment tensors 70
4.2 An approach to perturb 3D Earth structure *73*

Appendix Supporting information for Section 2 76

A.1 Derivation of combined covariance matrix *76*
A.2 The least-squares solution *78*
A.3 Bayesian inference for hierarchical inversion *79*

REFERENCES 102

LIST OF SYMBOLS, ABBREVIATIONS, and ACRONYMS 107

LIST OF FIGURES

Figure **Page**

Figure 1. *(a)* Various Earth models used in synthetic experiments and real data MT inversion in this study. *(b)* Waveforms filtered between 20 and 50 s, corresponding to the Earth models in a) for a common MT. *(c)* Similar to b) but waveforms are filtered between 10 and 50 s. *(d)* like b), but waveforms are filtered between 6 and 50 s. 5

Figure 2. Visualization of covariance matrices for each row corresponding to one of the three recording velocity components for station BKS: *(a)* vertical, *(b)* radial, and *(c)* tangential. ...11

Figure 3. Network configuration for the MT inverse problem. 14

Figure 4. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels……16

Figure 5. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels…… 18

Figure 6. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, h, are hyper-parameters to be recovered in the inversion. 20

Figure 7. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting for and not accounting for the structural uncertainty and unknown data noise levels.. 21

Figure 8. Evolution of McMC chains the iterative inversion scheme. 22

Figure 9. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration…….23

Figure 10. Systematic reduction of max-norm distance (defined in Equation (20)) from the mean MT solution in each iteration to the true MT input for three-period bands: *(a)* 6–50 s, *(b)* 10–50 s, *(c)* 20–50 s. 24

Figure 11. Comparison of inversion results in the two scenarios for real data inversion when accounting for and not accounting for structural uncertainty and unknown data noise levels.26

Figure 12. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme for the real data, which is the mean of the post-burning chain in the fifth iteration. 27

Figure 13. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. 29

Figure 14. Summary of sensitivity tests of MT solutions of the LV2 event considering (*a*) levels of structural uncertainty and (*b*) event depths. 30

Figure 15. Waveforms for different 1D Earth structures. 37

Figure 16. Schematic demonstration in two-dimensional parameter space of the stretched move used in the affine invariant McMC (Goodman & Weare, 2010)……...38

LIST OF FIGURES (Continued)

Page

Figure 17. Comparison of sampling efficacy between (left panel) the traditional Metropolis-Hasting method and (right panel) the ensemble samplers with stretched moves (Goodman & Weare, 2010)…… 39

Figure 18. Map of region showing the DPRK2017 test location (red star) and seven seismic stations (blue triangles).. 41

Figure 19. A waveform comparison of critically shallow isotropic and vertical CLVD moment-tensors recorded by regional stations (< 1200 km).. 42

Figure 20. The ambiguity of non-double-couple components of shallow seismic sources. 43

Figure 21. The variety of seismic MTs producing almost identical seismic waveforms…….44

Figure 22. MT inversion results considering uncorrelated data noise and 2D structural error with a hierarchical Bayesian framework for the synthetic scenario. 46

Figure 23. Recovered station-specific noise parameters.....47

Figure 24. Recovered station-specific time shifts. 48

Figure 25. MT decomposition and waveform fit at seven stations for the synthetic scenario ..49

Figure 26. Lune diagram of source type for recovered MT solutions in the scenarios of different signal-to-noise ratio (SNR): (a) SNR=30, (b) SNR=25, (c) SNR=20, (d) SNR=15, (e) SNR=10, and (f) SNR=5. 50

Figure 27. Lune diagram of source type for MT solutions recovered when the true source is at a different depth. 50

Figure 28. Lune diagram of source type for MT solutions recovered for different input source types. 51

Figure 29. Lune diagram of source type for MT solutions of five DPRK tests in 2009-2017.. 53

Figure 30. Fits between observed (black) and predicted (red) waveforms for five DPRK 2009-2017 explosions……54

Figure 31. Each station's logarithm of likelihood (Equation (30)) was calculated during MT inversion for the DRPK 2017 test.. 55

Figure 32. Lune diagrams of source types of MT solutions for the DPRK2017 test assuming the various source depths at 0.5km, 1.0km, 1.5km, 2.0km, and 2.5km in panel (a) - (e), respectively.. 56

Figure 33. MT Solutions were obtained by hierarchical Bayesian inversion considering uncorrelated noise and 2D structural error when removing two stations in Japan……...57

Figure 34. Same as Figure 33, except MT solutions were obtained by hierarchical Bayesian inversion considering uncorrelated noise and 2D structural error using another dataset. 58

Figure 35. Comparison for 3-component waveforms (displacement) at five standard stations for five DPRK 2009-2017 tests.....59

Figure 36. Pairwise correlation coefficients of the same components at MDJ station for five DPRK explosions from 2009 to 2017. 59

Figure 37. The mean time shift at each station. 60

LIST OF FIGURES (Continued)

Page

Figure 38. Lune diagram of source type of MT solutions when fixing the noise strength to the same as pre-event ambient noise for five DPRK tests from 2009-2017 with panel (a) - (e), respectively. 61

Figure 39. Fits between observed (black) and predicted waveforms (color-coded line) obtained from 5 previous studies and this study for DPRK 2017 explosion. 62

Figure 40. Six elementary basis mechanisms in Equation (34). (b) Plot of forces for the first elementary basis M1. 65

Figure 41. Principle of reciprocity…….. 67

Figure 42. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.05 Hz. 67

Figure 43. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.10 Hz. 68

Figure 44. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.15 Hz…….69

Figure 45. (left) Condition number of GtG as function of depth and frequency (Equation (38)). (right) Balance of amplitudes among elementary Green's Functions 71

Figure 46. Least-squares solutions for the seismic source at 0.5 km. 72

Figure 47. Least-squares solutions for the seismic source at 3.5 km. 73

Figure 48. A method to perturb 3D Earth models. 74

Figure 49. Green's functions corresponding to Monte-Carlo perturbed 3D Earth models using the method presented in Figure 48. 75

Figure 50. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels…….....82

Figure 51. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels.....83

Figure 52. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels…….84

Figure 53. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels. 85

Figure 54. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, h, are hyper-parameters to be recovered in the inversion. 86

Figure 55. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, h, are hyper-parameters to be recovered in the inversion. 87

Figure 56. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels……..88

LIST OF FIGURES (Continued)

Page

Figure 57. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels……...……89

Figure 58. Evolution of McMC chains the iterative inversion scheme……..…..90

Figure 59. Evolution of McMC chains of the iterative inversion scheme. 91

Figure 60. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration……...92

Figure 61. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. 93

Figure 62. Preliminary MT solutions by hierarchical Bayesian inversion and least-squares inversion.....94

Figure 63. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels……..95

Figure 64. Comparison of inversion results in the two scenarios for real data inversion when accounting for and not accounting for structural uncertainty and unknown data noise levels.96

Figure 65. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels. 97

Figure 66. Evolution of McMC chains in the iterative inversion scheme. 98

Figure 67. MT decomposition and waveform fit corresponding the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. ……99

Figure 68. Evolution of McMC chains in the iterative inversion scheme. 100

Figure 69. MT decomposition and waveform fit corresponding the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. ……101

LIST OF TABLES

Table **Page**

Table 1. True station-specific time shifts (unit: seconds) were used for the numerical experiment of MT inversion for the DPRK2017 test.....41

Table 2. Recovered station-specific time shifts (in seconds) for five DPRK 2009-2017 tests. For the 2013 test, used two stations in South Korea are CHNB and YNCB 52

Table 3. Glossary of Symbols…….81

This page is intentionally left blank.

1 SUMMARY

Title of Project: Towards Improved Understanding of Seismic Moment Tensors: Utilizing High-Resolution Earth Model with Accompanying Uncertainty Within Hierarchical Bayesian Inversion Framework

Principle Investigator: Prof. Hrvoje Tkalčić

Co-PI: Assoc. Prof. Seongryong Kim

Postdoctoral Researcher: Dr. Thanh-Son Pham

Graduate student: Jinyin Hu

Synopsis of Research: The main objective of this method is to develop inversion methods in the Bayesian framework to resolve highly non-double-couple seismic sources with appropriate accounting for both data noise and errors due to Earth models. Data noise is the error encountered in the acquisition process, primarily caused by background seismic noise at the recording sites. On the other hand, theory error is caused by theoretical approximations made in the physical model of the forward wave simulation solver. It is reasonable to assume that the error due to the imperfect representation of the actual Earth by a structural model significantly contributes to the theory error. In this project, we investigate two methods to characterize the structural uncertainty for this purpose:

- Firstly, we developed a method to construct empirical structural covariance matrices using a large ensemble of pre-computed waveforms corresponding to Earth models slightly perturbed from a 1D Earth model. This methodological framework helped demonstrate the importance of structural uncertainties when resolving a highly non-double-couple source of settings for volcanic earthquakes.
- Secondly, we develop a joint inversion scheme for MT solution and time shifts of individual stations. The station-specific time shifts account for the discrepancy between a structural model and actual Earth structures along the propagation paths. Besides, through numerical experiments, we describe the challenges of using regional full-waveform inversion to invert for focal mechanisms of critically shallow seismic sources, such as nuclear explosions. The inherent challenge inspires the use of ensemble samplers with invariant affine to effectively sample the MT parameter space. We demonstrate the advantages of the newly developed inversion scheme on six nuclear tests in DPRK.
- We are currently working towards incorporating 3D structural uncertainties most objectively into the inversion. The actual implementation remains challenging, but the task itself is computationally feasible. When applying the approach to a shallow nuclear explosion, the source type ambiguity in shallow depths remains a dominant issue that requires further refinement.

Approved for public release; distribution is unlimited.

2 Toward Improving Point-Source Moment-Tensor Inference by Incorporating 1D Earth Model's Uncertainty: Implications for the Long Valley Caldera Earthquakes

Summary

To infer seismic moment tensors (MT) of moderate earthquakes at regional scales, seismologists typically simulate waveforms using available structural models of the Earth to match the observed seismograms. This procedure is known as waveform seismic MT inversion. However, the seismic data are noisy, and the Earth model is inevitably different from the actual structure of the Earth; hence, there is a discrepancy between the predicted and the observed waveforms. This discrepancy arises from the noise in the data and imperfections in theoretical predictions, stemming most significantly from the Earth model. This study introduces structural uncertainty, estimated empirically, and referred to as "theory uncertainty," as part of the combined covariance matrix. In the synthetic setting, we first show through a series of synthetic experiments that the structural uncertainty plays a critical role in retrieving MT solutions, especially for short-period waveforms. The method is then benchmarked against the waveforms of non-double-couple earthquakes in Long Valley Caldera, California. We confirm the highly isotropic nature of the source in a pilot event but find a non-negligible CLVD component that was overlooked in past studies that ignored the Earth model uncertainty. Thus, careful consideration of the Earth model's uncertainty as part of MT inversion schemes will be necessary for future applications to better understand the complicated physics of earthquake sources.

Plain Language Summary

The representation of a seismic source, which can be underground faulting, a volcano-unrest event, or a manmade explosion, is often through an equivalent force system placed at a point, known as a point-source moment tensor (MT). Its inference from the observed seismic waveforms is referred to as an MT inverse problem. This study is at the forefront of solving this inverse problem by accounting for two primary sources of unavoidable uncertainties: (a) structural uncertainty related to the imperfection of the Earth model representing the true Earth, and (b) data noise stemming from the observed time series of ground motion. Although the former outweighs the latter in practice, it is often ignored. Here, we show that an appropriate accounting for both uncertainties is critical in yielding reliable MT solutions. It allows for the meaningful interpretation of the nature of seismic sources, including nondouble couple events and clandestine nuclear explosions. We illustrate this both through comprehensive synthetic experiments and an in-depth analysis of a volcanic event in the Long Valley Caldera, CA, for which a significant nondouble-couple component is now revealed.

2.1 Introduction

A moment tensor (MT), a 3×3 matrix, is a convenient mathematical representation of underground seismic sources, including faulting or volume sources, by a system of body forces under a point-source approximation (Aki & Richards, 2002, p. 37). Each element, m_{ij}, where $i, j = 1, 2, 3$, of the MT represents an opposite force-couple acting along coordinate i with the arm resting along coordinate j. The inversion for a seismic moment-tensor (MT) from seismological observations is a classical inverse problem in seismology. Reliable MT solutions play an essential role in forensic seismology as they are important evidence in determining the nature of source types, which distinguish between tectonic, volcanic, or explosive sources (Alvizuri et al., 2018; Ford et al., 2009; Mustać et al., 2020; Nettles & Ekström, 1998). Additionally, obtained MT solutions are essential interim products that are inputs for a wide range of other seismological applications, such as in structural tomography across scales (e.g., Bozdağ et al., 2016; Lloyd et al., 2020), earthquake and tsunami early warnings (e.g., Allen &

Melgar, 2019; Käufl et al., 2016), probabilistic seismic hazard maps (e.g., Gerstenberger et al., 2020) or studies of coda correlation (e.g., Tkalčić & Phạm, 2020). There exist catalogues routinely updated for recent moderate-to-large earthquakes at global scales (e.g., GCMT; Ekström et al., 2012), or at regional scales (e.g., CSCMT-3D; Hejrani et al., 2017).

The numerical procedure for inferring an MT solution from recorded seismic waveforms is referred to as MT inversion. Methods for MT inversion can be grouped in four categories: *(1)* methods based on first arrival polarities; *(2)* methods based on amplitude ratios; *(3)* methods based on full waveforms; and *(4)* hybrid methods. Shang and Tkalčić (2020) discussed the advantages and disadvantages of each category in previous studies. For moderate earthquakes (i.e., Mw ~4.0–6.0), full-waveform inversion is preferable as waveform data have a sufficiently high signal-to-noise ratio. As the full-waveform inversion approach utilizes amplitude time series, the comprehensive approach could be used to solve for the temporal dependence of energy release (i.e., source time function), centroid location, and MT solution (e.g., Pasyanos et al., 1996; Stähler & Sigloch, 2014; Vackář et al., 2017). The fact that source time functions and centroid locations are often unknown causes a severe issue of non-linearity in seismic inverse problems.

The crucial impact of Earth structure models on obtained MT solutions is well known. At the global scale, where very long seismic signals are being used (up to a few hundred seconds), reference Earth models such as ak135 (Kennett et al., 1995) were found to be efficient (e.g., Ekström et al., 2012; Weston et al., 2018). However, shorter period waveforms, i.e., a few tens of seconds, are used to invert for MT solutions when it comes to regional scales. Local structure models are preferred over global reference models in the MT inversion (e.g., D.S. Dreger et al., 2000; Panning et al., 2001). However, regardless of whether it's a 1D or 3D model, any Earth model is imperfect; in the full waveform inversion context, that means that there is always a discrepancy between the actual response of the Earth and the simulated seismograms based on the Earth model even given a perfectly known earthquake source. This discrepancy is then referred to as the error of the forward problem, which consists of a significant part of the theory uncertainty. This part of the error has often been ignored, where all the residuals between the observed and simulated waveforms were attributed to data noise uncertainty.

The importance of uncertainty in the forward problem was conceptualized in the 1980s (Tarantola & Valette, 1982) and further explored in Tarantola (2005). The different sources of uncertainty are represented by covariance matrices, including the theory uncertainty (noise) covariance matrix and the data uncertainty (noise) covariance matrix. The sum of them represents the uncertainty characteristics of the waveform residual between observed and simulated waveforms. Concerning seismic sources studies, Duputel et al. (2014) and Yagi and Fukahata (2011) were among the first attempts to develop the theory employing the structural uncertainty covariance matrix when applying to finite fault inversion.

This study is an attempt to incorporate the structural uncertainty of 1D Earth models in point-source MT inversion within the Bayesian framework, while the treatment of structural uncertainty in 3D is an important subject of near-future research. There have been several important attempts to quantify the uncertainty in the 1D Earth model in the form of a prediction covariance matrix. Duputel et al. (2012) considered first-arrival time uncertainty only, whereas Hallo & Gallovič (2016) presented a fast and cheap approximation of the covariance matrix. In the most recent study, Vasyura-Bathke et al. (2021) compared different methods in constructing the structure uncertainty covariance matrix and favoured the non-Toepliz form of the covariance matrix. Here we use the Monte-Carlo method to propagate the Earth-structure model uncertainty, in the form of an ensemble of Earth models, into an ensemble of predicted waveforms and then estimate the structural uncertainty covariance matrix from the collection

of waveform realizations. In addition, we consider the hierarchical aspect of the inverse problem, where the amplitude of the data noise covariance matrix is treated as a free parameter to be determined by input data. One can have additional assurance regarding how well the structural uncertainty is accounted for when the recovered data noise amplitudes are reasonably well estimated. We also examine the importance of incorporating structural uncertainty when several filter bands are used.

The rest of this manuscript is structured as follows. The theory for MT inversion will be elaborated in Section 2.2, adopting the framework presented in Duputel et al. (2014). An independent description of the theory in their work is adapted for the context of this work and included in the Supporting Information. In Section 2.3, the synthetic experiments are designed to demonstrate the importance of the structural uncertainty covariance matrix for simulated waveforms filtered at three different period ranges: 6–50 s, 10–50 s, and 20–50 s. An application to the Long Valley Caldera event is presented in Section 2.4, and it is followed by thoughts on the limitations of the current framework and the way forward.

2.2 Theoretical background

This section reviews the relevant theory and then defines essential elements to pose the seismic inverse problem for MTs. In a full-waveform inversion, we consider a collection of three-component seismograms (i.e., vertical, radial, and tangential) from $\boldsymbol{N}_k$ stations,

$$\boldsymbol{d}_{obs}^{kj}; k = 1, \dots, N_k; j = Z, R, T \tag{1}$$

to constrain the unknown MT, denoted as $\boldsymbol{m}$, of a seismic source. All observed seismograms have the same length of $\boldsymbol{N}$ points. Without any loss of generality, we work with seismograms being sampled at 1 Hz in this study. To simplify the notation, in this section, we drop the superscript kj and use the notation $\boldsymbol{d}_{obs}$ for a single component seismogram. The cumulative account of all seismograms will be presented at the end of this section. Note that $\boldsymbol{m}$ is generic notation for an MT, which can be parameterized differently depending on the purposes described in Section 2.3.1. We use bold-face symbols to denote column vectors or matrices through this manuscript (see Table S1 in the Supporting Information for a complete glossary of symbols).

An essential element to cast the seismic inverse problem is a forward problem,

$$\boldsymbol{d}_{pred} \coloneqq g(\Omega, \boldsymbol{m}), \tag{2}$$

which is an efficient solver of the wave propagation problem that outputs a waveform, $\boldsymbol{d}$, from an input MT, $\boldsymbol{m}$, in a velocity model Ω. The forward problem is also known as a deterministic prediction of output waveforms. As shown in Section 2.3, we use the routines packaged in the Computer Programs in Seismology (CPS; Herrmann, 2013) to solve the forward problem (practical details of the forward problem can be found in Section 2.3.2).

When the epicentre and origin time of the source is fixed, the forward problem can be conveniently expressed as a linear expression (e.g., Jost & Herrmann, 1989; Shang & Tkalčić, 2020),

$$\boldsymbol{d}_{pred} = \boldsymbol{G}^{\Omega} \cdot \overline{\boldsymbol{m}}. \tag{3}$$

The pre-computed Green's functions, $\boldsymbol{G}^{\Omega}$, represent the impulse response of a velocity model, Ω, corresponding to six elementary MTs that form the linear basis for an MT parameterization.

Note that we use the notation $\overline{\boldsymbol{m}}$ for a particular parameterization of the generic MT, so that the forward problem can be written in the linear form.

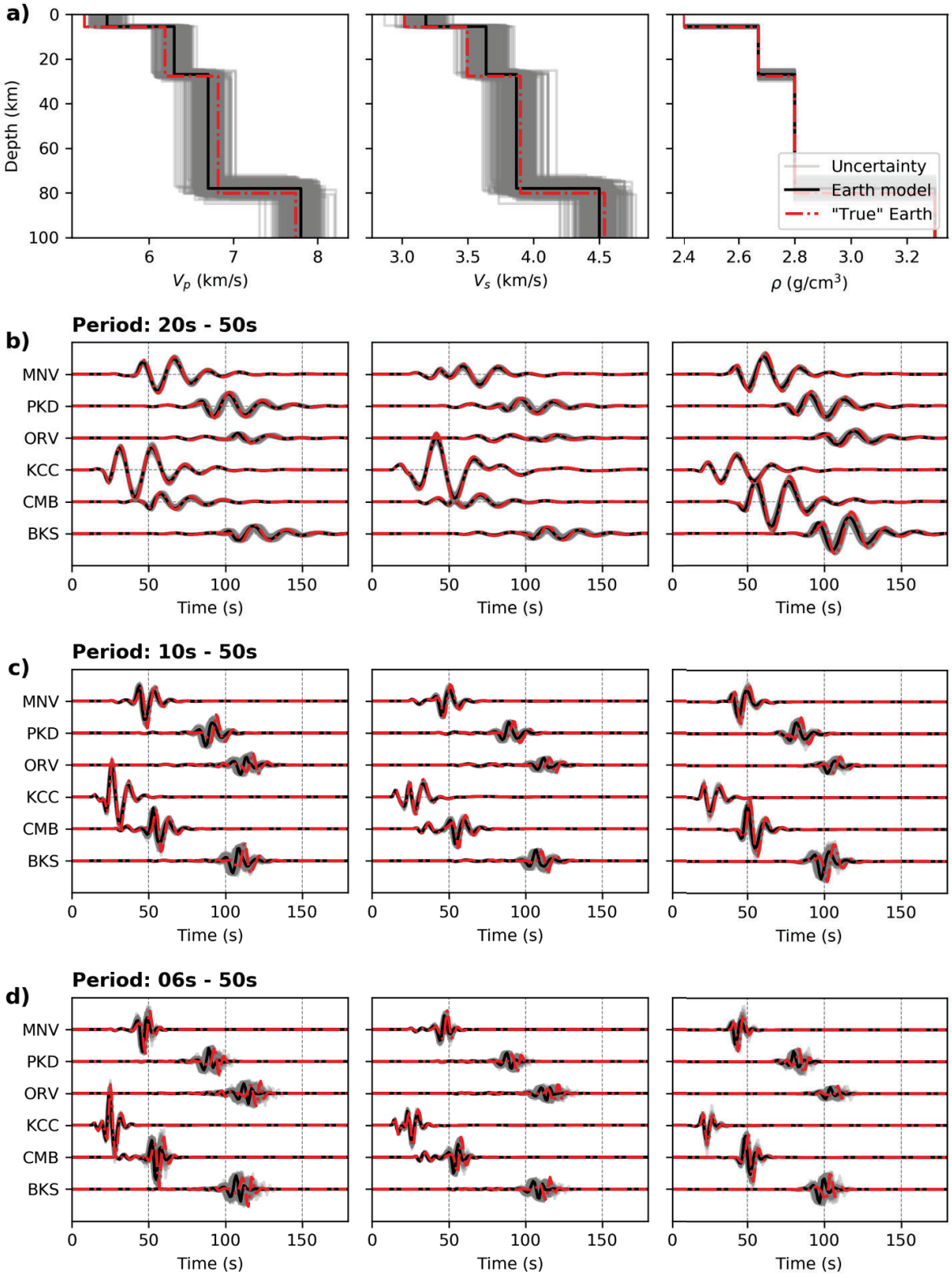

Figure 1. (a) Various Earth models used in synthetic experiments and real data MT inversion in this study

Black lines show the SoCal model (D. S. Dreger & Helmberger, 1990), denoted as $\boldsymbol{\Omega}$ (see Table S1), used as the reference Earth model to compute Green's functions for the Long Valley Caldera earthquakes. Gray lines represent a randomly perturbed ensemble of models from SoCal, denoted as $\{\boldsymbol{\Omega_i}\}$. Dotted-dashed red lines represent the imaginary Earth model that serves as the "true" Earth in our experiments, denoted $\widetilde{\boldsymbol{\Omega}}$. The "true" Earth is used to generate "fake data" for synthetic experiments. ***(b)*** **Waveforms filtered between 20 and 50 s, corresponding to the Earth models in a) for a common MT.** ***(c)*** **Similar to b) but waveforms are filtered between 10 and 50 s.** ***(d)*** **like b), but waveforms are filtered between 6 and 50 s.** See Figure 3 for the location map of stations and MT of the source.

Apart from defining a forward problem, the other essential element to cast the seismic inverse problem is the misfit between the observed and predicted waveforms. A good misfit function must account for relevant uncertainties in the observational process and the prediction theory. In a probabilistic framework, the misfit function is used to express how likely it is that an instance of the unknown model explains the observation. Noise, or uncertainty, as part of the observational process, is known as data noise. This noise is often assumed to be equivalent to the random ambient noise at the recording sites.

Theory uncertainty in the forward problem refers to the possible mismatch between the theoretical prediction and the actual observation. In our context, this would be a mismatch in the ground motion due to imperfect knowledge of the Earth structure, a 1D simplification of the 3D Earth, or an approximation made in computing the predicted Earth response, for example, a point-source approximation of a finite fault rupture. In this proof-of-concept study, we assume that the imperfect knowledge of the Earth structure is the most significant source of uncertainty in the theoretical prediction. More specifically, we consider a simple 1D structure over the entire study region and refer to it as the *"true Earth structure"*, $\widetilde{\Omega}$. The subjective knowledge of the velocity structure is the *reference Earth model*, $\widehat{\Omega}$. However, our knowledge of the Earth model is uncertain, so it is reasonable to represent it as a collection of structural model realizations, $\{\Omega_i, \quad i = 1, \dots, N_m\}$, representing the uncertainty deviating around the reference model. It can be said that the reference and its uncertainty are of good quality when the "true Earth" lies within the uncertainty estimate of the reference model, particularly, $\widetilde{\Omega} \in \{\Omega_i\}$.

The data noise is assumed due to the background noise at the recording sites, which is independent of actual ground motion induced by seismic waves, $\boldsymbol{d}$. The actual ground motion is the hypothetical ground motion recorded in the absence of observational noise. The stochastic characteristics of the data-noise can be described by the data covariance matrix, $\boldsymbol{C}_d$, which governs the stochastic behavior of each sampling point in time and possible correlation between them. Furthermore, given that the "true Earth" is within the uncertainty representation of the reference model, which is $\widetilde{\Omega} \in \{\Omega_i\}$, the theory noise covariance matrix can be estimated from the collection of structural models, $\{\Omega_i\}$,

$$\boldsymbol{C}_t(\boldsymbol{m}) = \frac{1}{N_m - 1}\sum_{i=1}^{N_m}([g(\Omega_i, \boldsymbol{m}) - \bar{g}(\boldsymbol{m})]^T \cdot [g(\Omega_i, \boldsymbol{m}) - \bar{g}(\boldsymbol{m})]), \tag{4}$$

where $\bar{g}(\boldsymbol{m}) = \frac{1}{N_m}\sum_i^{N_m} g(\Omega_i, \boldsymbol{m})$ is the mean Green's function. Ideally, the theory covariance matrix corresponding to the unknown MT, $\boldsymbol{m}$, is correctly estimated with the true MT, $\widetilde{\boldsymbol{m}}$. However, because the true solution is unknown, we can instead estimate the structural

uncertainty covariance matrix with a prior estimate of moment tensor, $\boldsymbol{m}^p$, which is expected to be close to $\widetilde{\boldsymbol{m}}$.

The likelihood of the observation $\boldsymbol{d}_{obs}$ given a moment tensor $\boldsymbol{m}$ considering structural uncertainty can be expressed as follows,

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \frac{1}{(2\pi)^N \sqrt{|\boldsymbol{C}_D|}} \exp\left(-\frac{1}{2}\left[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T \boldsymbol{C}_D^{-1} \left[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]\right) \quad (5)$$

given that the combined covariance matrix is defined as,

$$\boldsymbol{C}_D(\boldsymbol{m}^p) = \boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^p). \quad (6)$$

Operators $|\boldsymbol{C}_D|$ and $\boldsymbol{C}_D^{-1}$ are the determinant and inversion of the data combined covariance matrix and the superscript T denotes the transposed matrix/vector. The rigorous derivation of Equation (5) is presented in the Supporting Information. In this study, we also consider the hierarchical aspect of the problem by letting the noise strength σ be a free parameter, which will be determined by the data alongside the MT, $\boldsymbol{m}$. This is somewhat like previous work by Mustać & Tkalčić (2016, 2017), in which the authors inverted noise amplitudes of correlated noise, whereas this study only considers the strength of uncorrelated data noise. Therefore, the final form of the combined covariance matrix to be implemented, which describes the nature of uncertainties in residual waveforms, in the next section is,

$$\boldsymbol{C}_D(\sigma, \boldsymbol{m}^p) = \sigma^2 \boldsymbol{I} + \boldsymbol{C}_t(\boldsymbol{m}^p). \quad (7)$$

Note that by using the diagonal data noise covariance matrix, we assume uncorrelated data noise, the rationale behind this simplification is further discussed in Section 2.3.3. When all seismogram components are considered, the integrated likelihood function with a fully model-dependent covariance matrix is defined as

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \prod_{k=1}^{N_k} \prod_{j=1}^{3} \frac{1}{(2\pi)^N \sqrt{\left|\boldsymbol{C}_D^{kj}\right|}} \exp\left(-\frac{1}{2}\left[\boldsymbol{d}_{obs}^{kj} - g^{kj}\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T \left(C_D^{kj}\right)^{-1} \left[\boldsymbol{d}_{obs}^{kj} - g^{kj}\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]\right) \quad (8)$$

In the following, we rely on this likelihood function to build the Bayesian inference methodology incorporating the Earth model's uncertainty.

2.3 Methodology

2.3.1 Parameterization of seismic moment tensors

A seismic MT is a 3×3 matrix consisting of 9 element entries,

$$\begin{bmatrix} m_{xx} & m_{xy} & m_{xz} \\ m_{yx} & m_{yy} & m_{yz} \\ m_{zx} & m_{zy} & m_{zz} \end{bmatrix}. \quad (9)$$

It represents a body-force system equivalent to the discontinuities in displacement or strain due to seismic faults (chapter 3, Aki & Richards, 2002). Each matrix entry, m_{ij}, of the MT represents an opposite force couple with its arm along coordinate i and its direction along

coordinate j. Under the assumption of no net torque, reciprocal force couples are equal, that is, $m_{ij} = m_{ji}$, so the MT matrix is symmetric with six independent entries. Thus, an MT matrix can be described as a point in the $\boldsymbol{R}^6$ space. Here we will consider two complementary methods to parameterize a symmetric MT, equivalent via a bijective transformation. The first parameterization enables us to cast the forward problem in a linear matrix multiplication form (Equation (3)), while the second parameterization is uniform "*in the sense that equal volumes in the coordinate domain of the parametrization correspond to equal volumes of moment tensors*" (Tape & Tape, 2015). The uniformity can enhance the efficiency of the Bayesian sampling (e.g., Fichtner & Simutė, 2018; Stähler & Sigloch, 2014).

Firstly, we consider a parameterization of a seismic MT in the primitive coordinates,

$$\overline{\boldsymbol{m}} = \left[m_{xx}, m_{yy}, m_{zz}, m_{xy}, m_{xz}, m_{yz}\right]^T \in \boldsymbol{R}^6 \tag{10}$$

In this coordinate system, a complete seismogram can be represented as a linear combination of elementary Green's functions

$$\boldsymbol{d}_{pred} = \sum_{i=1}^{6} m_i \cdot \boldsymbol{g}_i^{\Omega}. \tag{11}$$

The numerical index, i, corresponds to elements in Equation (10). Elementary seismograms, $\boldsymbol{g}_i^{\Omega}$, correspond to 6 MT bases, where the corresponding element entry (Equation (9)) equals 1 while the others are 0. In this setting, the forward problem can be expressed as a linear matrix multiplication as in Equation (3), in which the Green's function matrix is defined as follows,

$$\boldsymbol{G}^{\Omega} = \left[\boldsymbol{g}_{xx}^{\Omega}, \boldsymbol{g}_{yy}^{\Omega}, \boldsymbol{g}_{zz}^{\Omega}, \boldsymbol{g}_{xy}^{\Omega}, \boldsymbol{g}_{xz}^{\Omega}, \boldsymbol{g}_{yz}^{\Omega}\right]^T. \tag{12}$$

An alternative method of linear parameterization was proposed originally by Kikuchi and Kanamori (1982) and has been used in several studies (e.g., Hejrani et al., 2017; Mustać & Tkalčić, 2016; Vackář et al., 2017). In this study, we prefer primitive parameterization (Equation (10)) solely because of its convenience in parameterization.

The second method of MT parameterization relies on a theoretical work in analytical geometry by Tashiro (1977) that defines a point uniformly in an *n*-sphere, a hyper-sphere in $\boldsymbol{R}^n$ space. When an MT (Equation (10)) is standardized by its own scalar moment,

$$M_0 = \sqrt{m_{xx}^2 + m_{yy}^2 + m_{zz}^2 + 2(m_{xy}^2 + m_{xz}^2 + m_{yz}^2)}, \tag{13}$$

it defines a point on a unit ellipsoid in $\boldsymbol{R}^6$ that has 5 degrees of freedom,

$$\frac{m_{xx}^2 + m_{yy}^2 + m_{zz}^2 + 2(m_{xy}^2 + m_{xz}^2 + m_{yz}^2)}{M_0^2} = 1 \tag{14}$$

Given that, $x_i, i = 1, \ldots 5$ are five real numbers bounded between 0 and 1, they determine a point in the unit *6*-ellipsoid (Stähler & Sigloch, 2014),

$$
\begin{aligned}
Y_3 = 1;\ Y_2 = Y_3\sqrt{x_2};\ Y_1 = Y_2 x_1 \\
m_{xx}/M_0 = \sqrt{Y_1}\cos(2\pi x_3)\sqrt{2} \\
m_{yy}/M_0 = \sqrt{Y_1}\sin(2\pi x_3)\sqrt{2} \\
m_{zz}/M_0 = \sqrt{Y_2 - Y_1}\cos(2\pi x_4)\sqrt{2} \\
m_{xy}/M_0 = \sqrt{Y_2 - Y_1}\sin(2\pi x_4) \\
m_{xz}/M_0 = \sqrt{Y_3 - Y_2}\cos(2\pi x_5) \\
m_{yz}/M_0 = \sqrt{Y_3 - Y_2}\sin(2\pi x_5)\,.
\end{aligned}
\tag{15}
$$

We then refer to this method as Tashiro's parameterization of the MT,

$$
\boldsymbol{m} \coloneqq (x_1, x_2, x_3, x_4, x_5, M_w) \in [0,1]^5 \times [3,6]. \tag{16}
$$

The moment magnitude, $M_w = 2(\log M_0 - 16.1)$ (Kanamori, 1977) is bounded between 3 and 6 for moderate earthquakes considered in this study. Readers are referred to Tape and Tape (2015) for more insights into uniform parameterization of MT and an alternative parameterization methodology (e.g., Alvizuri et al., 2018; Alvizuri & Tape, 2018a). Because the transformation system in Equation (15) is invertible, Tashiro's parameterization is equivalent to the primitive parameterization via a bijective transformation.

2.3.2 Forward simulation of full waveforms

To simulate full waveforms propagating in horizontally stratified media, we follow the method originally proposed by Jost and Herrmann (1989), which was then improved by Minson and Dreger (2008). Predicted seismograms of vertical, radial, and tangential components, $\boldsymbol{d}^Z_{pred}, \boldsymbol{d}^R_{pred}, \boldsymbol{d}^T_{pred}$, are linear combinations of pre-computed seismograms of elementary source types,

$$
\begin{aligned}
\boldsymbol{d}^Z_{pred} &= A_1 \cdot \boldsymbol{z}_{SS} + A_2 \cdot \boldsymbol{z}_{DS} + A_3 \cdot \boldsymbol{z}_{DD} + M_{iso} \cdot \boldsymbol{z}_{EX} \\
\boldsymbol{d}^R_{pred} &= A_1 \cdot \boldsymbol{r}_{SS} + A_2 \cdot \boldsymbol{r}_{DS} + A_3 \cdot \boldsymbol{r}_{DD} + M_{iso} \cdot \boldsymbol{r}_{EX} \\
\boldsymbol{d}^T_{pred} &= A_4 \cdot \boldsymbol{t}_{SS} + A_5 \cdot \boldsymbol{t}_{DS}
\end{aligned}
\tag{17}
$$

In Equation (17), ten basis seismograms, $\boldsymbol{z_X}, \boldsymbol{r}_X, \boldsymbol{t}_X$, correspond to four elementary source types: vertical strike-slip fault (SS), vertical dip-slip fault (DS), 45° dip-slip fault (DD), and an explosion (EX). The Green's functions are computed by the frequency-wavenumber method implemented in the Computer Program in Seismology package (Herrmann, 2013). Six coefficients, $A_{1\ldots5}$ and M_{iso} are computed from 6 MT components in the primitive parameterization, $\overline{\boldsymbol{m}}$, and the receiver's azimuthal angle, ϕ,

$$
\begin{aligned}
A_1 &= \frac{1}{2}\left(m_{xx} - m_{yy}\right)\cos 2\phi + m_{xy}\sin 2\phi \\
A_2 &= m_{xx}\cos\phi + m_{yz}\sin\phi \\
A_3 &= -\frac{1}{6}\left(m_{xx} + m_{yy} - 2m_{zz}\right) \\
A_4 &= \frac{1}{2}\left(m_{xx} - m_{yy}\right)\sin 2\phi - m_{xy}\cos 2\phi \\
A_5 &= -m_{yz}\cos\phi + m_{xz}\sin\phi \\
M_{iso} &= \frac{1}{3}\left(m_{xx} + m_{yy} + m_{zz}\right)
\end{aligned}
\tag{18}
$$

Seismic waveforms for an MT in Tashiro's parameterization can be computed via the transformation in Equation (15).

2.3.3 Data-noise covariance matrix and noise-magnitude hyper-parameters

As mentioned above, the data-noise covariance matrix accounts for the observation error, which is mainly the background ambient noise at the recording site. It is a part of the misfit function in the optimization framework or the likelihood function in the Bayesian framework. The data-noise covariance matrix contains information reflecting the stationary nature of uncertainty in the seismic wavefield. Hence, it has been an active research area to find an efficient way to describe the data noise in the form of a covariance matrix to obtain a reliable MT solution.

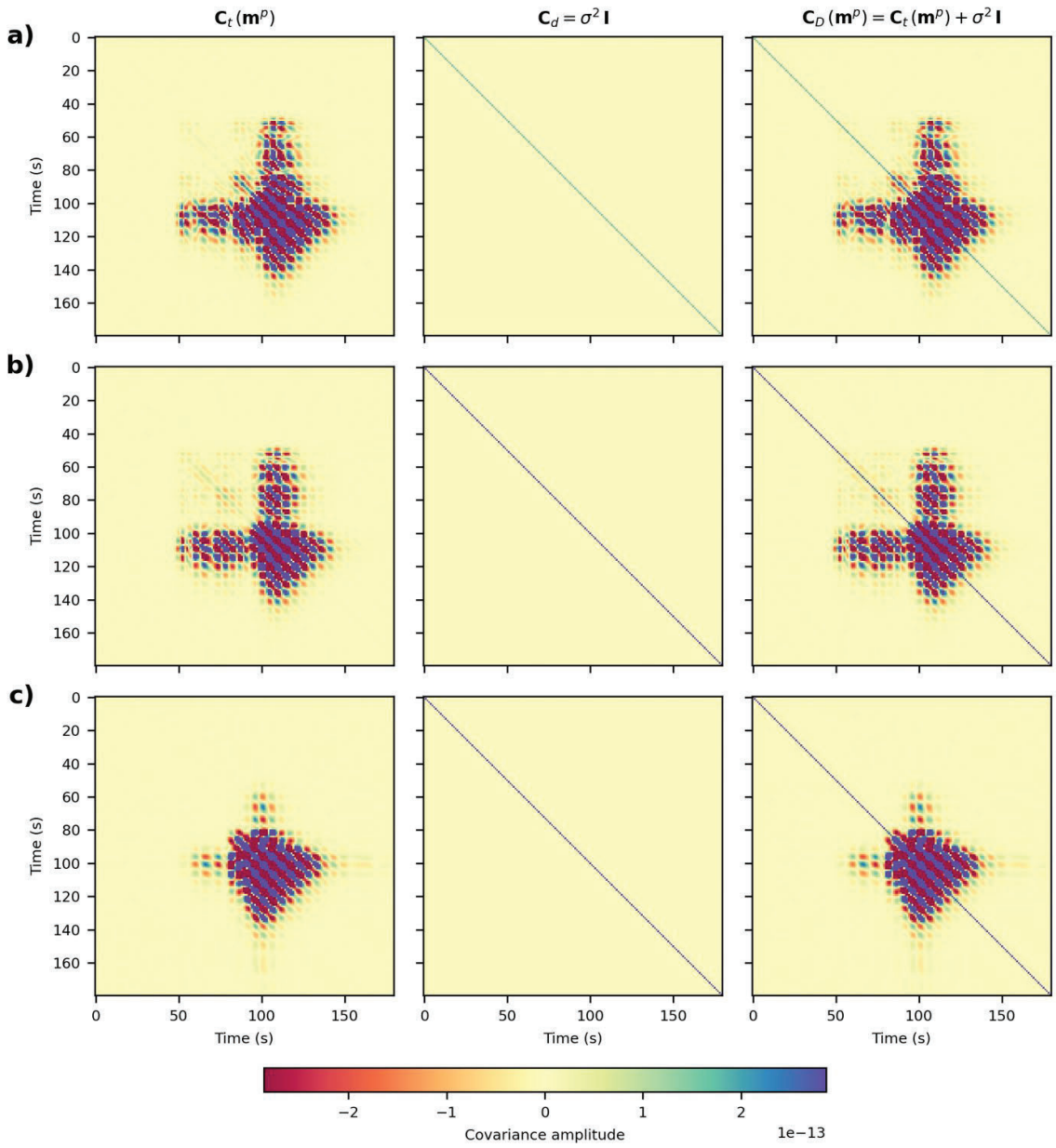

Figure 2. Visualization of covariance matrices for each row corresponding to one of the three recording velocity components for station BKS: *(a)* vertical, *(b)* radial, and *(c)* tangential.

The left column shows the structural uncertainty covariance matrices corresponding to the true MT input, that is $\mathbf{C_t}(\mathbf{m^p})$ where $\mathbf{m^p} = \mathbf{m^{true}}$. The middle column shows the diagonal covariance matrix representing uncorrelated data noise, that is, $\mathbf{C_d}$. The right column shows the combined covariance matrix, which is the sum of the covariance matrices in the left and middle columns, that is, $\mathbf{C_D}(\mathbf{m^{true}}) = \mathbf{C_t}(\mathbf{m^{true}}) + \mathbf{C_d}$.

A diagonal data-noise covariance matrix is a common choice in geophysical studies for the data noise covariance matrix, possibly because its inversion and determinants are trivial to compute. However, seismic noise has been shown to be correlated in time, which could be due to band-limited filters in data preparation and the band-limited nature of seismic noise. Progress has been made recently in considering this temporal correlation. For example, Duputel et al.

(2012) highlighted the importance of off-diagonal components in the covariance matrix to incorporate the correlated nature in seismic noise. Their study used a decaying exponential model to represent the time-dependent relationship. Mustać and Tkalčić (2016) took a step further and described the noise correlation by a two-attenuated-cosine function that best approximates empirical noise autocorrelation. In other attempts, Vackář et al. (2017) employed pre-event noise to construct empirical covariance matrices, while Mustać et al. (2020) followed Dettmer & Dosso (2012) to use waveform residuals with a preliminary prediction to estimate an empirical covariance matrix.

Unlike the studies mentioned above, we attempt to explicitly separate the two sources of uncertainties by having two covariance matrices (Equation (6)). When we select stations with excellent signal-to-noise ratio waveforms to conduct the inversion, it is likely that the residual due to the imperfect knowledge of Earth structure dominates the observational noise. Therefore, we simplistically assume that structural uncertainty is dominantly responsible for the correlated nature of residual waveforms, and data noise is uncorrelated and stationary over time. Thus, the data covariance matrix is proportional to the identity matrix. We then incorporate the Bayesian inversion's hierarchical aspect (Bodin et al., 2012; Malinverno & Briggs, 2004; Mustać & Tkalčić, 2016, 2017) to let the data determine the appropriate levels of data noise apart from structural uncertainty. If close to the pre-event noise, the recovered data noise levels would then confirm that our assumption of stationary ambient noise is acceptable.

Specifically, we employ the simplified approach where the data noise covariance matrices are set to be proportional to the identity matrix. The multiplication factors of the diagonal covariance matrices are hyper-parameters that determine the magnitude of the noise to be determined by the data. At each station, we define a noise level hyperparameter, $h_k, k = 1, N_k$, which plays a role as a scaling factor for noise magnitudes at that station. Three reference noise magnitudes, $\sigma_{kj}^{ref}, j = Z, R, T$, are estimated from 1-hour pre-event noise traces of three components of each station. The absolute noise magnitude of a station component during the event is scaled from the reference noise magnitude,

$$\sigma_{kj}^{abs} = h_k \sigma_{kj}^{ref}. \tag{19}$$

The rationale behind this parameterization is to avoid too wide of dynamic ranges of noise parameters as they vary significantly among stations. Furthermore, it is also convenient that if the data noise is stationary before and during the event, it is expected to have a noise scale of around 1. In the synthetic experiments presented in Section 2.4, we set $h_k = 1$ for all stations meaning that the noise levels in the simulated waveforms are equal to the real pre-event noise magnitudes at all components. In the inversion, we search for the noise levels in the interval between 0 and 300. Mathematically speaking, the hyper-parameter space is a subspace of the 3-D real space, $H = (0, 300)^{N_k}$. Figure 2 visualizes three data noise covariance matrices with their true amplitudes used in the synthetic experiments in Section 2.4.

2.3.4 Structure-uncertainty covariance matrix

The available 1D model, SoCal (D. S. Dreger & Helmberger, 1990), widely used to study regional earthquakes in the Long Valley Caldera region (e.g., Minson & Dreger, 2008; Mustać & Tkalčić, 2016), is used as the reference model, $\widehat{\Omega}$, in this study. The model consists of 4 layers and covers a depth of 100 km from Earth's surface. As mentioned in the introduction, we use the Monte Carlo method to quantify the uncertainty of the Earth model in the form of a structural uncertainty covariance matrix.

Firstly, the reference model, SoCal, is perturbed to obtain an ensemble of Earth models to represent the uncertainty. A single controlling parameter for the perturbation is the magnitude of uncertainty in percentage, κ. Although we test other values, we chose $\kappa = 5\%$ in numerical experiments for a demonstrative purpose in this study. The sensitivity of this subjectively chosen parameter will be examined in the real-data application. Values of density, P- and S-wave velocities, and the thickness of all layers vary, while the number of layers and the density profile are unchanged in random model realizations.

Moreover, rather than deviating directly from the (P- and S-wave) velocity profiles, we do so in the bulk and shear modulus fields to achieve a homogeneous distribution of elastic parameters (p. 164, Albert Tarantola, 2005). At each random realization, we replace each value in the bulk and shear moduli and the thickness of each layer in the original model, say v, is replaced by a new value, $v + \delta v$, where δv is a random number drawn from the Gaussian distribution centered at 0, with a standard of deviation of κv. Figure 1a shows an ensemble of the 300 perturbed models. Ensembles of waveforms corresponding to the models (Figure 1a, 1b, 1c) are used to compute the structural uncertainty covariance matrices, as shown in Figure 2.

2.3.5 Iterative seismic inversion scheme

As noted in the theoretical background (Section 2.2), there is a circular dependence of the prior MT, $\boldsymbol{m}^p$, that defines the structural uncertainty covariance matrix and the unknown MT, $\boldsymbol{m}$. Here we propose a practical solution to break the circular puzzle by an iterative inversion scheme. Initially, the prior MT can be chosen either randomly or using a rough estimate of the MT via, for example, the least squares solution without the structural uncertainty covariance matrix. The mean solution obtained in the Bayesian sampling chain can then be used as an estimate of the prior MT in the next iteration. The inversion using Bayesian sampling can be repeated several times to find a stable solution. In Section 2.4.4, we empirically prove the feasibility and performance of this iterative scheme through numerical examples.

2.4 Synthetic Experiments

2.4.1 Experiment setup

We use a realistic configuration of an event and seismic stations by deploying a well-studied non-DC event that occurred on 20th November 1997, as a representative of the earthquake sequence in Long Valley Caldera (known as event LV2 in Dreger et al. (2000) and Minson and Dreger (2008)). Figure 3 shows the location map of the event and the monitoring network. Available data are from six stations, including five from the Berkeley Digital Seismic Network (BDSN) and a single station (MNV) from the United States National Seismic Network. The addition of station MNV improves the azimuthal coverage for the targeted earthquake in the Nevada site relative to previous analyses of the same event (D.S. Dreger et al., 2000; Minson & Dreger, 2008; Mustać & Tkalčić, 2016). As shown in Figure 1 and Table S1, synthetic waveforms are simulated in an arbitrarily selected velocity model, which is indexed as the 110th model, among 300 random perturbations at 5% from the reference model SoCal (D. S. Dreger & Helmberger, 1990) (see Section 2.3.4 for further details of the model perturbation). The reference model SoCal is used to predict seismograms to match the simulated data. To simplify the inverse problem, we fix earthquake depth at 5 km in the present numerical analysis. The depth and the MT solution of the event are taken from Minson and Dreger (2008).

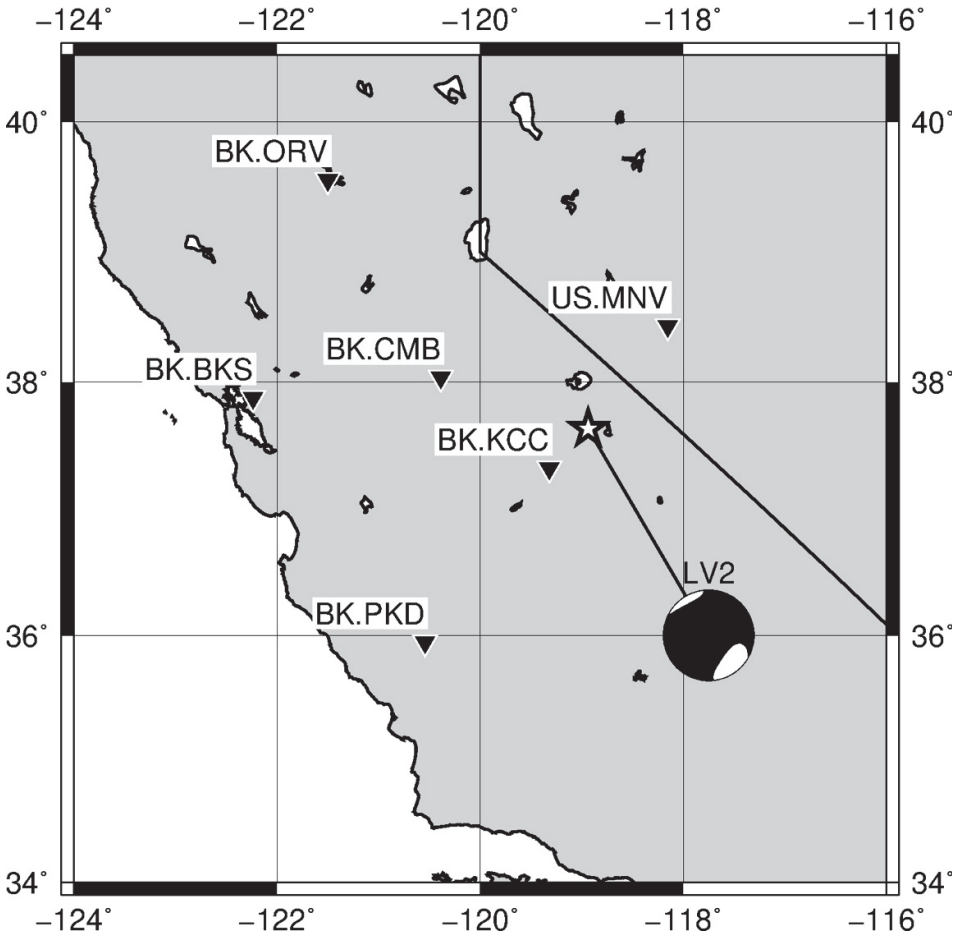

Figure 3. Network configuration for the MT inverse problem.

The non-double-couple event LV2 is marked as an open black star, and the beachball represent s the MT solution from Minson & Dreger (2008), which is later referred to as MD2008 solution in the following figures. Inverted triangles with white edges denote monitoring stations from the Berkeley Digital Seismic Network and US National Seismic Network. This source-receiver configuration is used in both synthetic and real data inversions.

Simulated waveform data are filtered in three period-bands: 20–50 s, 10–50 s, and 6–50 s, to which random noise filtered in the same period bands is added. Specifically, we generate time traces of Gaussian noise of the same length as the simulated waveforms. The random traces are filtered with the bandpass and then divided by their standard deviation. The standardized noise traces are multiplied by the pre-event root mean square (RMS), estimated for individual components, before being added to the corresponding simulated waveforms. For convenience, we will refer to the RMS as the absolute noise level. Using this approach for adding synthetic noise, the absolute noise levels of the added noise are equal to the absolute pre-event noise levels. Thus, the noise scaling factors for all stations, which we use to parameterize the data noise in the inverse problem, are uniformly equal to 1, that is, $h_i = 1, i = 1, \ldots, N_k$,where N_k is the number of seismic stations.

It is important to reiterate that we use a reference Earth model, $\widehat{\Omega}$, in the forward problem, different from the 'true' Earth model, $\widetilde{\Omega}$, used to compute the 'fake' input data in the following numerical experiment (Figure 1a). Thus, predicted waveforms calculated using the reference model with the true MT cannot match perfectly with the synthetic data. Therefore, even in our

synthetic environment, it is inevitable that the input MT is not perfectly recovered regardless of the inversion scheme. Rather than that, we demonstrate that inversion schemes incorporating structure uncertainty can robustly result in a better MT solution than the schemes ignoring the structural uncertainty.

2.4.2 True theory covariance matrices and true data noise levels

To illustrate the importance of incorporating the structural uncertainty in MT inversion in a simplistic manner, we compare the inversion results in the two hypothetical scenarios. In the first scenario, both the data noise amplitudes, σ^{true} and the prior MT, $\boldsymbol{m}^p$ that define the structure uncertainty covariance matrix are known, so the combined covariance matrix can be pre-determined since $\boldsymbol{C}_D = \boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^{true})$ where $\boldsymbol{C}_d = (\sigma^{true})^2\boldsymbol{I}$. In the second scenario, the structural uncertainty is not considered part of the combined covariance matrix, that is: $\boldsymbol{C}_D = \boldsymbol{C}_d$, with $\boldsymbol{C}_d$ defined as in the first scenario. Because the combined covariance matrices in both scenarios can be pre-determined, the least-square solution (Equation (52)) applies, and the solution is the maximum a posteriori estimate in the Bayesian approach (Aster et al., 2013). In the following sections, we present the Bayesian solutions, even though they are computationally expensive.

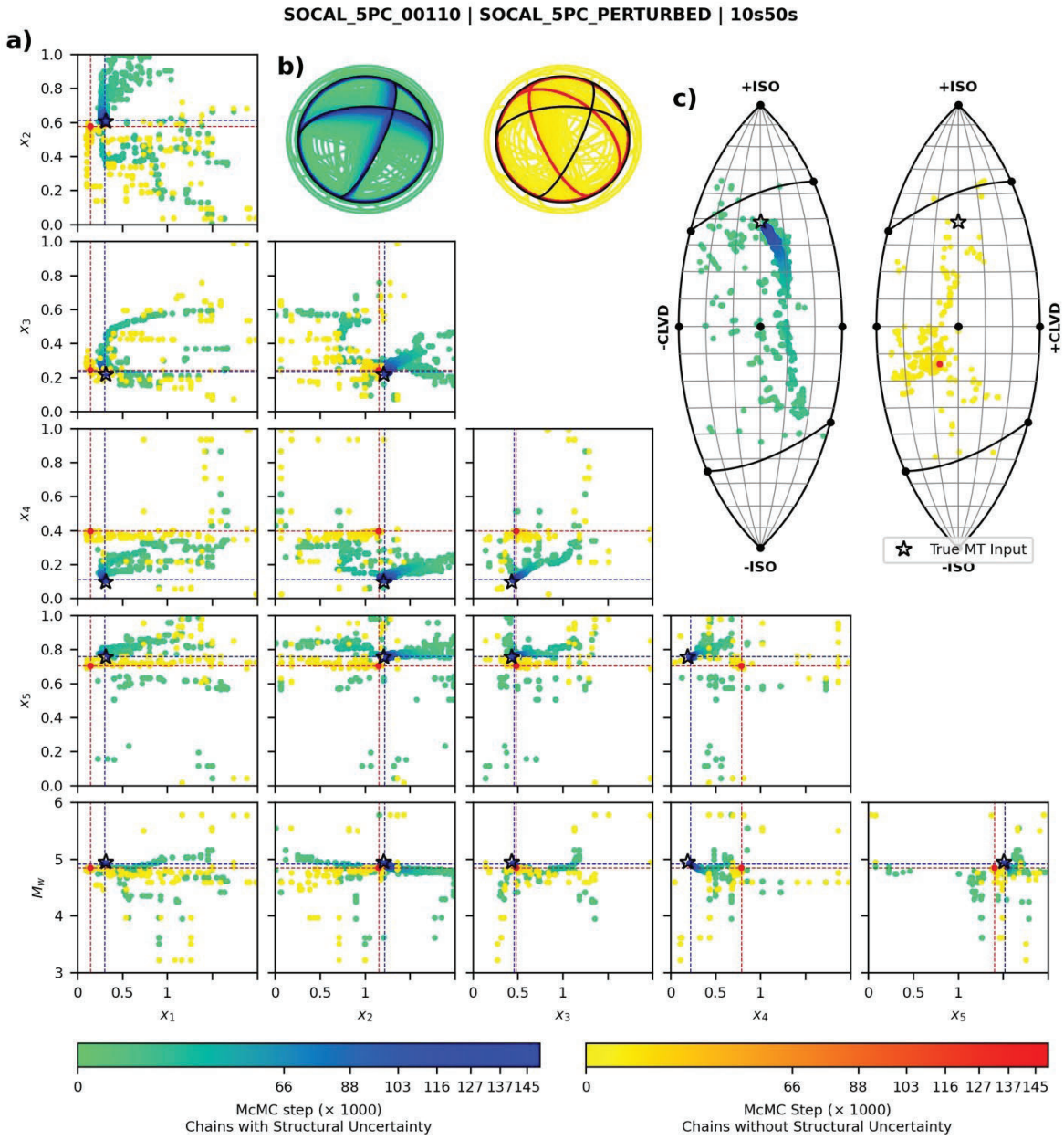

Figure 4. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels.

(See further description of the scenarios in Section 2.4.2). Synthetic data are filtered in the 10–50 s period band. A winter color scheme (transitioning from green to blue) is used for McMC steps in the MT inversion incorporating structural uncertainty, which is computed hypothetically with the true MT input. The autumn color scheme is used for McMC steps in the MT inversion that retains only the data noise components and ignores structural uncertainty. We use the power-law of 2/5 to normalize McMC steps in the color schemes to better visualize the variation at the early stage of the inversion chains, which are regarded as the burning in stage.

The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band considered. (a) Each sub-panel

shows a pair of parameters in Tashiro's (1977) parameterization of MT (see definition Equation (15)). Open black triangles show true input MT. The red and blue dashed lines indicate the position of the mean MT solutions of the post-burning stages of the chains. (b) Beach balls show the fault planes corresponding to the double couple part of the MT solutions in the inversion chains of the two scenarios. Fault planes corresponding to the true MT solution are in black. (c) In lune diagrams of source types (Tape & Tape, 2012), we compare the evolution of the MT solution in two scenarios. Open black stars mark the source type of the true MT input in both diagrams.

We run five independent Markov chains of 150,000 steps for each of the two scenarios presented above to guarantee their common convergence. In each chain, the first half is discarded, and the second half is thinned at every 20th sample; the five chains are later mixed to show the recovered solution in Figure 4. This figure compares the recovered MT solutions in the synthetic experiment, demonstrating the stark difference between accounting and not accounting for structure uncertainty in the period band 10–50 s. All six parameters of the recovered MT solution are significantly closer to the true MT when structure uncertainty is accounted for than when only data noise is considered (Figure 4a). Consequently, the DC component's fault geometries (shown by the beach ball diagram in Figure 4b) and the source types (shown in the lune diagram in Figure 4c) further demonstrate the comparison. A similar outcome is seen for synthetic waveforms filtered in a shorter period band, 6–50 s (Figure 50). In a relatively long period band, 20–50 s, the effect of accounting for structure uncertainty in the inversion is less prominent (Figure 51). Only slightly better solutions can be obtained with the incorporation of structural uncertainty.

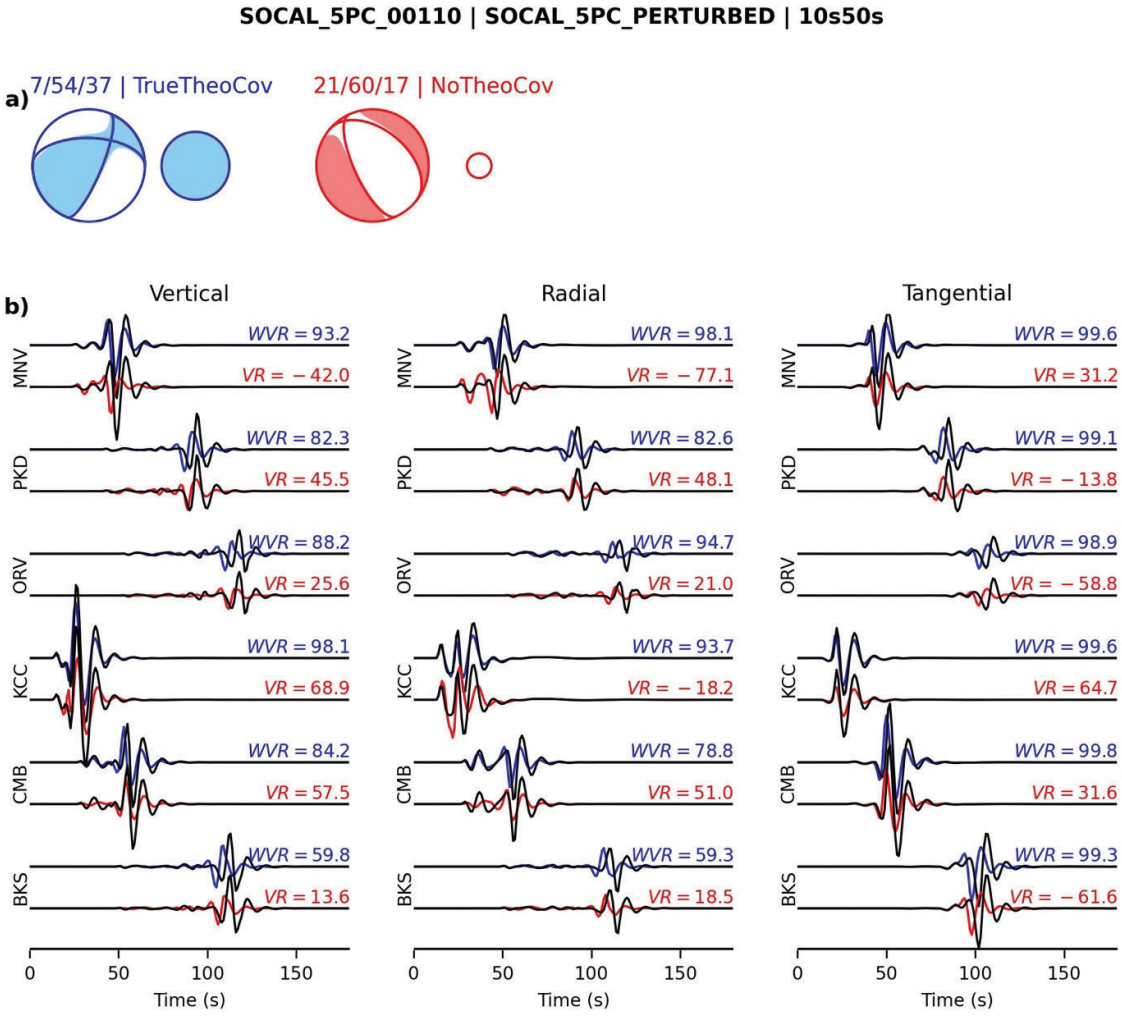

Figure 5. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels.

(See further description of the scenarios in Section 2.4.2). Synthetic data are bandpass filtered in the 10–50 s period band. In all sub-panels, the red color corresponds to the mean MT solution obtained when ignoring structural uncertainty. In contrast, blue corresponds to the MT solution inferred with structural uncertainty covariance accounted for (see the labels on top of the beach balls). The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$ used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band considered. (a) Decomposition of mean MT solutions shown in Figure 4 into deviatoric and isotropic parts. The beachball sizes are proportional to the percentages. Percentages of CLVD, DC, ISO parts in the full MT are displayed on the deviatoric beachballs. A filled isotropic ball means an explosion, while a non-filled isotropic ball means an implosion. b) Waveform fit corresponding to the mean MT solutions of the two numerical scenarios. Input synthetic seismograms are black traces, while the recovered waveforms are in red and blue colors. On the right-hand side of each sub-panel is the weighted variance reduction (WVR – defined in Equation (56-a) for waveforms recovered with structural uncertainty and the variance reduction (VR – defined in Equation 56-b) for waveforms recovered without structural uncertainty.

To gain additional insight into the influence of an appropriately quantified structural uncertainty, we consider the waveform fit corresponding to the recovered MT solutions of the two scenarios under consideration in Figure 5 for period band 10–50 s, while Figure 52 and Figure 53 are similar figures for period bands of 6–50s and 20–50 s respectively. The mean MT solution not accounting for structure uncertainty, visual fit, and variance reduction (VR – defined in Equation (56-b) percentages are particularly low, which is a direct consequence of using a "wrong" Earth model in the inversion described above. When accounting for the structural uncertainty, the combined covariance matrix effectively mitigates the residual between the observed and the predicted waveforms. As a result, the weighted variance reductions (WVR – defined in Equation (56-a) for all components are significantly higher.

The scenarios considered in this problem can be deemed hypothetical as the true MT input is required to correctly determine the structural uncertainty covariance matrix, that is $\boldsymbol{m}^p = \boldsymbol{m}^{true}$. Still, the numerical results in the section provide important empirical evidence to benchmark the theory developed in Section 2.2. Specifically, the addition of the structural uncertainty covariance matrix and the diagonal data covariance matrix (Equation (6)) can indeed represent the stochastic nature of the waveform residual being imposed in the present synthetic settings.

2.4.3 True theory noise covariance matrices and unknown data noise levels

The main difference between the numerical scenarios presented in this section with those in the previous is that we now relax the constraints on data noise levels and treat them as unknowns. We compare the inversion results for the combined covariance matrices, $\boldsymbol{C}_D = \boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^{true})$ versus $\boldsymbol{C}_D = \boldsymbol{C}_d$, where $\boldsymbol{C}_d = \sigma^2\boldsymbol{I}$. Like the previous cases, we run five independent Markov chains to sample the parameter space of MT, $\boldsymbol{m}$, and data noise scaling, $\boldsymbol{h}$, parameters. The recovered ensemble solutions are the mix of post-burning thinned chains, which are shown in Figure 6.

Approved for public release; distribution is unlimited.

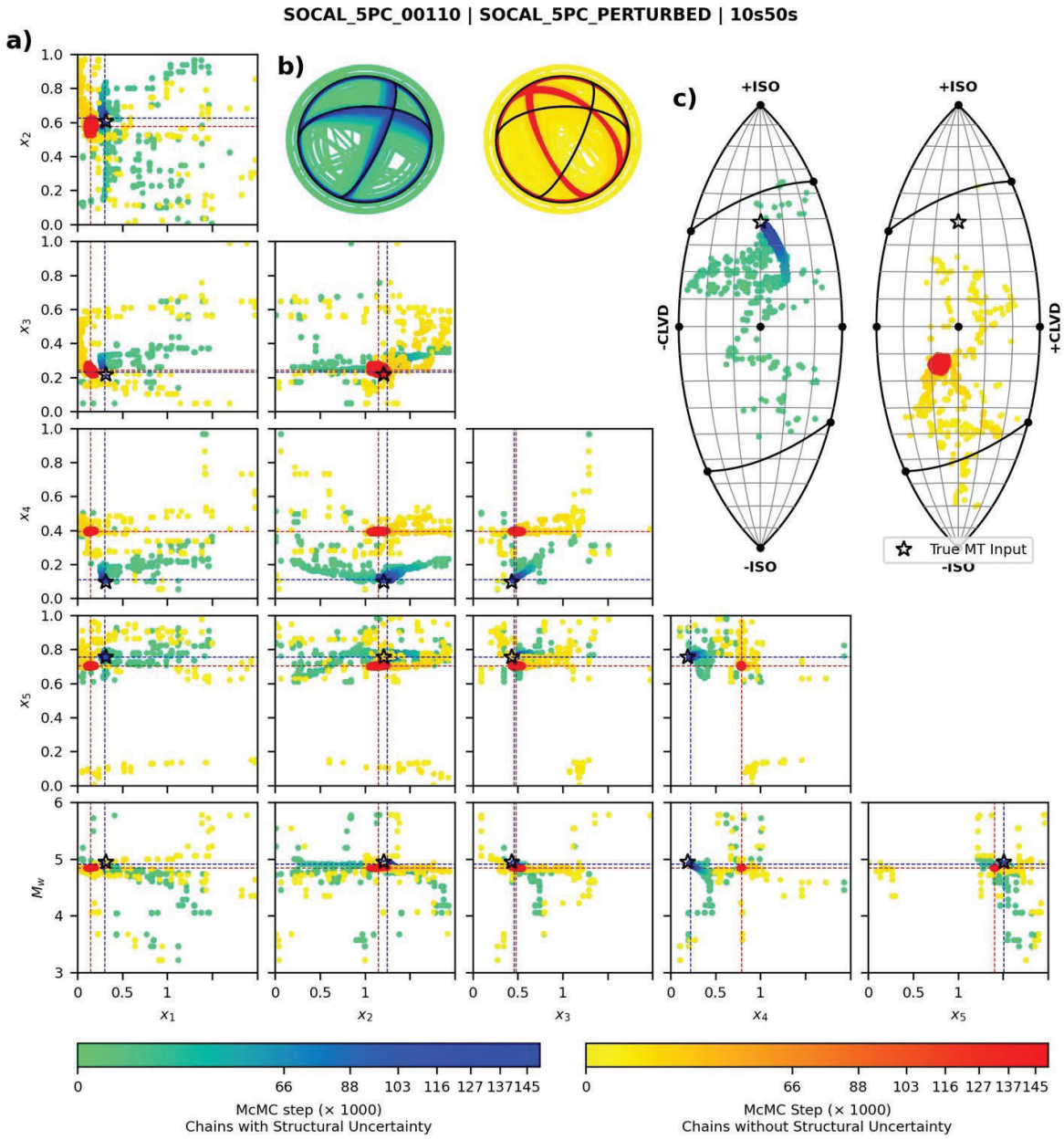

Figure 6. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, *h*, are hyper-parameters to be recovered in the inversion.

(See further description of the scenarios in Section 2.4.3). Synthetic data are filtered in the 10–50 s period band. See the caption of Figure 4 for all other details. We observe similar outcomes from these experiments to the ones presented in the previous section. In all period bands under investigation, incorporating structural uncertainty into the inversion scheme enables robust recovery of the MT solution. When the structural uncertainty is not accounted for, the hierarchical inversion fails to recover the input in shorter period cases: 6–50 s (Figure 4) and 10–50 s (Figure 6), while its effect is minimal for the 20–50 s case (Figure 55).

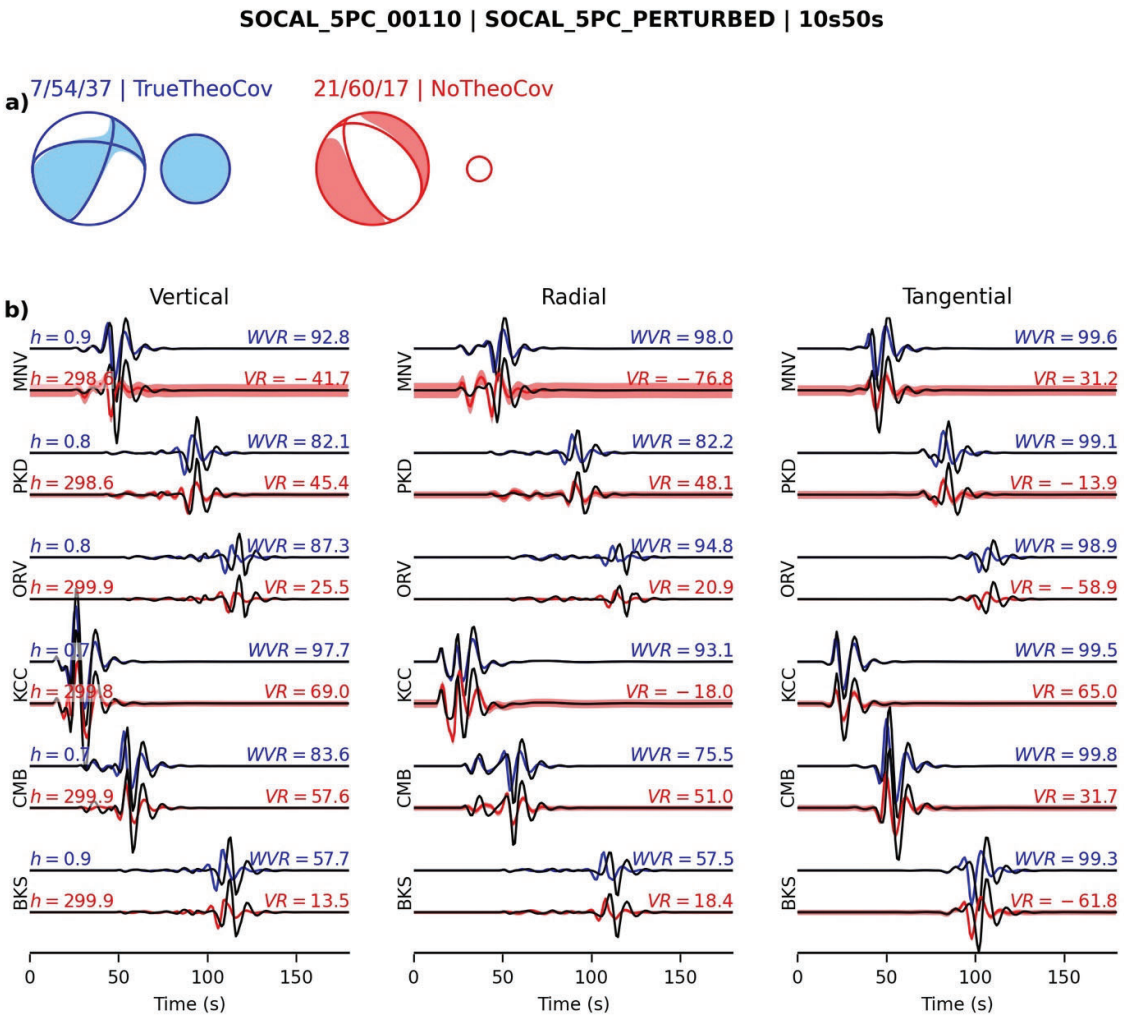

Figure 7. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting for and not accounting for the structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 2.4.3). Synthetic waveforms are filtered in the 10–50 s period band. See the caption of Figure 5 for other details. In addition to details in Figure 5, the numbers on the left side of the vertical traces are the recovered noise-scaling factors, $\mathbf{h_k}$, for the stations, and the thicknesses of these lines are proportional to the recovered noise levels.

Interestingly, the need to incorporate structural uncertainty is further justified when looking at the recovered data noise scaling factors in the mixed chain. When the structural uncertainty is accounted for, not only the MT solution is recovered well (Figure 6), but the noise scaling factors can also be determined well. In Figure 7, the mean scaling factors are smaller but very close to $h = 1.0$, which is the true scaling factor for the period band of 10–50 s, while similar observations for the period band of 6–50 s and 20–50 s are shown in Figure 56 and Figure 57 respectively. With the data noise factor determined from the data, these experiments demonstrate that the hierarchical inversion can effectively decompose the part relating to structural uncertainty from data-noise error in the waveform residual.

2.4.4 Unknown theory covariance matrices

In the above examples, we have made a hypothetical assumption on the prior MT, $\boldsymbol{m}^p$, required to construct the structural uncertainty covariance matrix, that it is the true MT, $\boldsymbol{m}^{true}$. Here, we empirically test the feasibility of recovering the input MT solution with minimal assumptions about the prior MT using the interactive inversion scheme sketched in Section 2.3.5. Here, we provide additional practical details of the approach for this problem. In each iteration, the structural uncertainty covariance matrix is pre-determined with a prior MT and kept unchanged during the iteration. We use the least squares solution as the initial prior MT. We use the mean of the post-burning thinned chain of the previous iteration as the current prior-MT from the second iteration. We also run five independent Markov chains for the same input settings and mix them to obtain the final MT solutions. Each chain consists of five iterations, and each iteration has 150,000 steps.

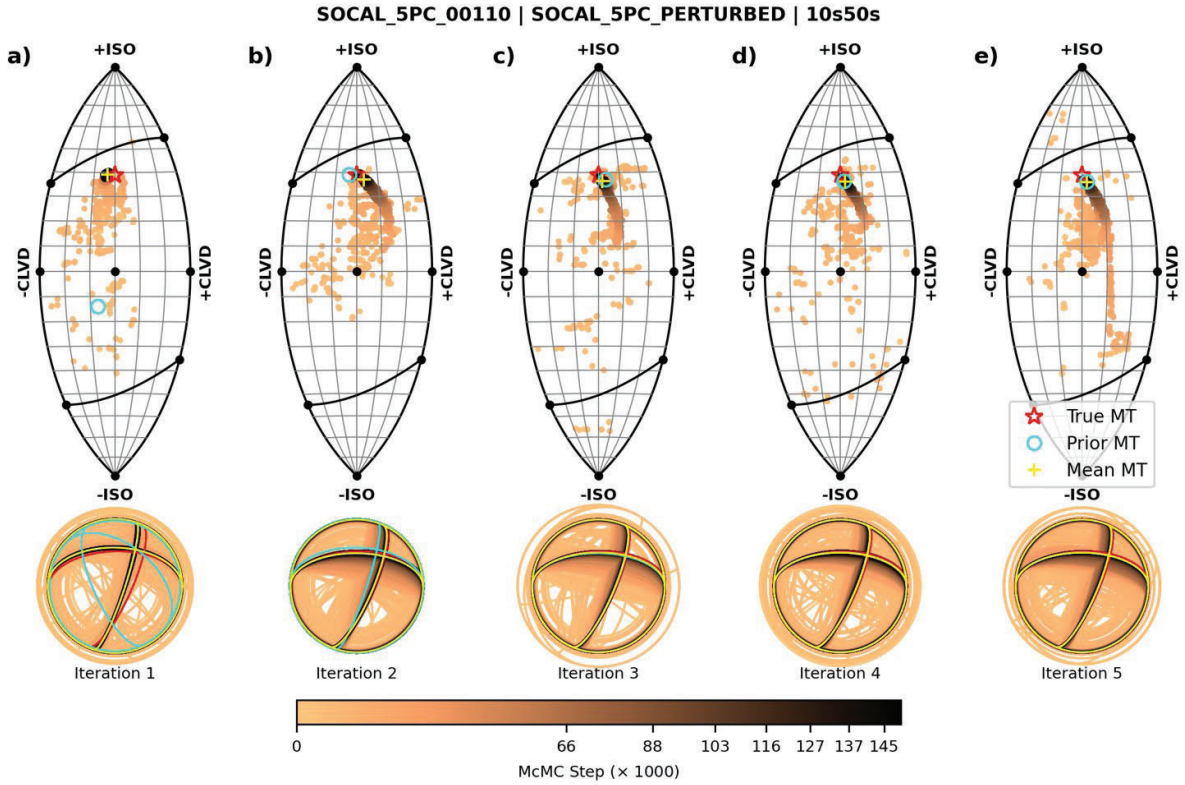

Figure 8. Evolution of McMC chains, the iterative inversion scheme. (See description of the scenario in Section 2.4.4). Synthetic data are filtered in the 10–50 s period band. Progression of MT solutions in the chains is shown by their locations in the lune diagrams representing source types and by double-couple parts via beach balls representing their fault planes. The title indicates the "true" Earth model, $\widetilde{\boldsymbol{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\boldsymbol{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band being considered. Five iterations involved in the scheme are plotted in five sub-figures from a) to e). In each iteration, the prior MT is required to compute the structural uncertainty covariance matrix, $\mathbf{m^p}$ – see Table S1. It is denoted by an open cyan circle in the lune diagram and is plotted in cyan in the beachball diagram. Similarly, the mean MT solution of the post-burning chain is shown by a yellow plus sign and plotted in yellow in the beachball, and the true input MT is shown by an open red star and plotted red in the beachball. As in Figure 4 and Figure 5, we use the power-law of 2/5 to normalize the steps in the reversed copper scale to visualize the evolution of the chains during the burning stage.

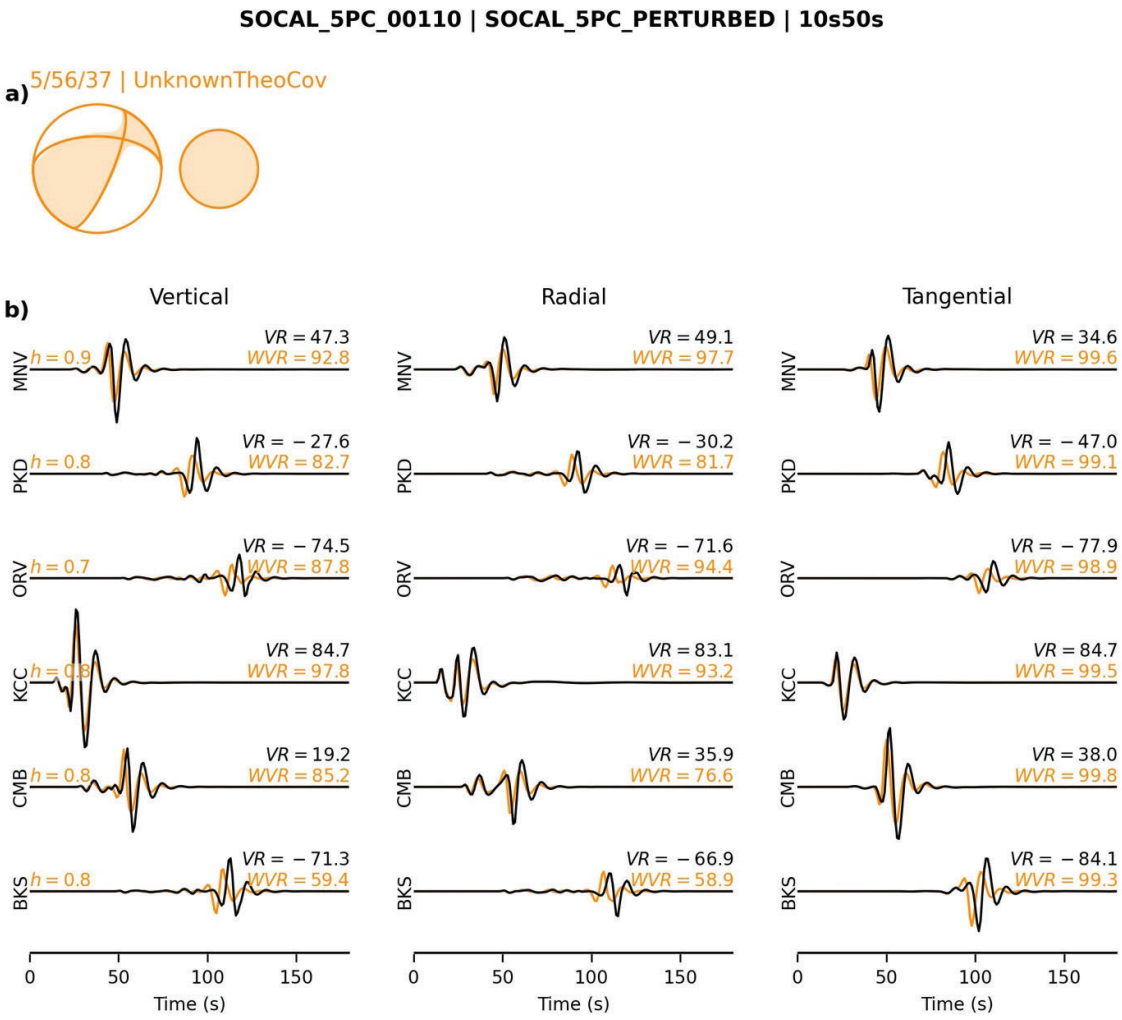

Figure 9. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration.

Synthetic data are bandpass filtered in the 10–50 s period band. The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band being considered. a) Decomposition of the mean MT solution into deviatoric and isotropic parts. The beachball sizes are proportional to their percentages in the full MT solution. Percentages of CLVD, DC, and ISO components of the full MT are displayed on top of the deviatoric beachballs. A filled isotropic ball signifies an explosion, while a non-filled isotropic ball signifies an implosion. b) Waveform fit corresponding to the mean MT solutions. Input synthetic seismograms are black traces, while the recovered waveforms are in copper color. On the right-hand side of each sub-panel are the weighted variance reduction (WVR – defined in Equation (56-a)) for waveforms recovered with structural uncertainty and the variance reduction (VR – defined in Equation (56-b)) between the recovered waveforms and the "fake" data. On the left are recovered data noise scaling factors, h k, for the stations.

Approved for public release; distribution is unlimited.

Figure 8 shows the evolution history of MT solutions for each of five iterations in each of the sub-figures for the period band of 10–50 s (similar results for 6–50 s and 20–50 s are in Figure 58 and Figure 59). The MT solutions quickly approach the true MT input after three iterations even though the initial prior MT is poorly chosen (marked by the open cyan circle in Figure 8a). The pair of prior MT and mean MT solutions overlaps and are very close to the true MT input from the fourth iteration onward. This indicates successful convergence of the inversion scheme. Consequently, the mean MT of the post-burning chains in the fifth iteration will be taken as the final result. Figure 9, Figure 60, and Figure 61 show the MT decompositions and waveform fit for the MT solutions in three period bands under consideration. Note that the data noise scaling factors are also well recovered in all cases and are around 1.

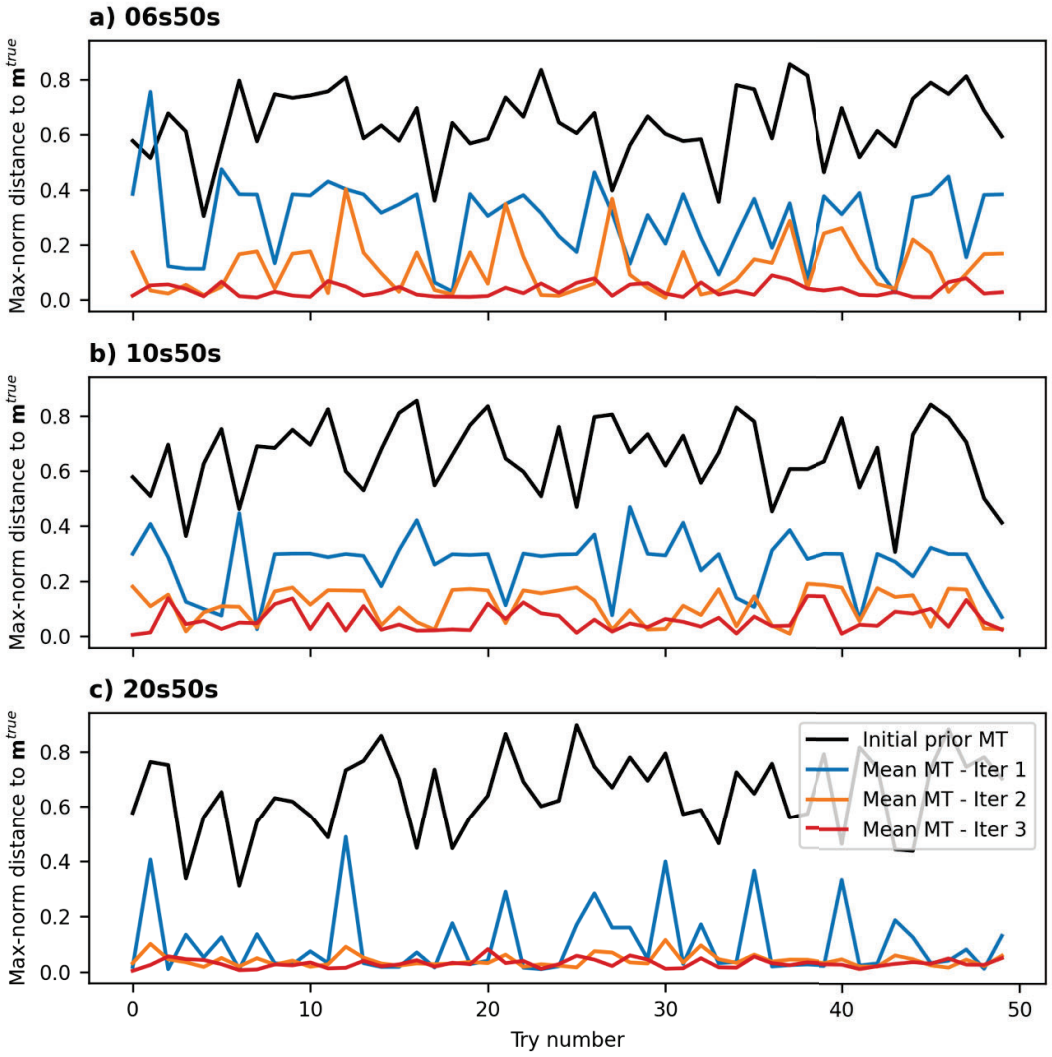

Figure 10. Systematic reduction of max-norm distance (defined in Equation (20)) from the mean MT solution in each iteration to the true MT input for three-period bands: *(a)* 6–50 s, *(b)* 10–50 s, *(c)* 20–50 s. Black lines show the max-norm distance from 50 randomly picked prior MTs. Blue, orange, and red lines show "the distance" of the mean MT solutions from the post-burning, thinned chains in the first, second, and third iterations used as interim, prior MTs for each subsequent iteration.

Furthermore, Figure 10 demonstrates this iterative approach is robust in recovering the true MT in a few iterations with the least dependence on the initial prior MT chosen to start the first iteration. We performed 50 runs of the iterative inversion with randomly selected prior MTs. We define the max-norm distance between the two MT solutions as,

$$\left|\left|\boldsymbol{m}-\boldsymbol{n}\right|\right|_{max} = \max_{\mathrm{i}=1,6} |m_i - n_i|. \tag{20}$$

The M_w parameter in the sixth position is linearly scaled from the range between 3 and 6 to the range between 0 and 1. In Figure 10, because the initial prior MTs are randomly selected, the distance in max-norm with the true MT input is relatively large (black line). After three iterations, the distance reduces systematically and approaches the true MT input. The number of iterations is chosen to be 3 in this study in an ad-hoc manner, which can be adjusted for other applications. Interestingly, quick convergence to the true MT happens in all three period bands under consideration.

2.5 Implications for a volcanic earthquake in Long Valley Caldera

2.5.1 Data Preparation

Armed with the new knowledge based on the synthetic experiments in Section 2.4, we now apply the seismic MT inversion scheme, which incorporates structural uncertainty, to a realistic geological setting in one of the best-studied volcanic earthquake regions – the Long Valley Caldera. More specifically, we apply the method to a pilot event, LV2 – the event in the 1997 Long Valley Caldera sequence, introduced earlier in Section 2.4.1. The recorded waveforms are corrected for instrumental response and band-pass filtered between 20 and 50 s using a Butterworth acausal filter with four poles. In the time window of consideration, the filter is causal to retain the first arrivals from the nearest station, which is just 100 km from the station. The waveforms are then windowed to 180 seconds from the event origin time reported by Minson and Dreger (2008).

To compute synthetic waveforms and estimate the structural uncertainty covariance matrix, we use the SoCal model (Dreger & Helmberger, 1990) as the Earth reference model (i.e., $\widehat{\Omega}$ – see Table S1). At first, the reference model is perturbed by κ = 5% uncertainty (see details in Section 2.3.4), and the event depth in our inversion is fixed at 5 km, as estimated by Minson and Dreger (2008). Because this parametrization could strongly influence the recovered MT solution, especially the non-double-couple components, we examine their influence on our results in the later analysis.

Initially, we run a hierarchical Bayesian inversion on the prepared waveform data to search for a preliminary MT solution and noise amplitudes (see Figure 62). This preliminary MT solution is referred to as $\boldsymbol{m}_0$ (see Table S1) hereafter. Then, we search for the time shifts for the observed waveforms according to the maximum cross-correlation with the predicted waveforms regarding the preliminary MT solution, $\boldsymbol{m}_0$. The time shifts are necessary for the study region that the 1D reference model cannot accommodate due to the crust and upper mantle complexity due to the volcanic systems in Long Valley Caldera. Specifically, we allow the same shift time for vertical and radial for each station, while the shift for the tangential component is independent of the other two. The rationale behind this implementation is that the vertical and radial components are dominated by Rayleigh waves, while Love waves dominate the tangential component. Near the end of the next section, we will further discuss

the robustness of the MT solutions when incorporating structural uncertainty for different time-shift schemes.

2.5.2 Moment-tensor inversion results

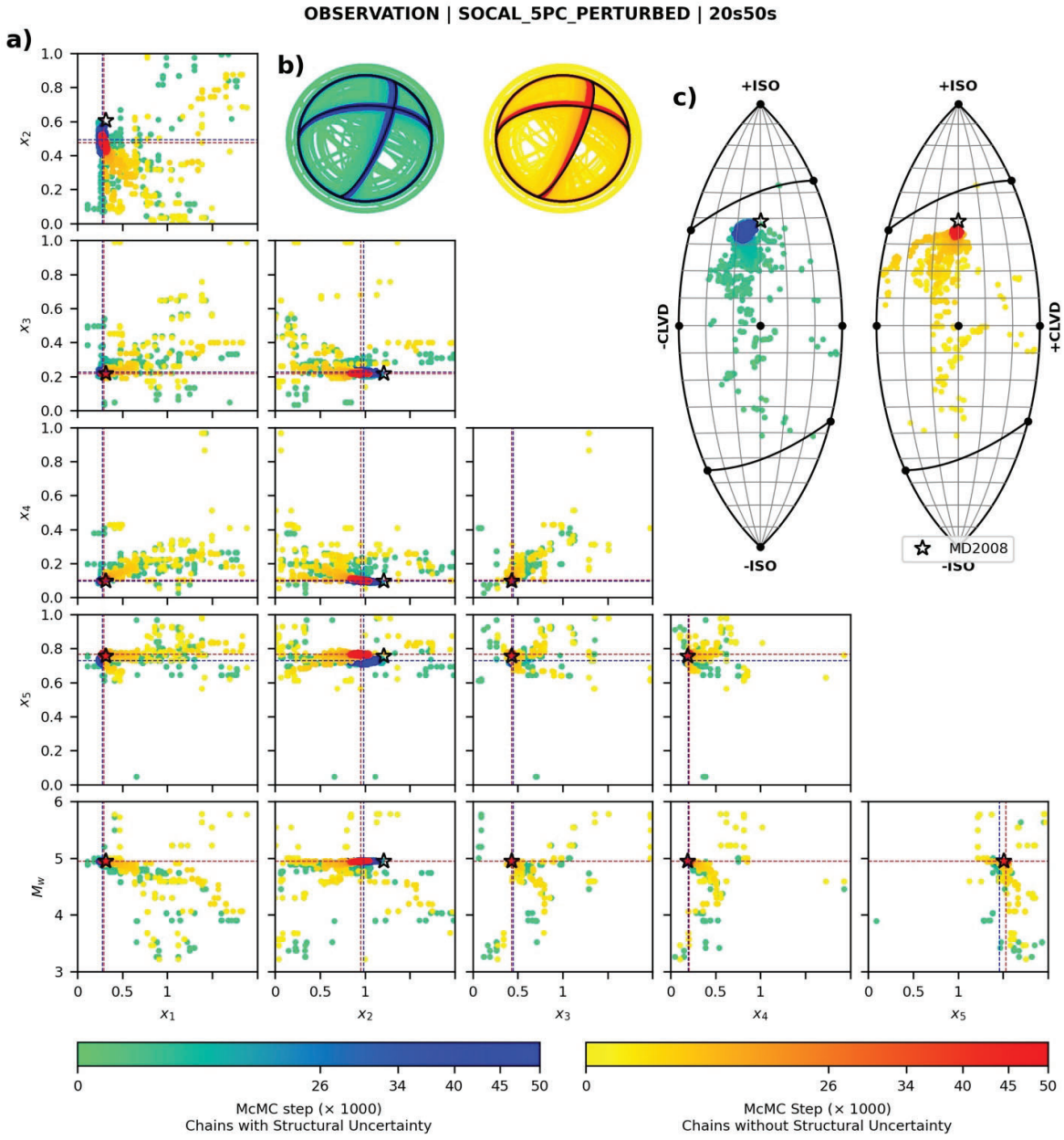

Figure 11. Comparison of inversion results in the two scenarios for real data inversion when accounting for and not accounting for structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 2.4.2). Observed data are filtered in the 20–50 s period band. See the caption of Figure 4 for all details, except we use the open black stars for the MD2008 solution (Minson & Dreger, 2008) rather than being the true MT input in the real application.

We compare the performance of two MT inversion schemes for this event for two cases: (1) using the combined covariance matrix, $\boldsymbol{C}_D$, where the structural uncertainty part is computed with the prior MT solution from Minson and Dreger (2008) – the MD2008 solution, and (2) ignoring the structural uncertainty covariance matrix part. Both attempts are hierarchical in the sense that the noise hyperparameters are also inverted for as unknowns, which is like the synthetic cases presented in Section 2.4.3.

To guarantee the convergence of the Markov chains of 50,000 steps, we performed five independent chains on the same input settings and established that they are well-mixed. We retain the second half of the chains and thin them at 20th sample to obtain the final ensemble of the solutions, including the MT and six noise scaling factors for six stations. Figure 11 compares the evolution of McMC chains in the two scenarios, and Figure 63 visualizes the mean MT solutions via beach ball diagrams with the deviatoric and isotropic parts being decomposed, and the waveform fits corresponding to the mean MT solutions. As shown in Figure 63, the waveform fits for both solutions, when the structural uncertainty covariance matrix is enabled and not, are both of high quality. This is in good agreement with the synthetic experiment for this period band, 20–50 s (see Sections 2.4.2 and 2.4.3), which shows the minimal influence of the structural uncertainty in the MT solution. We note that the data noise scaling factors are systematically lower when the structural uncertainty is in effect. We believe that the waveform differences between the observed and predicted waveforms are effectively separated between data noise and structural uncertainty.

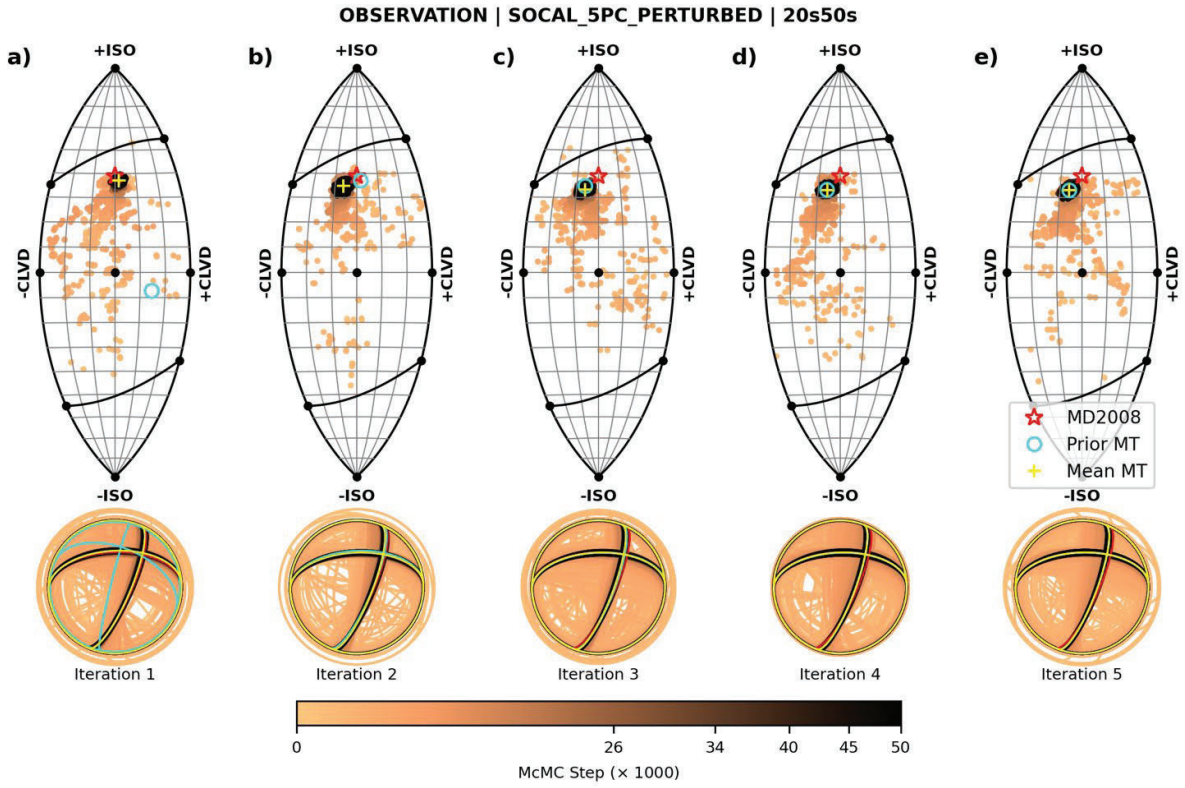

Figure 12. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme for the real data, which is the mean of the post-burning chain in the fifth iteration. Observed data are filtered in the 20–50 s period band. See the caption of Figure 8 for all details.

We next compute the percentages of isotropic, CLVD, and DC components of the mean MT solutions after Vavryčuk (2001). MT solutions of both scenarios demonstrate a high percentage of isotropic components (~35%), which agrees well with Minson & Dreger (2008) (marked as MD2008; see Figure 11). Interestingly, there is a significant percentage of the CLVD component in the solution with the structural uncertainty, while both the solution without structural uncertainty and the MD2008 solution show almost negligible CLVD parts. Subsequently, we present a series of experiments to test the recovery of the non-double couple components against several subjective choices made in the current inversion.

Firstly, we slightly alternate the time shifts of observed waveforms before the inversion and compare them with the solutions presented above. Specifically, instead of using the preliminary MT solution obtained by the Bayesian inversion, we use the least-squares solution (Equation (52)) for that purpose, which results in slightly different time shifts to the observed waveforms (see the comparison of shift times in Figure 62). Figure 64 is like Figure 11 but showing the MT solution obtained with the new shifting, while Figure 65 shows MT decomposition and waveform fit like Figure 63. There is a significant difference in source types between the MT solutions without the structural uncertainty in two attempts to determine the initial time-shifting, highlighting the great sensitivity of the MT solutions on the initial shifts, which are often subjectively determined by researchers. On the other hand, the MT solutions obtained with the structural uncertainty enabled are very similar when comparing Figure 11 and Figure 64. The robustness of the solutions against time shifts could be seen as an advantage of accounting for structure uncertainty inversion to reduce manual interference in the analysis.

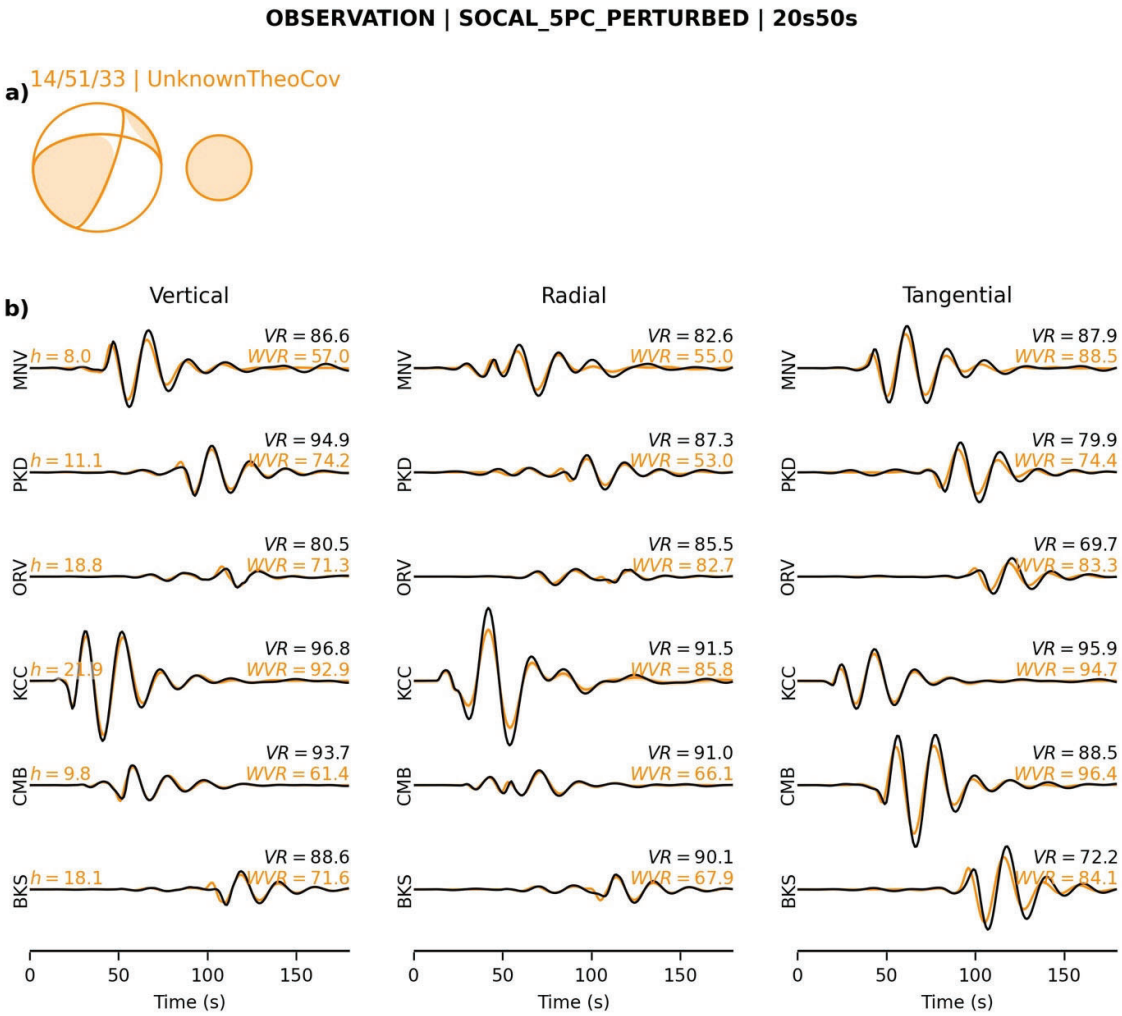

Figure 13. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. Observed data are filtered in the 20–50 s period band. See the caption of Figure 9 for all details.

Secondly, we apply the iterative inversion scheme to relax the assumption of the prior MT being equal to the MD2008 solution to construct the structural uncertainty covariance matrix. In the first iteration, we use the preliminary solution ($\boldsymbol{m}_0$) determined by the Bayesian inversion to compute the initial structure-uncertainty covariance matrix. The inversion scheme updates the prior MT five times at every 50,000 steps. We also run the iterative inversion in 5 independent chains with identical input settings to guarantee the convergence of the obtained solution when the chains are mixed well. In Figure 12, the prior MT, $\boldsymbol{m}^p$, quickly converges to the MT solution, $\boldsymbol{m}$, after two iterations and remains unchanged afterward. Like the synthetic case (shown in Figure 8), we will take the mean MT in the fifth iteration as the final MT solution of the inversion. MT decomposition and waveform fits corresponding to this MT solution are represented in Figure 13. This solution is very similar to the solution with the enabled structural uncertainty in Figure 11, which further confirms the observed non-double-couple in the above cases.

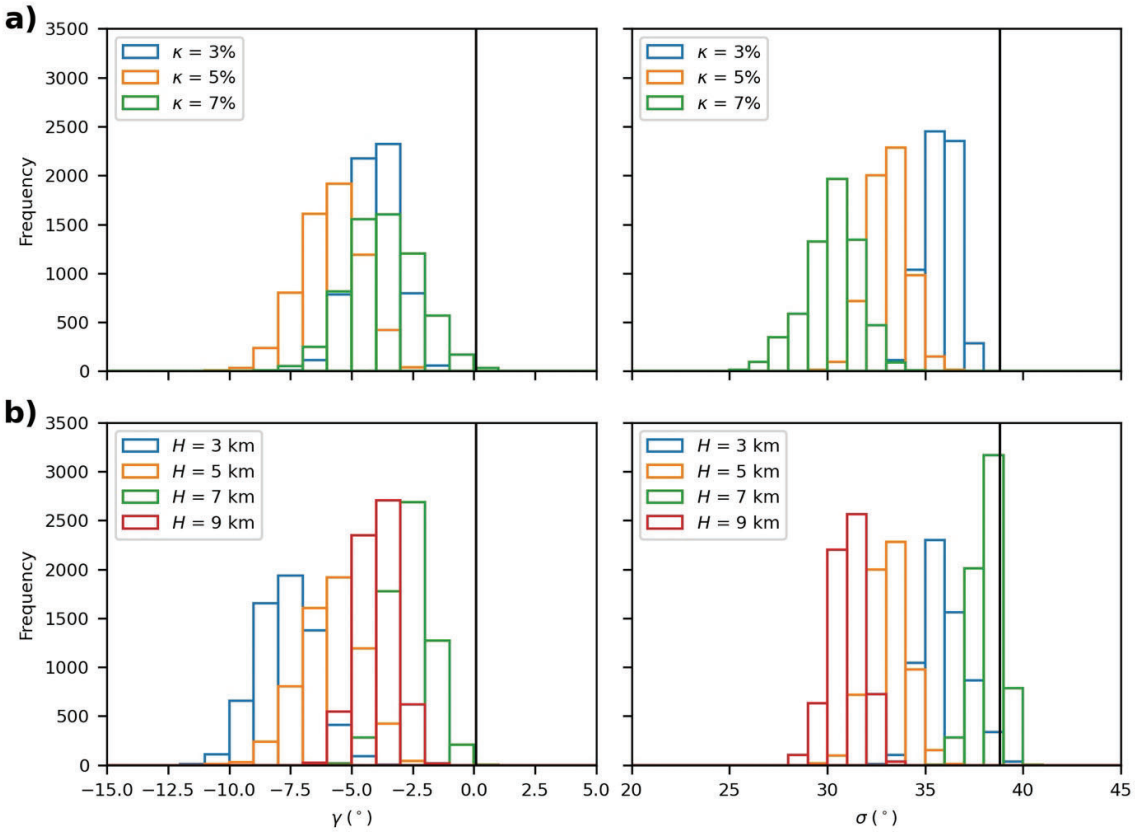

Figure 14. Summary of sensitivity tests of MT solutions of the LV2 event considering (*a*) levels of structural uncertainty and (*b*) event depths. Left/right panels show the histograms of latitude $\boldsymbol{\gamma}$/longitude $\boldsymbol{\sigma}$ in the lune diagram of MT solutions in the well-mixed second half of the fifth-iteration chains, plotted in Figure 12e. Histograms are colour-coded corresponding to the test cases. Black vertical lines mark the values calculated for the MT solution from Minson & Dreger (2008).

Then, we examine the non-linear effect of the pre-assumed level of structural uncertainty, κ =5%, and the event depth, $h = 5$ km, on the recovered MT solutions. Particularly, we repeat the iterative inversion scheme (described in the above paragraph) for two additional levels of structural uncertainty, κ = 3% (Figure 66 and Figure 67) and κ = 7% (Figure 68 and Figure 69). Similarly, we test three different depths: 3 km, 7 km, and 9 km, which involves re-calculation of all ensembles of elementary Green's functions corresponding to these depths. Figure 14 summarizes these results via the histograms of the lune coordinates implying non-DC source types of the MT solutions. For each test case, these coordinates are thinned from the second half of the fifth iteration, which is similar to those plotted in Figure 12e. We see the significant variation of the recovered source types across the test cases, emphasizing the importance of testing the solution's structural uncertainty level and event depth. However, all instances consistently reveal significant positive isotropic components (represented by the lune latitudes) and negative CLVD components (represented by the lune longitudes), although there is a wide confidence interval of ~10° for both parameters. Despite the considerable solution uncertainty, this analysis reveals that the isotropic component is slightly lower than in Minson and Dreger (2008), and the negative CLVD is not negligible.

In summary, the MT solution obtained for this event confirms a significant isotropic component, which agrees with Minson and Dreger (2008). Additionally, there is a non-negligible CLVD (see Figure 11, Figure 12, and Figure 13) in the MT solution when the structural uncertainty is employed in the inversion through several sensitivity tests, even though the definitive percentage of the CLVD component is subject to refinement in future work. This highlights the importance of incorporating the uncertainty of Earth's model in the MT inversion scheme, especially for non-double-couple events.

2.5.3 Short-period waveform examples

Working with short-period waveforms is desirable in the context of incorporating the effects of Earth heterogeneity. The scattering effect could be an issue for short periods, and the 1D Earth model can be inadequate to explain the recorded waveforms. 3D earth structural models will be important in explaining 3D effects. In this proof-of-concept study, we have not approached this problem because the simulations of short-period waveforms in 3D media are computationally expensive. These simulations must be repeated for each of the perturbed models. Therefore, the extension to 3D structural uncertainty is the subject of near-future studies.

2.6 Discussions

This work is at the forefront of attempts to advance seismic source studies by incorporating the Earth model's uncertainty in the full waveform inversion of the point source moment tensor. Here is a summary discussion of some inherent difficulties encountered in the newly developed MT inverse problem framework and some thoughts on the way forward.

Of particular importance is recognizing that the primary goal of incorporating the Earth model's uncertainty in the MT inverse problem is to retain a meaningful interpretation of the recovered MT solution. This interpretation embraces primary sources of uncertainties involved in the problem rather than obtaining an MT solution of absolute accuracy. The accuracy of the MT solution is governed by the quality of available Earth structure models. The widely used approach to circumvent the strict requirement of high-quality Earth structure is to utilize long-period waveforms, for example, 20–50 s, as presented in Section 2.5.1. In the regional settings with complex geology, Hejrani et al. (2017) showed that 3D structure is necessary to recover MT solutions.

In this study, we calculate the structural uncertainty covariance matrix, $\boldsymbol{C}_t(\boldsymbol{m})$, using an ensemble of structural models that are randomly perturbed from an existing Earth model. In our numerical experiments, we assume that the true Earth structure is not known but confined within the known random distribution from the reference model. Under this important assumption, the MT solutions can be recovered at high accuracy across period bands under consideration. It is difficult to use a 1D model with uncertainty to appropriately represent the complication of 3D Earth structures, which may introduce multipathing or scattering effects. Therefore, it is still the subject of future research.

A model ensemble provided by the Bayesian approach (e.g., Kim et al., 2016) can be utilized in seismic source studies by its integration in the MT inversion. However, another important question remains about the most efficient and practical way to incorporate the 3D structural models and their uncertainty in the current MT inversion framework. The numerical solution of the wave propagation in 3D is computationally expensive, but worth exploring because it captures the 3D effects reasonably. In a region with complex geological settings such as volcanic areas, considering composite models accounting for the azimuthal variations in Earth structure could be an interim step before pursuing the complete 3D structure uncertainty.

Compared to previous attempts to perform MT inversion in a hierarchical Bayesian framework (e.g., Mustać and Tkalčić, 2016; 2017), our approach to incorporate empirically estimated structural uncertainty is more computationally expensive due to two critical factors. Firstly, there is a significant expense to pre-compute multiple sets of elementary Green's functions for an ensemble of many perturbed Earth's models. The computational cost is linearly proportional to the number of perturbed models. This factor becomes increasingly considerable in a 3D Earth model environment, in which full-waveform simulations, such as difference method or spectral element method, must be performed multiple times. Secondly, when a noise amplitude is proposed, a considerable portion of steps of the Markov chains requires evaluation of the determinant and inversion of the combined covariance matrix, $\boldsymbol{C}_D$. The key factors influencing the computational time are the frequency in which a data-noise hyper-parameter is proposed, and the size of the covariance matrix defined by the number of points in the observed waveforms. In our examples, the total inversion time is about four to five times more than in cases where the expensive operations can be avoided. Note that the figures are subject to changes in other settings of covariance matrix size or proposal frequency of the hyperparameters.

Here we analyzed an event that was well studied in the past, so it was reasonable to fix the source location in time and space thanks to knowledge from previous studies. However, an essential next step will be to incorporate the possibility to simultaneously constrain the source location and MT solution in the current framework. It is also worth mentioning that in this study, we ignore effects of anisotropy, lateral heterogeneity, or attenuation in the Earth structure model, which could account for an important part of the theory uncertainty. The point-source MT approximation could be another potential source of theory error that is currently ignored. The inclusion of single-force components in the inversion framework is a topic of ongoing research.

2.7 Concluding remarks

This study attempts to incorporate the Earth structure uncertainty into the MT inversion problem within the Bayesian framework. To represent stochastic properties of the residuals between the observed and simulated (theoretical) waveforms, we construct a combined uncertainty covariance matrix as a superposition of diagonal data noise covariance matrices and theory noise covariance matrices empirically estimated from a randomly perturbed model of the existing Earth structure. We demonstrate the particular importance of considering the structural uncertainty when relatively short-period bands are used as input data (6–50 s and 10–50 s) via synthetic experiments. As short period data have been shown to be critical in resolving the properties of shallow earthquakes, it is essential to include the theory noise in future studies. For relatively longer period data (20–50 s), it appears to be less critical to incorporate the structural uncertainty in the synthetic settings. Still, we showed that including the structural uncertainty helps retain a robust MT solution with minimal dependence on subjective choices prior to the inversion. To demonstrate the implications of the proposed approach, we re-examined the MT solution of a well-studied volcanic event in the Long Valley Caldera – event LV2. We confirmed a high percentage of an isotropic component as Minson & Dreger (2008) previously determined. However, we identified a non-negligible negative CLVD component in the MT solution, which was overlooked in the previous analysis that did not consider the structural uncertainty. Therefore, we believe that structure uncertainty should be incorporated in future studies of regional earthquakes, especially those with non-double-couple nature.

3 Full moment tensors inversion using a Bayesian hierarchical approach with 2D structure uncertainty: Implications for DPRK tests 2009-2017

Summary

The seismic moment tensor (MT, a 3×3 matrix) is widely used to represent seismic sources such as tectonic earthquakes, volcanic eruptions, and manmade explosions. It can be decomposed into a double couple (DC), isotropic (ISO), and compensated linear vector dipole (CLVD). To infer the MT, a typical way is to simulate waveforms using an available Earth structure to fit the recorded waveforms through full waveform MT inversion. However, the inversion involves two uncertainty aspects: the data noise in conjunction with the ambient noise and acquisition process and the theory error primarily due to imperfect knowledge of Earth structure (hereafter called structural uncertainty). For shallow events such as underground nuclear explosions, additional theory uncertainty is caused by the theoretical trade-off between pure ISO and vertical CLVD, which plays a significant role. Therefore, a complete uncertainty treatment should consider data noise and theoretical error. Despite recent advancements on this topic that mainly focused on including 1D Earth structure uncertainty, incorporating 2D or 3D structural errors in the inversion is a highly desired, cutting-edge contribution. In this study, we propose an MT inversion scheme within the hierarchical Bayesian framework that accounts for both data noise and 1D and 2D structural uncertainties as a step toward considering 3D structural uncertainty. An advanced Markov chain Monte Carlo sampling algorithm allows us to benefit by exploring the theoretical ISO-CLVD trade-off solution space. Using synthetic experiments with shallow events, we demonstrate the feasibility of our new approach. Subsequently, we use it to perform a MT inversion for five of the six Democratic People's Republic of Korea (DPRK) underground nuclear explosions. Our analysis shows a remarkable similarity among the events' MTs in terms of high ISO and relatively significant CLVD components.

3.1 Introduction

Ground motions emanated from seismic sources, such as earthquakes, volcano eruptions, and underground explosions, propagate through the Earth's interior and are recorded by seismic stations. Recorded ground motions, i.e., seismograms/waveforms, carry rich information about the source and Earth's interior they pass through. Utilizing the seismograms to study the Earth's structure and seismic sources comprise two broad, complementary themes of modern seismology. This study will focus on seismic MT inversion, a procedure to infer the physics of the source via fitting seismological observations with the theoretical predictions. Based on the types of observations, common source inversion methods can be categorized into four groups: (1) methods fitting P-wave first motion polarity; (2) methods fitting P- and S-wave amplitude or their ratio; (3) methods fitting the full waveform; and (4) hybrid methods. We refer the readers to Shang & Tkalčić (2020), who presented a comprehensive review of the advantages and disadvantages of each of these method categories. In this study, we focus on waveform inversion methods within regional distances because the waveforms present sufficient signal-to-noise (S/N) ratio for moderate-size earthquakes, the target events of this study.

A moment tensor (MT) is a widely used source representation, which is a 3×3 matrix that contains nine force couples and combines DC and non-DC components. Furthermore, under the assumption of no mass advection between the source region and the rest of the earth, the faulting within the source region is considered an internal process, so the momentum and angular momentum must be conserved. Therefore, the total force and torques exerted by the equivalent force system are required to be zero at each instance (Aki & Richards, 2002; Takei & Kumazawa, 1994). Consequently, the MT is symmetric and contains only six independent

elements. MT inversion refers to a framework for inferring the MT representation of a seismic source.

Rigorous uncertainty estimation has long been at the frontier of seismic source inversion. A complete uncertainty treatment should consider both data noise involved in the data acquisition and theoretical uncertainty primarily caused by imperfect Earth structure (i.e., structural uncertainty). Data noise has been estimated with different noise models, such as a Gaussian or an exponentially decaying noise model (e.g., Bodin et al., 2012; Duputel et al., 2012), empirical noise models from data residuals (e.g., Dettmer et al., 2007), from synthetic noise series (e.g., Gouveia & Scales, 1998; Piana Agostinetti & Malinverno, 2010; Sambridge, 1999), or model with two-attenuated cosine functions from pre-event ambient noise (Mustać et al., 2018; Mustać & Tkalčić, 2016). Recent studies incorporated 1D structural uncertainty by assuming a Gaussian noise distribution for teleseismic Green's Function (Yagi & Fukahata, 2011), estimating a covariance matrix from linear perturbation of Green's functions (Duputel et al., 2014), or estimating a covariance matrix from synthetically generated Green's Functions with Earth structure perturbations (Hallo & Gallovič, 2016). These studies involved remarkable efforts to handle data noise and theoretical error separately. Most recent advances have been made to jointly treat theoretical error data noise and theoretical errors for the MT inversion (e.g., Phạm & Tkalčić, 2021; Vasyura-Bathke et al., 2021). These works provide a pathway to estimating 1D structural error considering the overall average structural effect for all stations. In some complex regions, however, 2D structural errors considering specific source-receiver path differences can better constrain source parameters. In this study, therefore, we propose an improved MT inversion scheme within the hierarchical Bayesian framework that accounts for data and 1D & 2D structural uncertainties, which will be applied to determine the source physics of five DPRK underground nuclear tests. There will be follow-up work focusing on applications of the method in different geological settings.

3.2 Methodology

3.2.1 Bayesian inference of seismic moment tensors

We consider an MT inverse problem with the observation, $\mathbf{d} \coloneqq \{d_i\}$, consisting of $M = n_s \times 3$ seismograms (n_s is the number of stations with three components each). When assuming the source location and time are known precisely, which is a reasonable assumption for manmade explosions, the forward problem is written in the linear form,

$$g_i(\mathbf{m}) = \mathbf{G}_i\mathbf{m}. \tag{21}$$

The predicted waveform $g_i(\mathbf{m})$ and the pre-computed Green's function $\mathbf{G}_i \in \mathbf{R}^{M\times 6}$ correspond to the observed seismogram d_i, given an MT $\mathbf{m} \in \mathbf{R}^6$. The likelihood function for component i of the observed data d_i given the moment tensor $\mathbf{m}$ is the probability,

$$p(d_i|\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^N|C_i|}} exp\left(-\frac{(g_i(\mathbf{m}) - d_i)^T C_i^{-1}(g_i(\mathbf{m}) - d_i)}{2}\right) \tag{22}$$

where N is the number of data points in each component d_i, and their uncertainty covariance matrices are $\mathbf{C} \coloneqq \{C_i\}$. For simplicity, C_i is assumed to be independent among all components, even if they belong to the same station. Therefore, the aggregated likelihood function for all components is the product of all individual components' likelihood functions,

$$p(\mathbf{d}|\mathbf{m}) = \prod_{i=1}^{M} p(d_i|\mathbf{m}) \tag{23}$$

and the logarithm of the likelihood for all seismograms to be observed simultaneously is,

$$\log\left(p(\mathbf{d}|\mathbf{m})\right) = -\frac{1}{2}\sum_{i=1}^{M}\Big(N\log(2\pi) + \log|C_i| + (g_i(\mathbf{m}) - d_i)^T C_i^{-1}(g_i(\mathbf{m}) - d_i)\Big) \tag{24}$$

Based on Bayes' theorem (Bayes & Price, 1763), the logarithm of the posterior probability is,

$$\log\left(p(\mathbf{m}|\mathbf{d})\right) = \log\left(p(\mathbf{d}|\mathbf{m})\right) + \log\left(p(\mathbf{m})\right) - \log\left(p(\mathbf{d})\right) \tag{25}$$

where we assume an uninformative prior, $p(\mathbf{m}) = c$, and $p(\mathbf{d})$ is the evidence of the data, which is also a constant. By dropping all additive constants which do not affect the relative landscape of the posterior probability, the log-probability in Equation (25) used for sampling is finally written as,

$$\log\big(p(\mathbf{m}|\mathbf{d})\big) = -\frac{1}{2}\sum_{i=1}^{M}\Big(\log|C_i| + (g_i(\mathbf{m}) - d_i)^T C_i^{-1}(g_i(\mathbf{m}) - d_i)\Big). \tag{26}$$

The covariance matrix C_i is the entry for consideration of uncertainties in the inversion problem, which can be solely data noise, empirical theory uncertainty, or their combination. Firstly, data noise is mainly caused by background ambient noise at the recording site and instrumental noise in the data acquisition. Secondly, the theory uncertainties, or uncertainties relating to the forward problem, are any source of errors due to theoretical approximations in the forward problem. It is reasonable to assume that the largest contribution to the theory error is due to our imperfect knowledge of the Earth's interior structure, which is also referred to as structural errors in this report.

The data noise and structural error have been studied separately using a specific covariance matrix C_i, as mentioned in the introduction. For the complete consideration of the uncertainty in an MT inversion problem, the covariance matrix accounts for both sources of uncertainties (e.g., Duputel et al., 2012, 2014). Recent advancements treating data noise and structural errors jointly in MT inversion have been made by Phạm & Tkalčić (2021) and Vasyura-Bathke et al. (2021). They both proposed the use of a combined covariance matrix,

$$C_i = C_i^d + C_i^t, \tag{27}$$

where C_i^d and C_i^t are covariance matrices accounting for the data noise and structural error, respectively. The structural covariance matrix, C_i^t, is estimated empirically by perturbating a 1D Earth model using a Monte-Carlo simulation, which is computationally expensive, especially when 3D Earth is considered in a future attempt. Furthermore, the empirical estimation of the structural covariance matrix requires subjective choices for scale and parameterization of the Earth model perturbations, which are currently subjected to future research.

Here, we propose a simplified treatment of the structural uncertainty, in which the structural errors are accounted for using site-specific time shifts (more details to be considered in Section 3.2.2). In a further simplification, data noise on each component is assumed to be uncorrelated,

and the covariance matrix becomes diagonal when the signal-to-noise ratios of inverted waveforms are small. To reduce the number of noise parameters and avoid a wide search range for them, we follow the approach presented by Phạm & Tkalčić (2021) to parameterize the covariance matrix in Equation (26) as,

$$C_i = h \cdot \left(\sigma_i^{ref}\right)^2 \boldsymbol{I}, \tag{28}$$

where σ_i^{ref} is the reference noise strength for each component pre-computed from the 1-hour pre-event ambient noise, and h is the station-specific noise level hyper-parameter whose dynamic range is set to [0, 400].

Additionally, at regional distances, there is an ambiguity between pure ISO and vertical CLVD mechanisms in long-period waveform MT inversion for shallow events such as manmade explosions (e.g., Chiang et al., 2014; Ford et al., 2012). The reason is that the long-period waveforms at regional stations are most sensitive to the radiated energy along the equator of a focal sphere, where the pure ISO and vertical CLVD radiate similar Rayleigh waves. The ISO and CLVD cannot be differentiated well using only long-period regional waveforms. This results in a theoretical trade-off between them. This theoretical ISO-CLVD trade-off is another significant theory error in MT inversion for shallow sources, but it cannot be treated with the covariance matrix. Therefore, an advanced sampling algorithm is required to explore this trade-off in the solution space, which is also an essential component of this study.

3.2.2 Accounting for Earth model uncertainty by station-specific time shifts

Accounting for the influence of uncertainty of Earth structure on inverted moment-tensor solutions is still at the forefront of seismic source research. Initial attempts have been made to estimate the covariance matrix for structural errors (e.g., Duputel et al., 2014; Phạm & Tkalčić, 2021; Vasyura-Bathke et al., 2021), but the construction of structural uncertainty covariance matrices remains complicated, with many parameters that must be subjectively tuned. Alternatively, Tkalčić et al. (2009) and Hallo & Gallovič (2016) noted that a significant source of GFs' uncertainty is due to the misalignment between predicted waveforms and observations when using a 1D layered model to represent the medium between the source and receivers. In particular, Figure 15 shows the waveform comparison for an ensemble of velocity models generated by random perturbations from the reference model with 5% uncertainty. Given the same MT, all seismograms share a similar shape and amplitude and are shifted horizontally at long periods. Changes in amplitude and waveform shape are more significant, but time-shifting remains the largest source of uncertainty.

In this study, we introduce a simplified scheme to account for the misalignment by allowing the time shift at each station to be an unknown in the inversion. The time shift parameters represent the uncertainty in the Earth model. In particular, the station-specific time shifts between observed and predicted waveforms act as a proxy for 2D structure error and relax them as free parameters during the inversion. Even though the time shifts on different components of the same station might be different (e.g., due to anisotropy), we assume three components of a station share the same time shift to simplify the problem. The time shift also takes care of the error from inaccurate origin times of the events.

Mathematically speaking, the time-shifting of a waveform can be done analytically as,

$$g_i'(\mathbf{m}) = F^{-1}\left[F[g_i(\mathbf{m})] \cdot e^{-i\omega\tau}\right], \tag{29}$$

in which F, F^{-1}denote forward and inverse Fourier transforms, respectively. τ is a station-specific time shift parameter. In this work, the τ is bounded by [-10, 10] to avoid cycle skipping for waveforms filtered between 20 -50 s.

Therefore, the complete model to invert for is defined as$[\mathbf{m}, \mathbf{h}, \boldsymbol{\tau}]$ where $\mathbf{m} = \left(M_{xx}, M_{yy}, M_{zz}, M_{xy}, M_{xz}, M_{yz}\right)$ parameterizes a complete moment tensor, $\mathbf{h} = (h_1, h_1 \cdots h_{n_s})$ parameterizes station-specific data noise strengths, and $\boldsymbol{\tau} = (\tau_1, \tau_1 \cdots \tau_{n_s})$ are the station-specific time shifts. Finally, the log-probability in Equation (26) is rewritten as

$$\log\left(p(\mathbf{h}, \mathbf{t}, \boldsymbol{\tau}|\mathbf{m})\right) = -\frac{1}{2}\sum_{i=1}^{M}\left(\log|C_i| + (g_i'(\mathbf{m}) - d_i)^T C_i^{-1}(g_i'(\mathbf{m}) - d_i)\right). \quad (30)$$

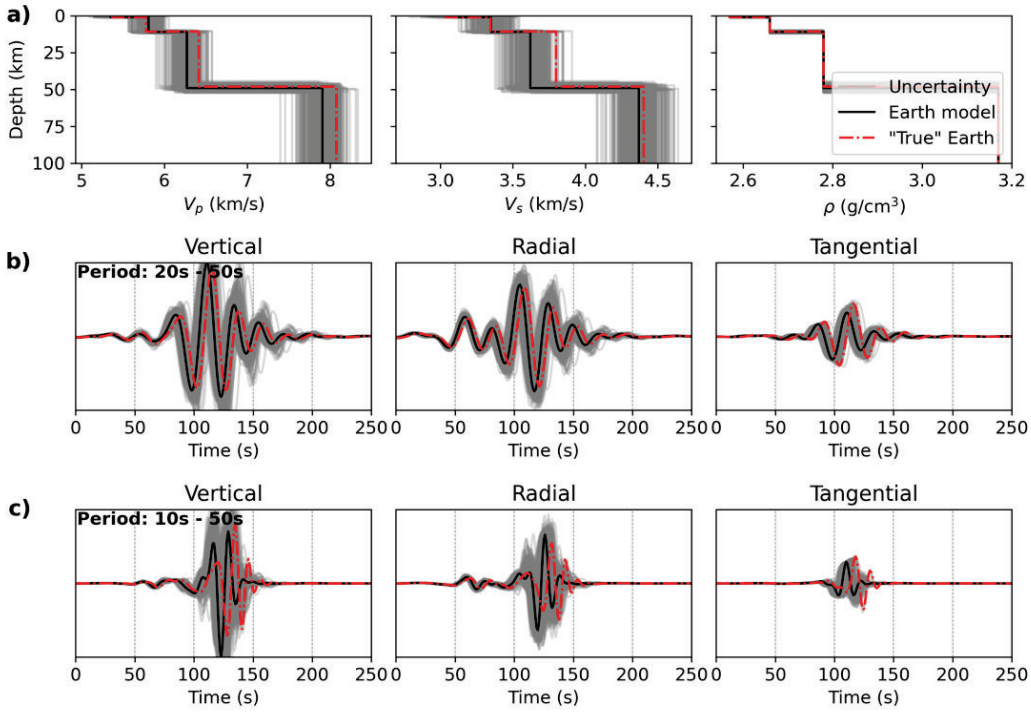

Figure 15. Waveforms for different 1D Earth structures. (a) shows P-wave and S-wave velocity and density of 'True' Earth (red), an estimated structure (black), and randomly perturbed structures (gray) from the estimated structure perturbed by 5% uncertainty (Phạm & Tkalčić, 2021). (b) and (c) show the waveforms at the same station (i.e., MDJ in Figure 18) corresponding to structures in panel (a) with the same source in 20-50 s and 10-50 s period bands, respectively.

3.2.3 Exploring the parameter space using affine invariant ensemble samplers

The posterior probability distribution in the parameter space can be numerically estimated using Markov chain Monte Carlo methods (Sambridge & Mosegaard, 2002). In this study, we use an advanced sampling method with ensemble samplers proposed by Goodman & Weare (2010), demonstrating a remarkable efficiency boost compared to the traditional Metropolis-Hastings sampling method (Hastings, 1970; Metropolis et al., 1953). Instead of a single walker used in MHA, this algorithm utilizes simultaneously evolving random walkers starting with

different initial positions (*i.e.*, vectors in N-dimensional parameter space) to explore the solution space. These walkers communicate and exchange information.

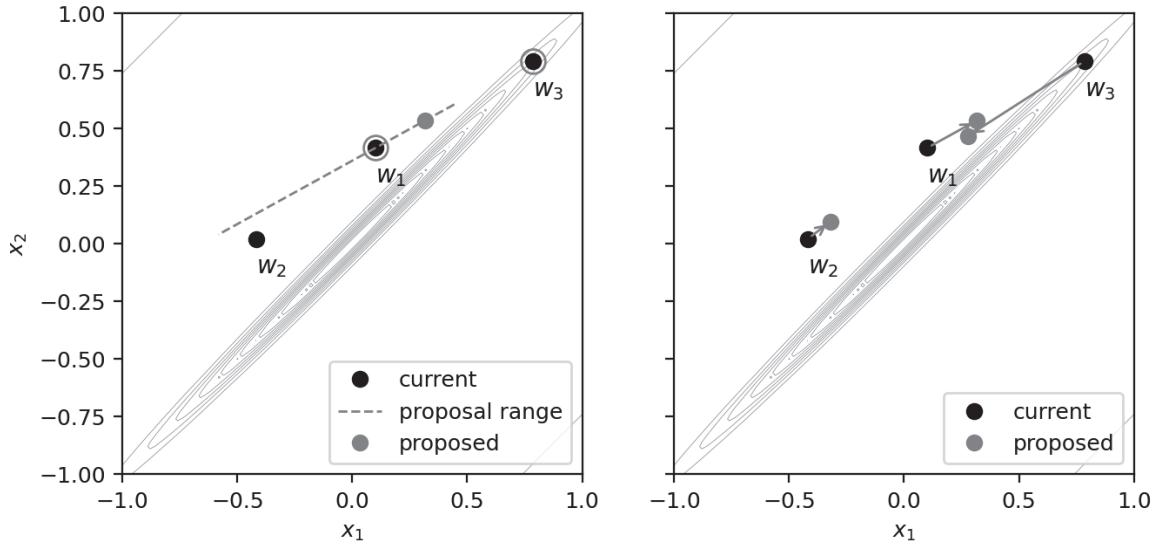

Figure 16. Schematic demonstration in two-dimensional parameter space of the stretched move used in the affine invariant McMC (Goodman & Weare, 2010). The background shows the contours of the probabilistic distribution to be sampled. On the left panel, black dots mark the current positions of three walkers. The grey dot is a proposed move for the walker $\mathbf{w_1}$, with a randomly chosen partner $\mathbf{w_3}$. The dashed grey line shows the range of possible proposals for the next move of $\mathbf{w_1}$. On the right panel, grey dots are proposed to move for all three walkers from their current positions, which will be accepted or rejected randomly.

Let's consider an ensemble of **N** walkers to explore the parameter space. Goodman & Weare (2010) suggested the 'stretch move' proposal scheme. In each iteration, a motion for the walker $\mathbf{m}_i$ is proposed in two steps. First, a partner is chosen from the complementary walkers in the ensemble, say $\mathbf{m}_j$. Then, the proposed move is drawn randomly along the line connecting the two walkers,

$$\mathbf{m}_i' = \mathbf{m}_j + Z \cdot (\mathbf{m}_i - \mathbf{m}_j). \tag{31}$$

In Equation (31), Z is a random, positive number drawn from a probability distribution $g(z)$ in the $[1/a, a]$ interval,

$$g(z) \propto \begin{cases} \frac{1}{\sqrt{z}} & \text{if } z \in \left[\frac{1}{a}, a\right] \\ 0 & \text{otherwise} \end{cases}. \tag{32}$$

The parameter 'a', where $a > 1$, is the only parameter to adjust the performance of the stretch move. Furthermore, $a = 2$ has empirically been found to be an optimal choice in many large-scale inverse problems (Foreman-Mackey et al., 2013; Goodman & Weare, 2010). This proposed move of the walker $\mathbf{m}_i$ is accepted based on a probability involving the probabilities of the current coordinate and the proposed move in the parameter space,

$$q = \min\left(1, Z^{N-1}\frac{p(\mathbf{m}_i')}{p(\mathbf{m}_i)}\right). \tag{33}$$

The stretch move is iterated for other walkers in the ensemble before proceeding to the next iteration.

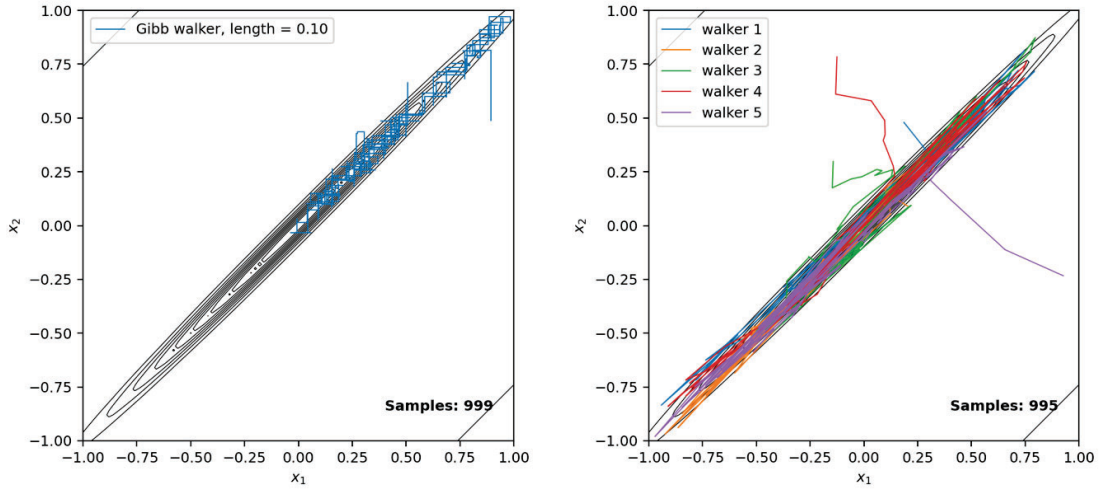

Figure 17. Comparison of sampling efficacy between (left panel) the traditional Metropolis-Hasting method and (right panel) the ensemble samplers with stretched moves (Goodman & Weare, 2010). The background contours show the target probability distribution. Each colored trace represents the trajectory of a walker. There are 1000 random samples drawn in both cases.

The ensemble samplers, designed as above, possess the affine invariant property, whose performance is not affected by an affine transformation of the coordinates. Such transformations are often caused by the linearly dependent parameters, which leads to a highly anisotropic probability distribution, as demonstrated in Figure 16 and Figure 17. As shown in Section 3.3, there is a strong linear dependence among the diagonal component elements of an MT due to the ambiguity of isotropic and CLVD components of seismic sources located at critically shallow depths. This phenomenon causes significant difficulty in resolving shallow seismic sources in previous studies (e.g., Chiang et al., 2014; Ford et al., 2010a, 2012), where additional constraints such as teleseismic P wave polarity are brought to bear apart from regional seismic waveforms primarily consisting of surface wave arrivals.

In the following numerical experiments and applications to data, we will demonstrate the advantages of the ensemble samplers (Goodman & Weare, 2010) for the MT inversion problem of non-double-couple components in shallow seismic sources.

3.3 Synthetic Experiments

In a series of synthetic experiments, we present the inherent challenges in resolving the non-double couple components for critically shallow seismic sources using regional seismic records (< 1200 km). We then show that the advanced sampling method using ensemble samplers has clear advantages in mitigating the challenges.

3.3.1 Experiment configuration

To demonstrate the feasibility of this approach for MT inversion for very shallow sources, we use the realistic configuration of an underground nuclear explosion conducted by the DPRK on 3 September 2017 at 03:30 UTC at the Punggye-ri test site. The DPRK conducted six nuclear tests in 2006, 2009, 2013, 2016 (twice), and 2017. Among all reported tests, the 2017 event is the most significant, with a body-wave magnitude (m_b) of 6.3 reported by the United States Geological Survey (USGS). Figure 18 shows the event location and seven stations, including three stations from the New China Digital Seismograph Network (NCDSN), two stations from the Global Seismograph Network (GSN), one station from the French Global Network of Seismological Broadband Stations (GEOSCOPE), and another one from Korean Seismic Network (KIGAM). They provide excellent azimuth coverage from 370 km up to 1100km. A four-layer 1D velocity model MDJ2 (Ford et al., 2009) is used to compute predicted seismograms to fit the simulated waveforms. The source depth for such shallow explosions is challenging to resolve using only long-period data (Ford et al., 2009). Wang et al. (2018) proposed a depth of 450 m by combining space-borne geodetic and seismic records. Dreger et al. (2021) assumed a depth of 800 m for joint regional long-period waveforms and regional and teleseismic P-wave polarity data. To simplify the problem, the depth is fixed at 0.5 km in this numerical experiment. We take the MT solution from Alvizuri & Tape (2018) to generate the simulated waveforms.

The simulated waveforms are produced with both data noise and 2D structural error. We use the wavenumber integration method (Bouchon, 1981) as implemented in Computer Programs in Seismology (CPS; Herrmann, 2013) for Green's Functions calculation. Then, we follow the method proposed initially by Jost & Herrmann (1989), which was then improved by Minson & Dreger (2008) to compute noise-free simulated waveforms. The noise-free simulated waveforms are filtered by a 4-corner acausal Butterworth filter in period bands of 20-50 s. Real three-component pre-event ambient noise at each station is preprocessed in the same way as the simulated seismograms. Then each noise time series is added to its corresponding noise-free seismogram. Based on Equation (28), the reference noise strengths σ_i^{ref} are pre-computed from the ambient noise. All input noise levels, $h_1, h_2 \ldots h_7$, are equal to one. To introduce the 2D structural error, we shift the data with station-specific times (see Table 1), i.e., positive time shifts for three stations in China and two stations in South Korea, and negative time shifts for two stations in Japan. Consequently, the MDJ2 model is faster for paths through the continental crust and slower for paths across the oceanic crust. Therefore, the simulated waveforms include uncorrelated data noise and 2D structural error.

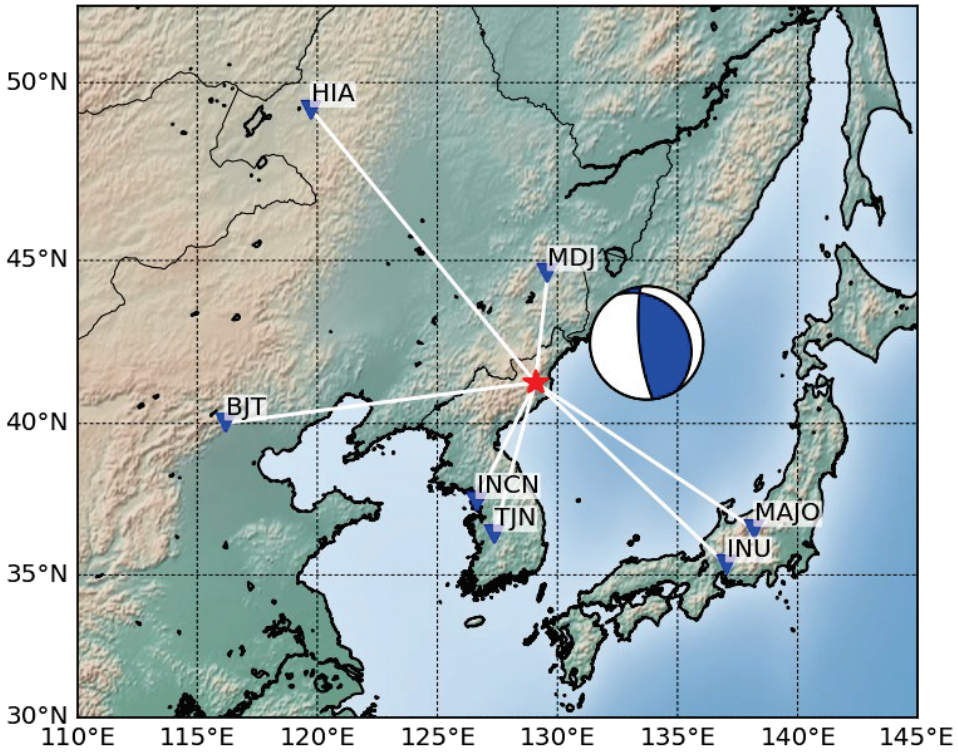

Figure 18. Map of region showing the DPRK2017 test location (red star) and seven seismic stations (blue triangles). The input MT is shown by the beachball from Alvizuri & Tape (2018). The decomposition of this source is 63.7% ISO, 6.4% CLVD, and 29.8% DC. The moment magnitude, $\boldsymbol{M_w}$, is 5.21.

Table 1. True station-specific time shifts (unit: seconds) were used for the numerical experiment of MT inversion for the DPRK2017 test

Explosion	IC.MDJ	IC.BJT	IC.HIA	IU.INCN	KG.TJN	IU.MAJO	G.INU
2017	4.0	3.7	4.0	2.0	1.5	-4.5	-5.5

3.3.2 The inherent ambiguity of non-double-couple source types for shallow seismic sources

Manmade seismic sources, including nuclear explosions, are conducted near the Earth's surface (mostly shallower than 1 km deep) (e.g., Yang et al., 2021). At these shallow depths, seismic waveforms recorded by regional stations (< 1200 km) are dominated by surface waves and have no sensitivities to the vertical force couple. This explains the high similarity between waveforms generated by a purely isotropic source and a vertical CLVD at 0.5 km depth, shown in Figure 19, which is meant to reproduce the comparison by Kawakatsu (1996). The waveform similarity leads to the non-uniqueness of non-double-couple components for shallow seismic sources (Figure 20), which was investigated in previous studies considering nuclear explosions (Chiang et al., 2014; Ford et al., 2010b, 2012). Here we demonstrate that careful consideration of data noise levels and efficient sampling methods can mitigate the inherent ambiguity of external sources when the uncertainty level is low. However, it is worth noting that complete resolvability for mechanisms of non-double-couple sources remains an inherent challenge. The

presence of significant data noise or structure uncertainty is likely to dominate the weak sensitivity to ISO and CLVD mechanisms.

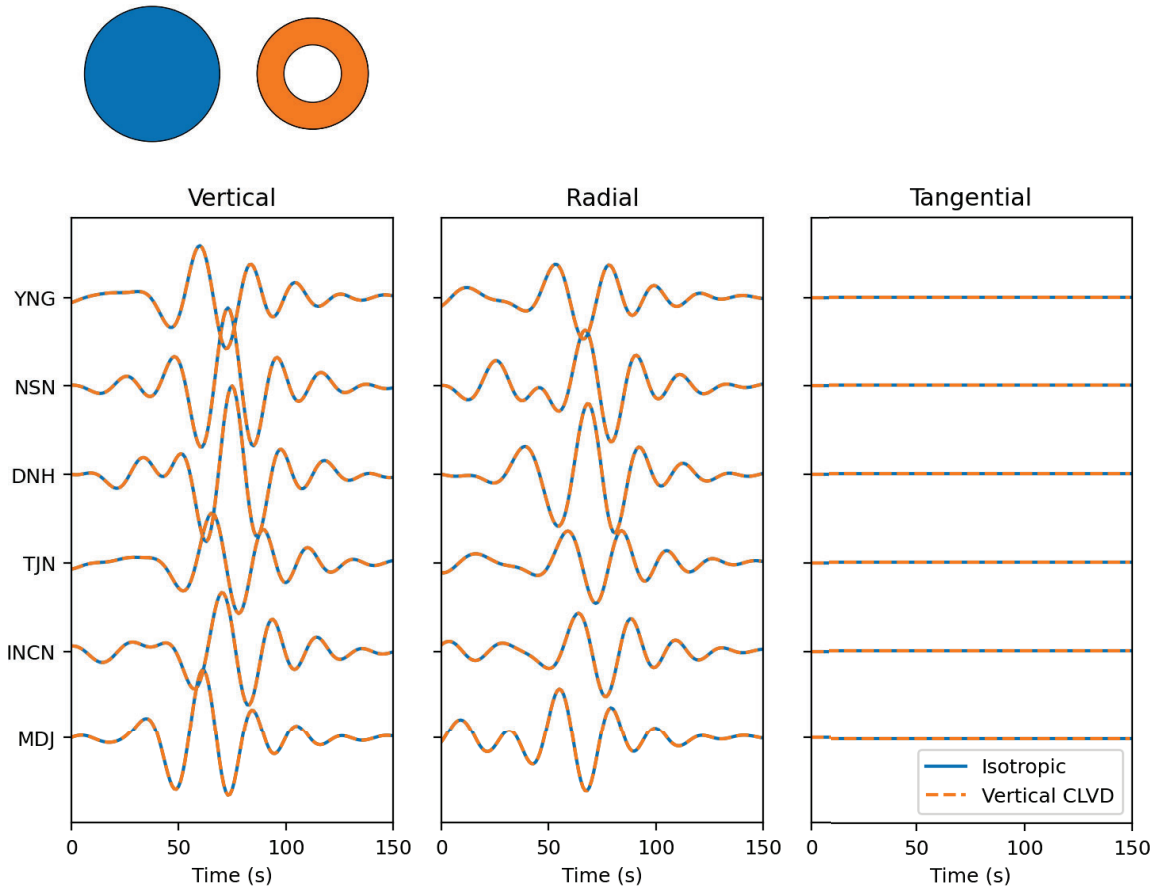

Figure 19. A waveform comparison of critically shallow isotropic and vertical CLVD moment-tensors recorded by regional stations (< 1200 km). Waveforms are bandpass-filtered between 20–50 seconds, and both sources are placed at 0.5 km depth.

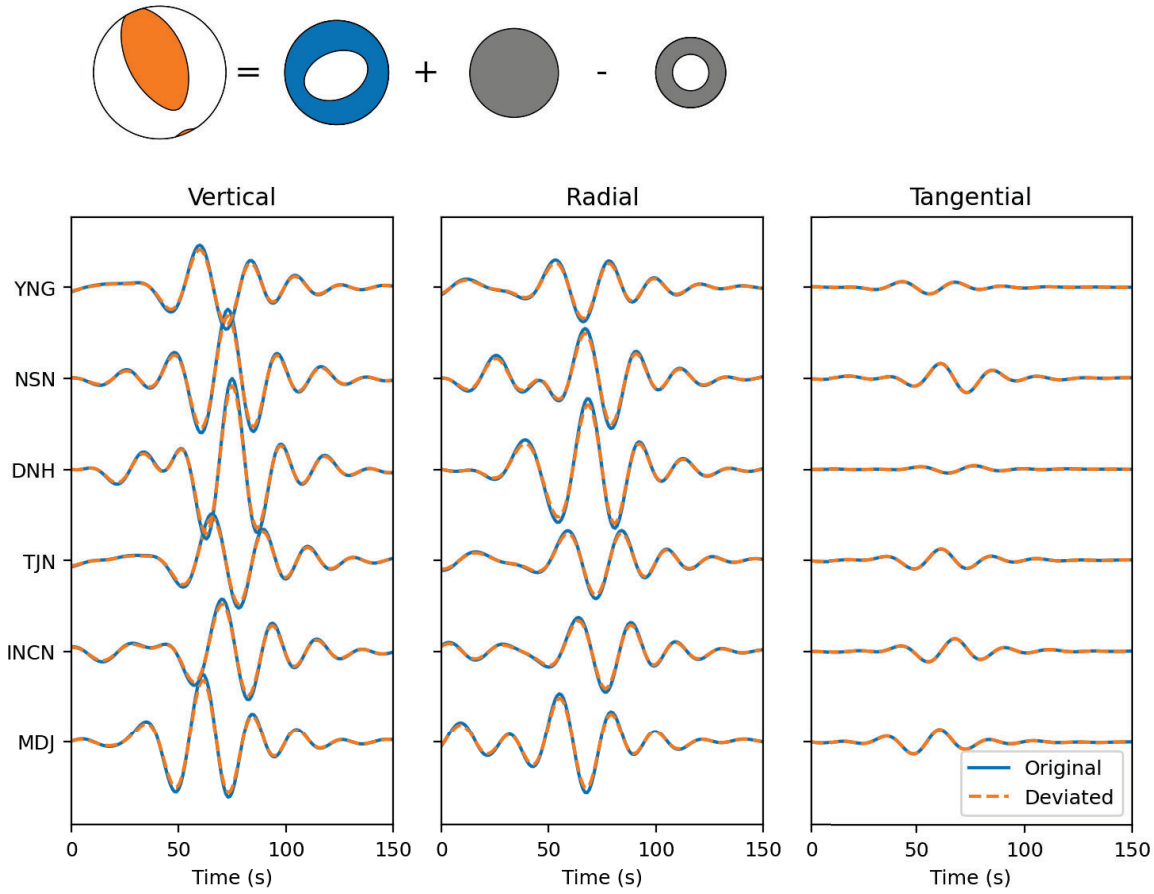

Figure 20. The ambiguity of non-double-couple components of shallow seismic sources. The deviated seismic MT, whose tension quadrants are filled with orange, is calculated by adding a significant isotropic component and subtracting an equivalent amount of vertical CLVD components.

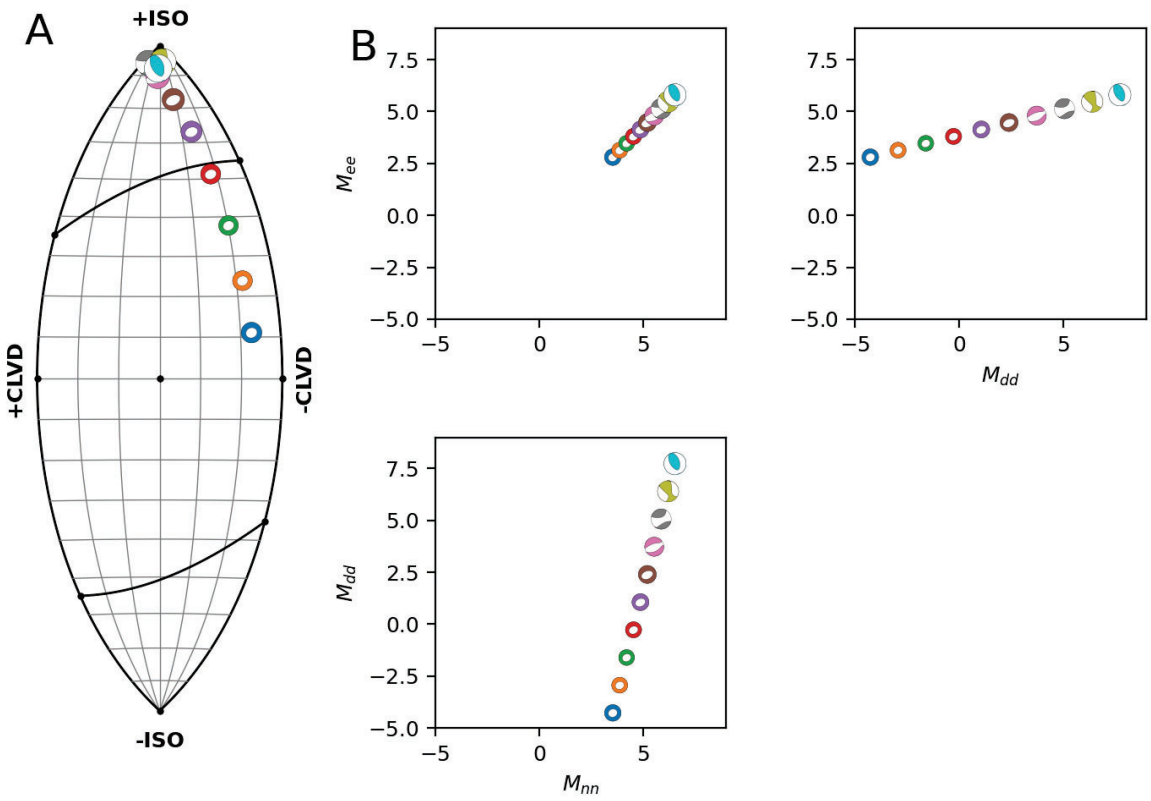

Figure 21. The variety of seismic MTs producing almost identical seismic waveforms.

3.3.3 Inversion results from a shallow synthetic explosion

To demonstrate the resolvability of shallow seismic sources, we set up a synthetic experiment with six regional seismic stations monitoring an explosion at a depth of 0.5 km (Figure 18). Simulated waveforms used as the input for the MT inversion algorithm are shown in Figure 25. The results of the inversion are summarized in Figure 22, Figure 23, and Figure 24. As can be seen in Figure 22, the true input MT (denoted by the magenta start) is successfully recovered by the mean of the converged chains after discarding the burn-in periods (see Figure 22c).

The linear correlation between M_{xx}, M_{yy} and M_{zz} in Figure 22a is the result of the theoretical trade-off between pure ISO and vertical-CLVD components for shallow sources, discussed in Section 3.3.2. As can be seen in the lune diagram of Figure 22c, the algorithm with ensemble samplers effectively explores the parameter space. Initially, a wide variety of source types is explored (black dots) in the lune diagram. Then the samplers converge to a stripe in the lune diagram to explore the ISO-CLVD trade-off with higher posterior probabilities (dark brown dots). The samplers then converge to a small area corresponding to the highest posterior probability (bright copper dots), whose mean of these most probable solutions is denoted by the magenta star.

The station-specific noise levels (Figure 23) and time shifts (Figure 24) are also recovered by the ensemble samplers. As mentioned before, all noise levels are fixed in the current numerical experiment, and the recovered mean noise levels for all stations are close to the desired value. Besides, the recovered time shifts (Figure 24) are also generally close to the input time shifts (Table 1). The recovered noise parameters and time-shift parameters are also found to be uncorrelated. A good waveform fit (>VR=99%) between observed and predicted waveforms

using the mean MT and time shifts is obtained in Figure 25. Therefore, we conclude that the inversion framework using regional stations is successful in this setting.

Now, to investigate the impact of noise on the resolvability of shallow seismic sources, we design six different sets of simulated waveforms corresponding to high, intermediate, and low SNR. The SNR is defined by $SNR = 20log_{10}(\frac{A_s}{C \cdot A_n})$ where A_s, A_n are the root mean squares of simulated waveforms and pre-event ambient noise amplitudes, respectively. C is a component-based coefficient that scales the ambient noise to generate waveforms of specific SNR. This results in six datasets of different SNR from 5 – 30, with increments of five units. Because of the band limited nature of seismic waveforms, the data noise is correlated and should be considered in the noise model in an inversion problem; we argue that assuming uncorrelated noise is, however, reasonable when the SNR is low.

As shown in Figure 26, the shallow source can be recovered in the case of high SNR ($SNR = 30$). The MT converges to a small area in orange, which is close to the true source (magenta star) on the lune diagram with slight uncertainty. As the correlated data noise becomes more significant (i.e., SNR=25, or 20), the solution uncertainty becomes more significant, and the theoretical tradeoff due to shallow depths become more significant despite their good waveform fits, i.e., VR=99.7 for SNR=25 and VR=99 for SNR=20. Therefore, there is still a chance to retrieve the source parameters by only considering uncorrelated noise for intermediate-size earthquakes whose data SNR are usually above 20. In the case of low SNR data ($SNR = 10$, or $SNR = 5$), our algorithm with the assumption of uncorrelated noise cannot obtain the acceptable MT. The solution uncertainty is substantial, as shown by orange dots in Figure 26e and Figure 26f, and the mean MT is far away from the true one. Besides, the theoretical trade-off between ISO and CLVD remains unsolved due to the inappropriate noise estimate. This always happens when noisy stations are involved, or the earthquake is small.

To demonstrate the ability of our algorithm to constrain the source parameters for different shallow events at different depths, we designed two numerical experiments. Firstly, we also use the MT from Alvizuri & Tape (2018a) as the true source but put it at different depths from 0.5 – 3.0 km with a 0.5 km increment. The SNR of simulated waveforms is fixed to 25, which is typical for intermediate-size earthquakes. The recovered MT of the sources at different depths are plotted in Figure 27. This algorithm constrains its source parameters well for shallow events at depths from 0.5 – 3.0 km. The recovered MTs are very close to true MT in Figure 27. Secondly, we fix the source depth at 0.5 km but consider five different non-DC sources, including ISO-dominated sources and CLVD-dominated sources, as the true sources to validate the resolvability of the trade-off between ISO and CLVD. The five different MTs are shown by magenta stars in the lune diagram and beachballs in Figure 28. The SNR in this experiment series is also fixed to 25. The MT solutions for these five sources show the trade-off between ISO and CLVD. However, each mean MT is still close to its true source (the magenta star). Even though the same noise level (i.e., SNR) and 2D structural error are introduced into the waveforms, the uncertainty of MT solutions for different sources appears different. For the near-deviatoric source in Figure 28d, the recovered source parameters contain considerable uncertainty, including an apparent trade-off between ISO and CLVD.

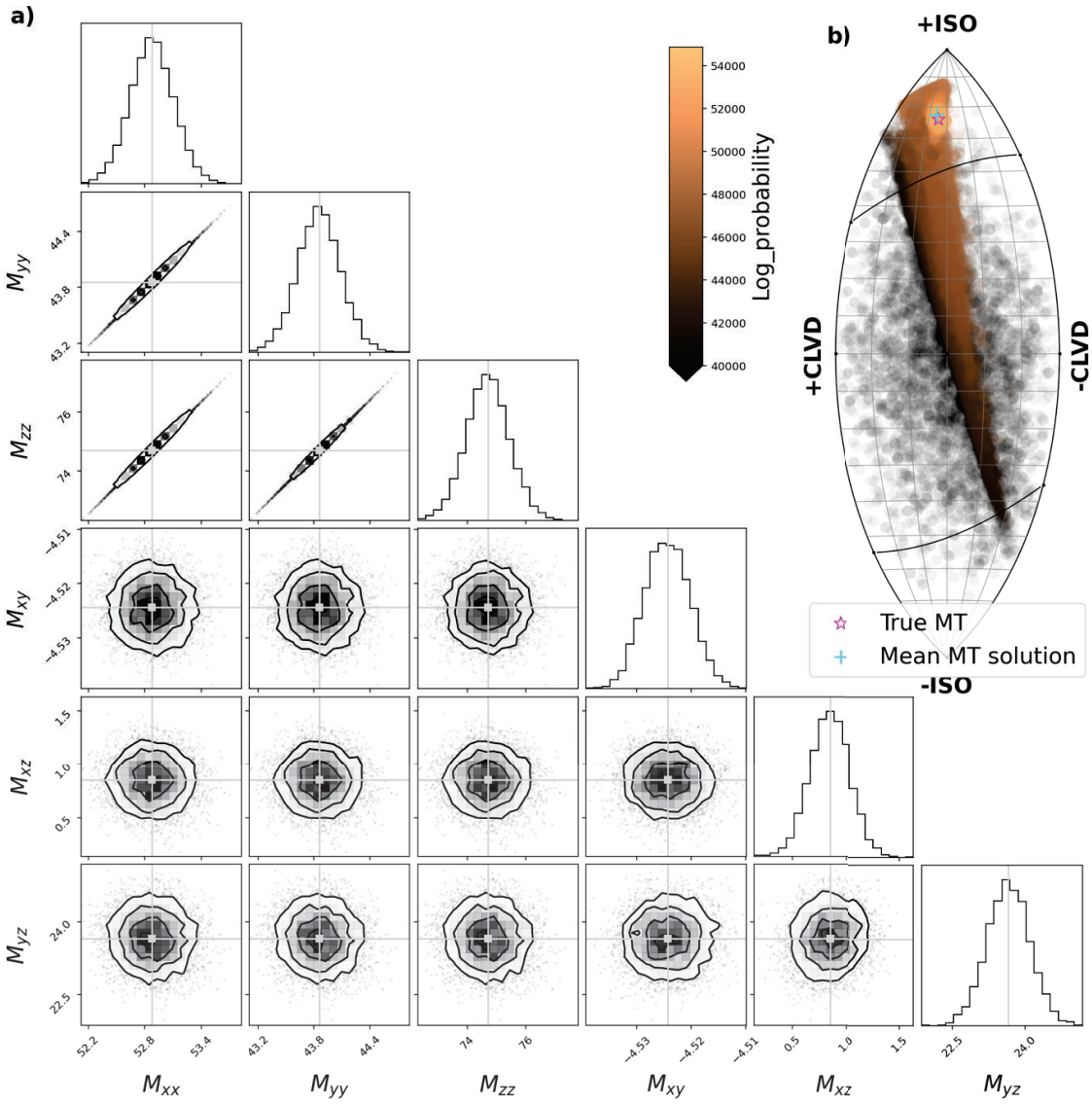

Figure 22. MT inversion results considering uncorrelated data noise and 2D structural error with a hierarchical Bayesian framework for the synthetic scenario. The source depth is 0.5 km. Synthetic data are filtered in the 20-50 s period band. (a) Each sub-panel shows a pair of parameters in 6 MT parameters only in the converging stage of the inversion. The unit of MT parameters is 10^{15} Nm. The light gray lines show mean MT parameters. (b) The lune diagram of source types (Tape & Tape, 2012) shows the evolution of every 5 MT solutions during the inversion. The magenta star marks the source type of the true MT input. The cyan cross marks the mean MT solution in the converging stage. The color bar is used for log-probability in Equation (30). All log probabilities under 4.0×10^4 are black to visualize the later stage of the inversion better.

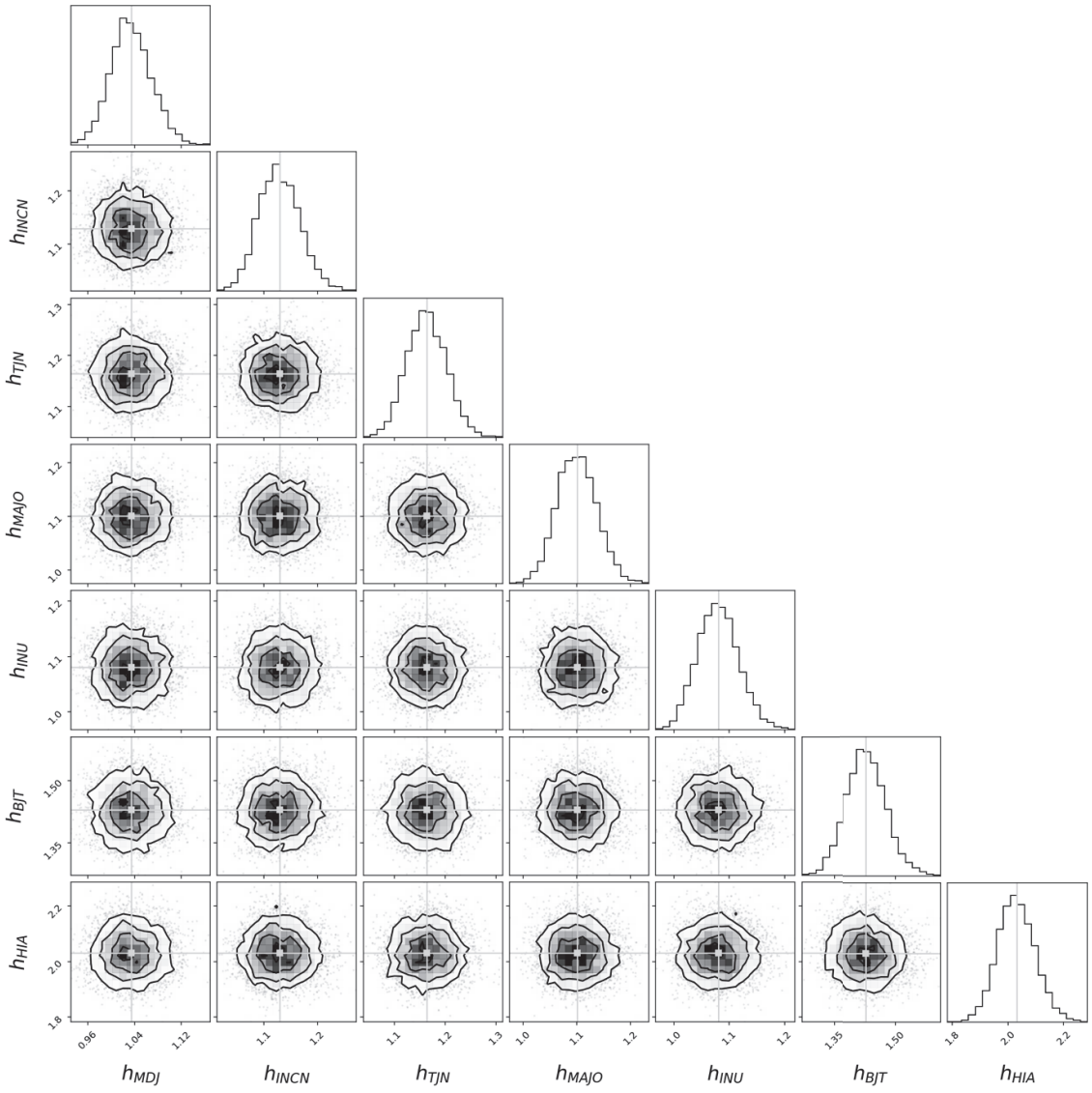

Figure 23. Recovered station-specific noise parameters. Each sub-panel shows a pair of parameters in seven noise parameters in the converging stage of the inversion. The light gray lines show mean noise parameters.

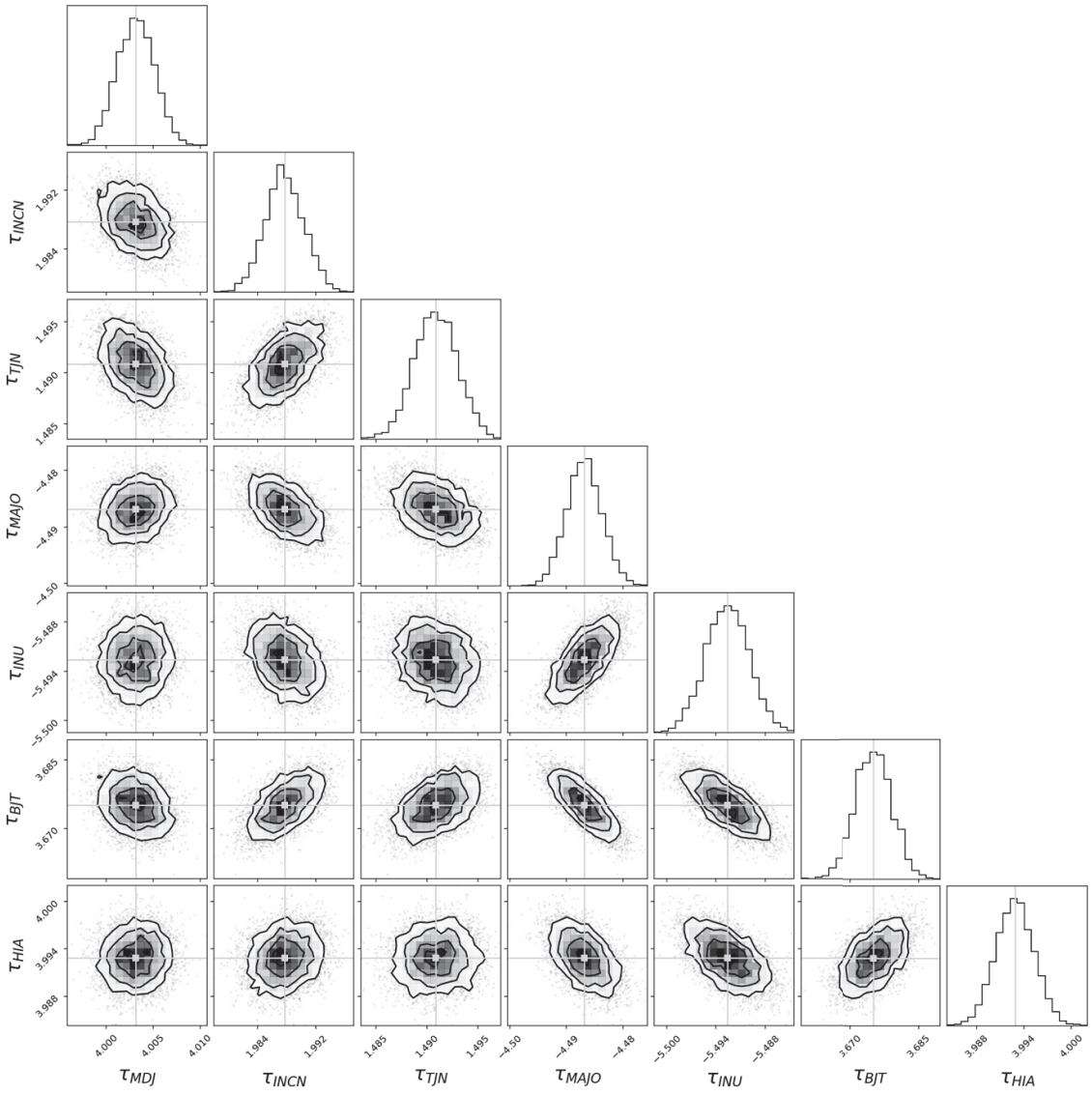

Figure 24. Recovered station-specific time shifts. Each sub-panel shows a pair of parameters in seven-time shift parameters in the converging stage of the inversion. The light gray lines show mean time shifts.

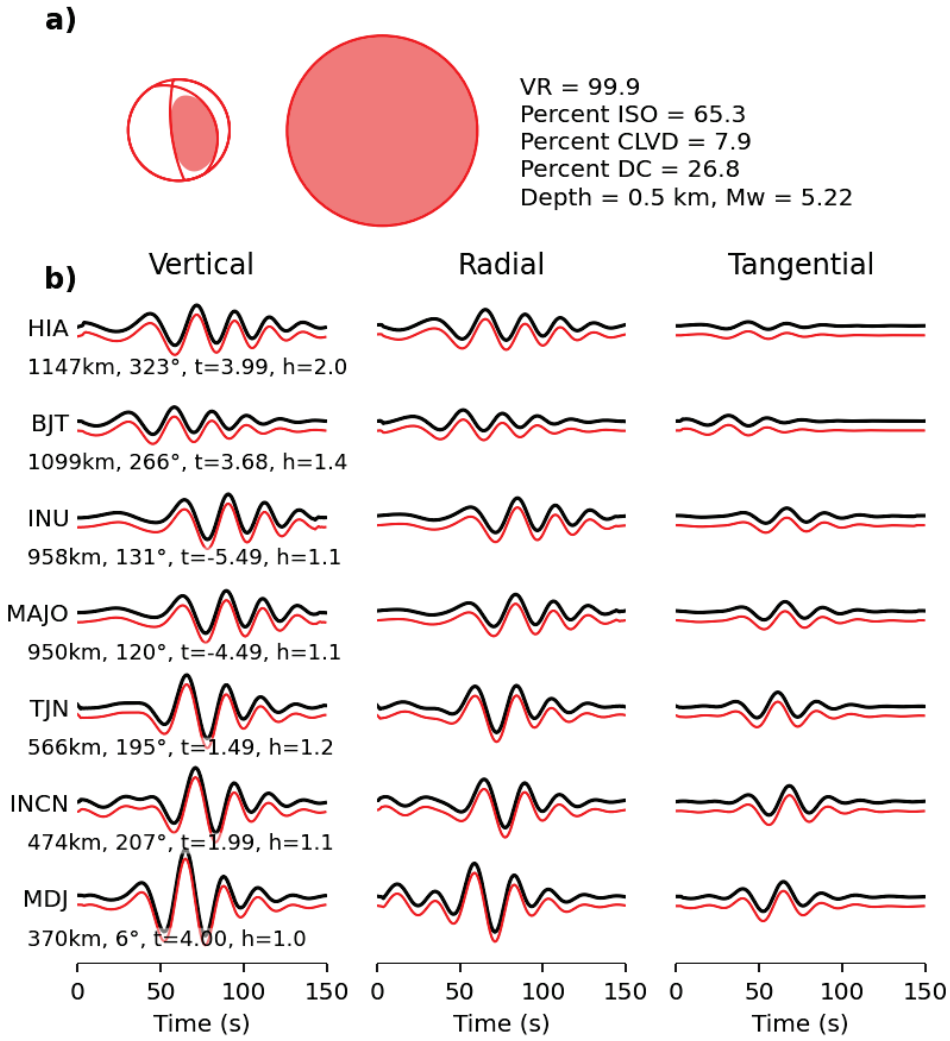

Figure 25. MT decomposition and waveform fit at seven stations for the synthetic scenario. (a) Decomposition of MT solution into deviatoric and isotropic parts. The beachball sizes are proportional to the percentages. (b) Waveform fit between 'observed' (black) and predicted (red) waveforms from the MT solution is shown in (a), which is measured by the variance reduction (VR, defined in Equation 56b of Phạm & Tkalčić, 2021). Waveform are shifted vertically for clearer comparison. The numbers on the lower-left corner of the left panel are source-receiver distance, azimuth, recovered station-specific time shift, and noise parameter, respectively.

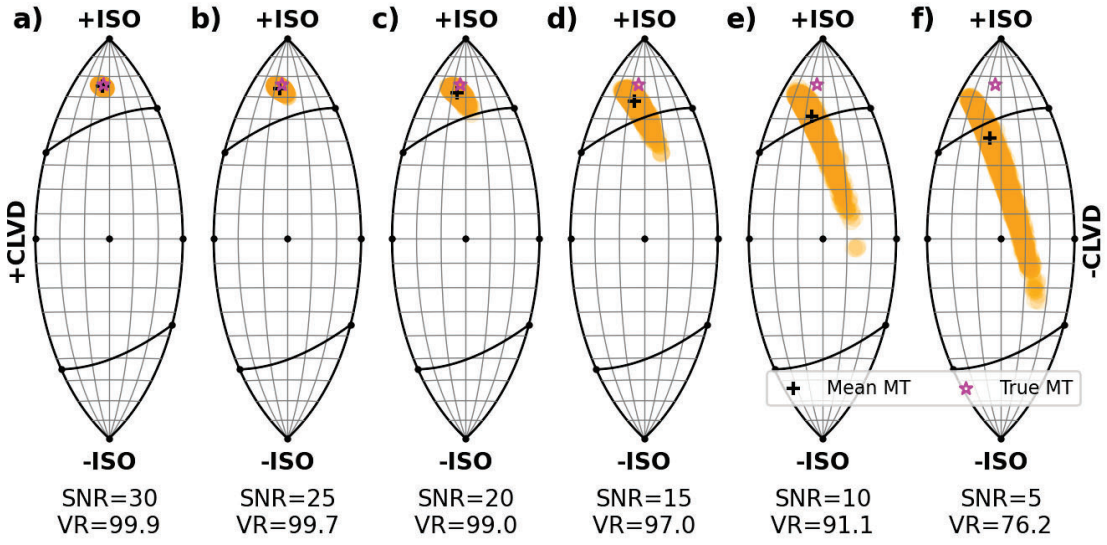

Figure 26. Lune diagram of source type for recovered MT solutions in the scenarios of different signal-to-noise ratio (SNR): (a) SNR=30, (b) SNR=25, (c) SNR=20, (d) SNR=15, (e) SNR=10, and (f) SNR=5. A magenta star denotes the true MT in six scenarios. Overlapped orange dots are MT solutions in the converging stage. Their mean MT is plotted as a black cross. The noise in the simulated waveform is pre-event ambient noise multiplied by different factors to obtain noise waveforms with given SNR.

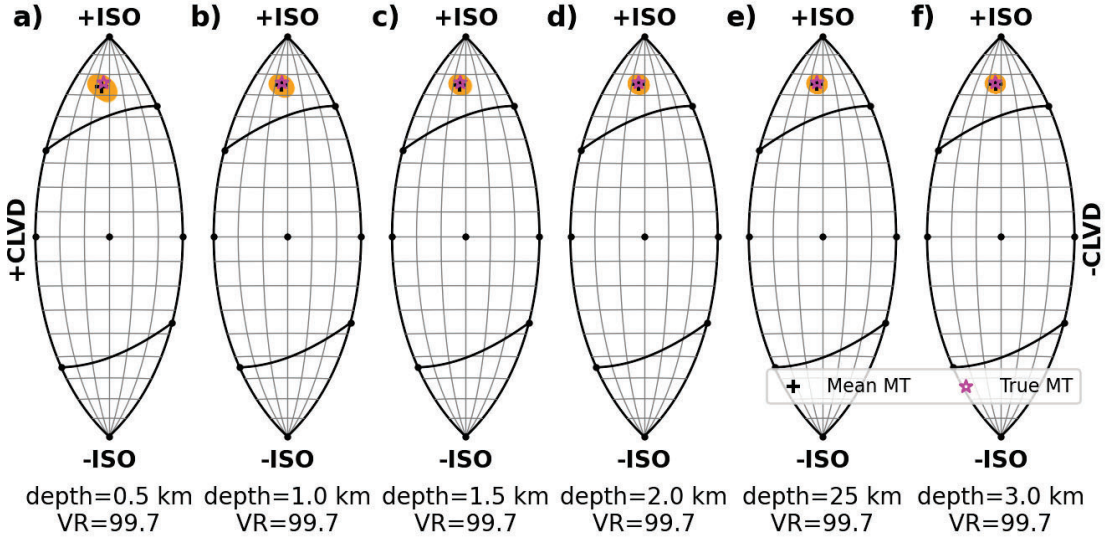

Figure 27. Lune diagram of source type for MT solutions recovered when the true source is at a different depth. The true sources are at depths of 0.5 km, 1.0 km, 1.5 km, 2.0 km, 2.5 km, and 3.0 km, corresponding to panels (a) – (f), respectively. The true depth is used for inversion. The SNR is 25. The orange dots in each panel are MT solutions in the converging stage. The black cross is their mean MT. The magenta star shows the true MT. The two numbers below each lune diagram are the source depth and the variance reduction between 'observed' and predicted waveforms.

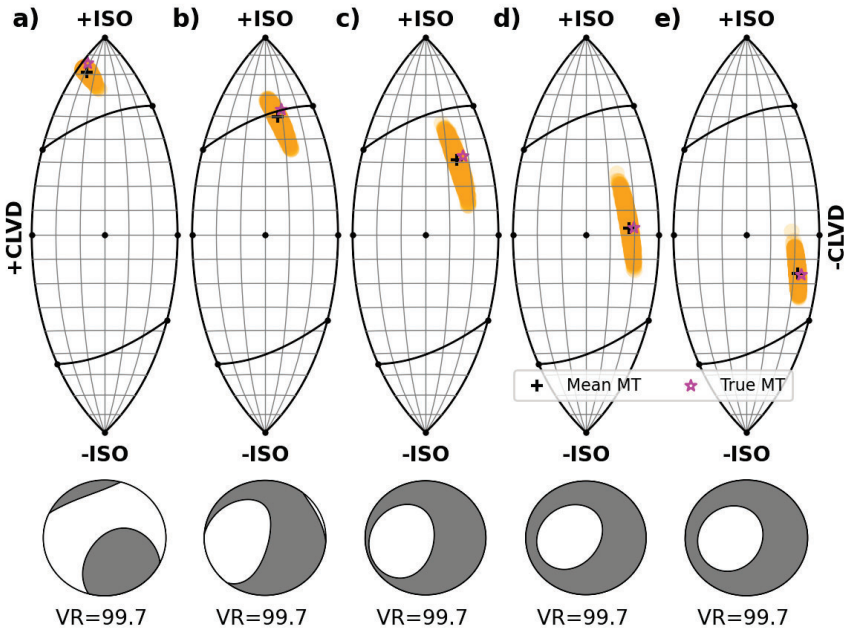

Figure 28. Lune diagram of source type for MT solutions recovered for different input source types. The source depth is 0.5 km. The SNR is 25. The orange dots in each panel are MT solutions in the converging stage. The black cross is their mean MT. The magenta star shows the true MT; the corresponding beachball is plotted below. The number below each beachball is the variance reduction between 'observed' and predicted waveforms.

3.4 Application to DPRK nuclear tests

3.4.1 Data preparation

After the synthetic experiments, we applied the developed inversion framework to five DPRK nuclear tests between 2009 and 2017. To perform MT inversions consistently, we endeavor to use the same set of stations for all events. We choose five standard stations (i.e., MDJ, MAJO, INU, BJT, and HIA, as shown in Figure 18) with sufficient SNR for each nuclear explosion. To fill the azimuthal coverage gap in South Korea, the INCN station is added for the 2009 test, CHNB and YNCB for the 2013 test, and INCN and TJN for the three tests in 2016-2017. Finally, we use six stations for the 2009 and seven for the 2013-2017 tests (Figure 32). The recorded 3-component waveforms are corrected for the instrumental response to obtain displacements and filtered in the 20–50 second period band using a 4-corner acausal Butterworth bandpass filter. The waveforms are then cut into a 150-second window from a manually picked starting time after the origin time of each explosion. The epicenter location and origin time used in this study are from Table 1 of Alvizuri & Tape (2018).

To calculate the synthetic waveforms, we use the same four-layer 1D velocity model MDJ2 (Ford et al., 2009), a revised version of the surface-wave-derived MDJ model (Nguyen, 1994) to calibrate the significant time shifts required for stations BJT and HIA. Green's functions are pre-computed by the wavenumber integration method (Bouchon, 1981) as implemented in Computer Programs in Seismology (CPS; Herrmann, 2013). The buried depths are 618 m, 587 m, 561 m, 586 m, and 307 m for five explosions in 2009-2017, respectively, as estimated by

Myers et al. (2018). Given that the uncertainty of their depth is 200 m, we fix the depth to 500 m for all explosions and examine their influence on our results in the later analysis.

3.4.2 MT inversion results of DPRK 2009-2017 tests

The MT inversion results assuming the same source depth of 0.5km for five DPRK explosions in 2009-2017 are in Figure 29. On the lune diagrams of source-type (Tape & Tape, 2012) for each explosion, starting with randomly chosen initial MTs, our method explores a wide variety of source types, including the ISO-CLVD trade-off area (the darker stripe in each sub-panel) with a higher posterior probability. Finally, the MT solutions converge to a small area with the highest posterior probability (consisting of black dots in each sub-panel in Figure 29). The mean MT solution of each explosion, i.e., the cyan cross in each sub-panel, is calculated by averaging the MTs in these converging areas. The synthetic waveforms computed from the mean MT fit the observed waveforms well, as shown in Figure 30. Some small residual time shifts remain in tangential components due to ignoring the velocity difference between Rayleigh-wave and Love-wave by applying a common time shift for three components at each station. The three diagonal elements of mean MT solutions, MM_{xxxx}, MM_{xxxx},MM_{xxxx} are all positive and larger than the off-diagonal elements, MM_{xxxx}, MM_{xxxx}, MM_{xxxx}. Furthermore, MM_{xxxx} and MM_{xxxx} are almost equal and are smaller than MM_{xxxx}, which indicates these five explosions are close to a crack source. Specifically, five sources are dominated by large ISO components, i.e., 43% in the DPRK2009 test and DPRK2013 test, 50% in the three DPRK2016-2017 tests, respectively, which indicates the explosive nature of the sources. The results also show significant CLVD components in these five explosions (>=30%) and small DC components, e.g., only 13% DC component for the 2017 explosion. Their moment magnitudes are $M_w = 4.69$, $M_w = 4.93$, $M_w = 5.0$, $M_w = 5.13$, and $M_w = 5.79$, respectively.

The uncorrelated noise and time shift at the individual stations are recovered as free parameters. The time shifts at the same stations among events are consistent among five explosions (Table 2). Three stations in China (MDJ, BJT, and HIA) have positive time shifts (up to 4.72 s), the stations in South Korea (INCN, TJN, CHNB, and YNCB) have smaller positive time shifts (0.64 – 2.95 s), while the two stations in Japan require negative time shifts (up to -5.51 s). The time shifts of the same station remain of the same sign even though the actual values vary. The reason is that the possible errors in event origin times also contribute to the time shifts in the observed data. The noise parameter is relative to the standard deviation calculated from 1-hour pre-event ambient noise records. As listed in Figure 30, the noise parameter of MDJ is the smallest for all explosions. At the same time, MAJO and INU stations obtain the largest noise parameters, which agree with the perfect waveform fit at MDJ, and the poorer waveform fits at MAJO and INU stations.

Table 2. Recovered station-specific time shifts (in seconds) for five DPRK 2009-2017 tests. For the 2013 test, the two South Korean stations used were CHNB and YNCB

Explosions	IC.MDJ	IC.BJT	IC.HIA	IU.INCN	KG.TJN	IU.MAJO	G.INU
2009	3.64	4.53	3.03	2.95		-2.7	-4.23
2013	4.06	3.55	3.3	2.41(CHNB)	2.05(YNCB)	-2.54	-4.34
2016a	4.09	4.4	4.35	2.43	1.54	-2.9	-4.8
2016b	4.72	4.4	4.68	2.39	0.96	-2.73	-4.6
2017	3.64	3.69	3.92	1.49	0.64	-3.61	-5.51

The contribution of each station is quantified by the likelihood function instead of only data noise strength. In Equation (22), the data noise CC_{ii} has two competing effects on the logarithm of the likelihood function (Bodin et al., 2012). The resulting likelihood reflects the importance of each station (e.g., Shang & Tkalčić, 2020). For example, in the inversion results of the DPRK 2017, station MDJ plays the most critical role based on the logarithm of the likelihood for all stations (Figure 31). It is the closest station to the sources and has a high SNR. Note that the vertical axis of Figure 31 is on the log scale. The visual comparison of individual station contributions corresponds to much larger relative contributions to the overall likelihood. Therefore, MDJ, INCN, and BJT are the most important stations to the MT inversion of the DPRK 2017 test, while stations MAJO and INU on the Japanese side make only minimal contributions. Thus, similar MT solutions can be obtained if the two latter stations are removed from the analysis.

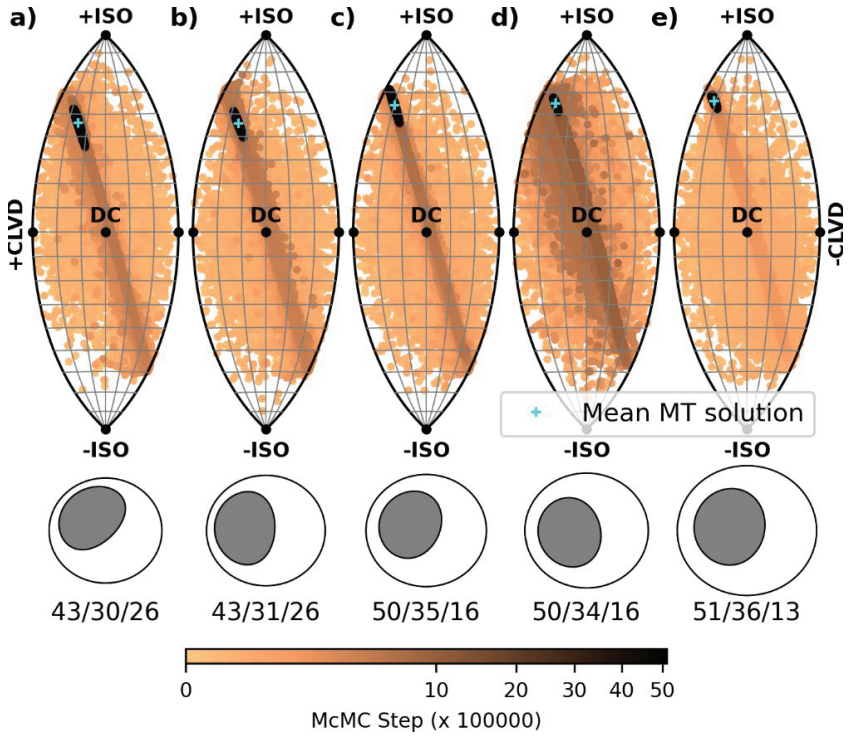

Figure 29. Lune diagram of source type for MT solutions of five DPRK tests in 2009-2017. (a) DPRK2009 test; (b) DPRK2013 test; (c) DPRK2016a (6 Jan 2016) test; (d) DPRK2016b (9 Sep 2016) test; (e) DPRK2017 test. The color bar indicates the equivalent inversion steps with the power law normalization of 2/5 to visualize the unconverging stage better. In each panel, the overlapping color-coded dots show the MT evolution as the inversion step increases. The cyan cross is the mean MT of the converging stage for each explosion. They are also plotted by beachballs below. The size of each beachball is proportional to its moment magnitude. The numbers below each beachball are a percentage of ISO, CLVD, and DC, respectively.

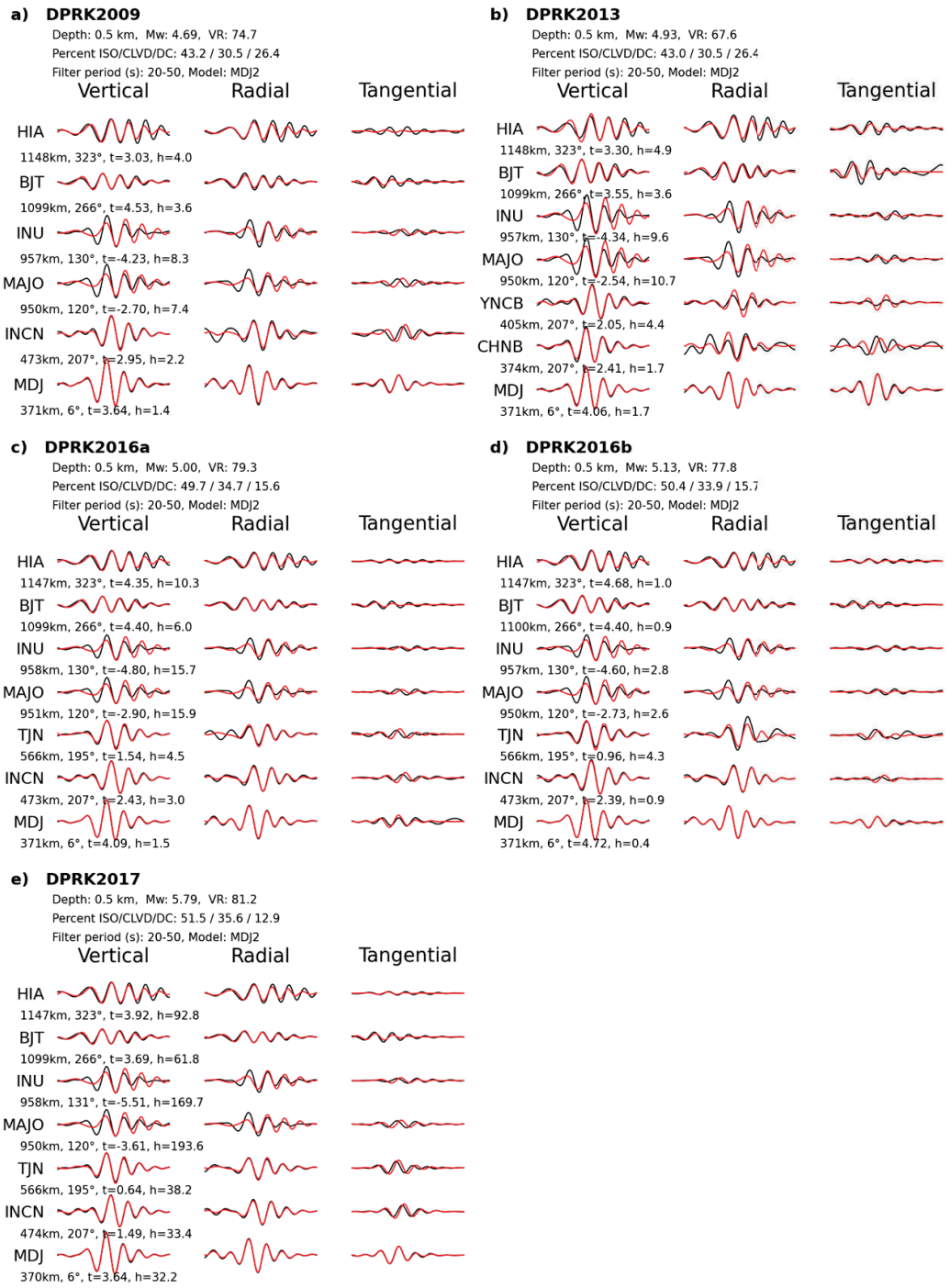

Figure 30. Fits between observed (black) and predicted (red) waveforms for five DPRK 2009-2017 explosions. (a) DPRK2009 test; (b) DPRK2013 test; (c) DPRK2016a (6 Jan 2016) test; (d) DPRK2016b (9 Sep 2016) test; (e) DPRK2017 test. The same 4-corner acausal bandpass (20–50 s) filter was used for each explosion. The five common stations used are MDJ, MAJO, INU, BJT, and HIA. To fill in the azimuth gap in South Korea, INCN was added for the 2009 test, CHNB and YNCB for the 2013 test, and INCN and TJN for the 2016-2017 tests. Numbers beneath each station are the source-station distance, azimuth, the recovered time shift in seconds, and the noise parameter.

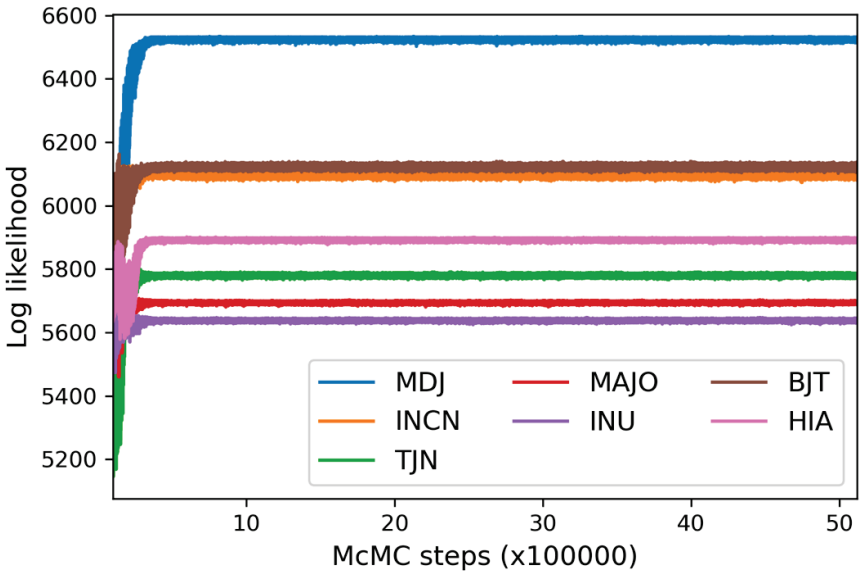

Figure 31. Each station's logarithm of likelihood (Equation (30)) was calculated during MT inversion for the DRPK 2017 test. Most warm-up steps are discarded to illustrate better the likelihood function in the converging stage.

Resolving the depth of shallow sources has been proven difficult with only long-period seismic data Ford et al. (2010b). Alvizuri & Tape (2018a) used a depth of 1 km to perform the full MT inversion. Chiang et al. (2018) thought the DPRK tests should be shallower than 760 m and used 1 km for the 2006 and 0.6 km for five tests from 2009-2017. Some efforts have been made towards constraining the depth of DPRK explosions. Wang et al. (2018) proposed a depth of 450 m by combining space-borne geodetic and seismic records to analyze the DRPK 2017 test. Dreger et al. (2021) assumed a depth of 800 m for joint regional long-period waveforms and regional and teleseismic P-waves polarity data. Cesca et al. (2017) and Liu et al. (2018) performed a grid search over depth to find the best-fit solution. Mustać et al. (2020) thoroughly analyzed the frequency-dependent depth of the DPRK 2013 test using a grid search over depth in several frequency bands. Similarly, we use trial source depths of 0.5–2.5 km, with a 0.5-km increment. Taking the DPRK 2017 test as an example (Figure 32), the ISO component increases by ~7% and the CLVD component decreases by ~9% as the depth increases from 0.5 km to 2.5 km, while the DC component is much more stable, only increasing ~2%. The retrieved moment magnitude M_w trades off with the depth, decreasing from $M_w = 5.79$ to $M_w = 5.48$, also mentioned by Liu et al. (2018). Overall, the MTs over the depth range of 0.5 km – 2.5 km remain stable, i.e., containing large ISO, significant CLVD, and small DC components.

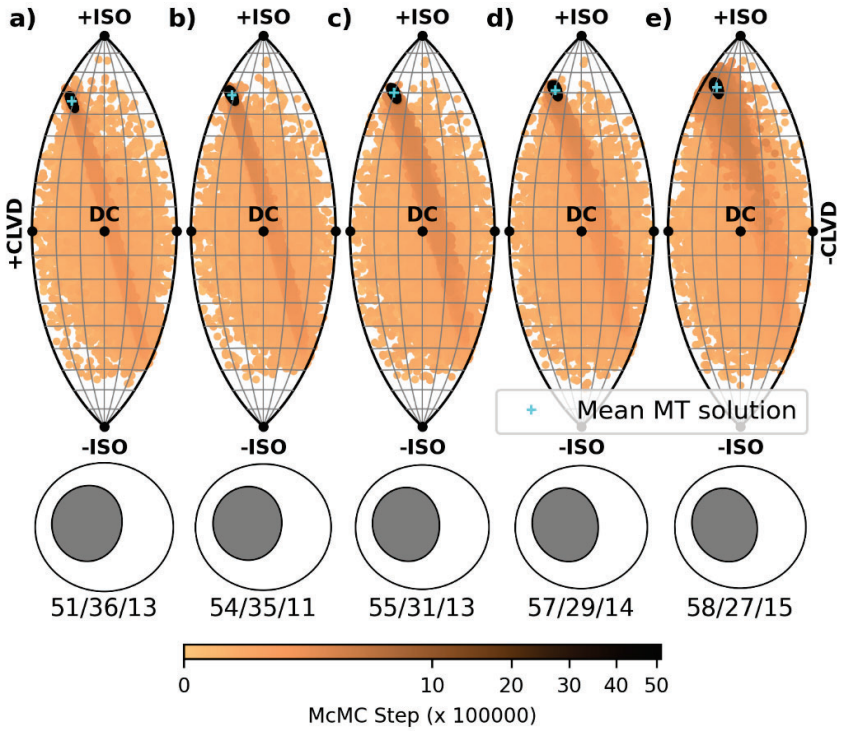

Figure 32. Lune diagrams of source types of MT solutions for the DPRK2017 test assuming the various source depths at 0.5km, 1.0km, 1.5km, 2.0km, and 2.5km in panel (a) - (e), respectively. The moment magnitudes are $\mathbf{M_w = 5.79}$, $\mathbf{M_w 5.67}$, $\mathbf{M_w = 5.58}$, $\mathbf{M_w = 5.53}$ and $\mathbf{M_w = 5.48}$, respectively. See the caption of Figure 29 for other details.

3.4.3 Robustness of the MT inversion

In this study, the contribution of each component driving the recovered MT solution is its log-likelihood weighted by the noise amplitudes (Equation (30)). As noted above, two stations in Japan, i.e., MAJO and INU, play the least important role in comparison to the other five for all events. Therefore, we remove these two Japanese stations and only use the other five stations in Korea and China to invert for the DPRK2017 explosion's MT. The solution is shown in Figure 33, close to a crack source mechanism, with 52% ISO, 37% CLVD, 11% DC, and M_w = 5.8. It is consistent with the source obtained from seven stations in Figure 29e. The recovered station-specific time shifts and noise parameters (Figure 33c) also remain stable compared with those of seven stations in Figure 30e. The variance reduction of waveform fit improves from 81.2% to 92.2% because two poorer fitted stations are neglected in the inversion. Besides, we also use another unique dataset from seven other stations that are closer to the Punggye-ri test site (Figure 34a) to invert for the DPRK2017 event. We apply the same band-pass filter to the waveforms and manually pick 150s-window waveforms. The inversion result using the same velocity model (i.e., MDJ2) shows a similar character to the previous dataset in Figure 29e. The source is dominated by an ISO=54% and is close to the crack source. The CLVD component is up to 38%, and the DC component is negligible (only 7.7%) whose contribution is even smaller than the result in Figure 29e. The pattern of recovered station-specific time shifts (panels a and c of Figure 34) agrees with those in Table 2. Four stations (KSA, CHNB, CHC2, and OKEB), whose surface waves propagate through both thin oceanic and thick continental crust, require a small positive time shift. Three stations (NSN, MDJ, and DACB) need more significant time shifts because the surface waves mainly propagate in the thick continental crust. In addition, these two datasets include a common station, MDJ. The time

shift and noise parameter at this station from two inversions remain stable; specifically, ~3.6 s time shift and a noise parameter of ~32. Therefore, this hierarchical Bayesian MT inversion algorithm is robust under the same assumption of Earth's structure.

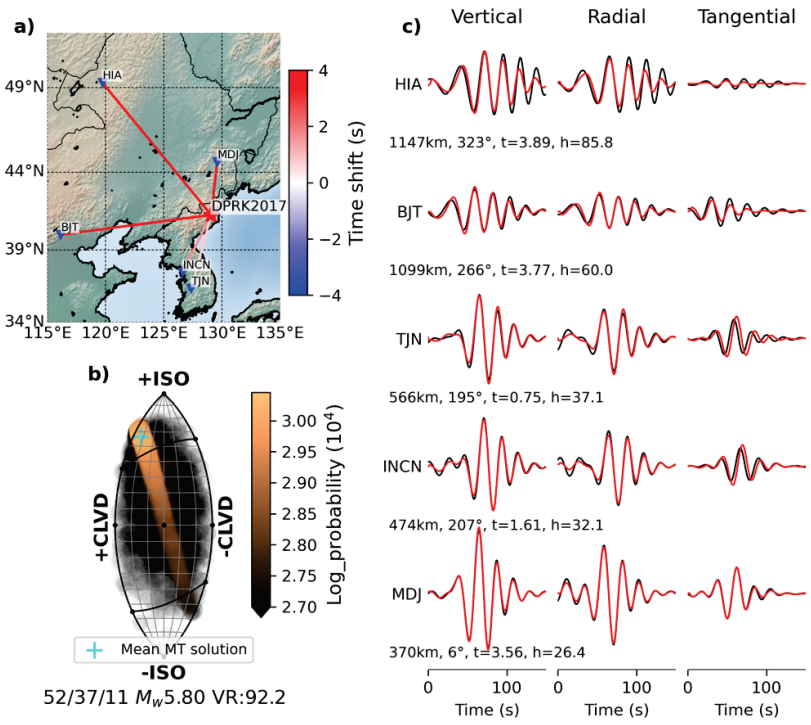

Figure 33. MT Solutions were obtained by hierarchical Bayesian inversion considering uncorrelated noise and 2D structural error when removing two stations in Japan. (a) Map of region showing five stations and the recovered station-based time shift. (b) Lune source type diagram shows MT solutions' evolution during the inversion. The cyan cross marks the source type of the mean MT solution. The color bar is used for log-probability in Equation (30). All log-probability under 2.7×10^4 is set to be black to visualize the later stage of the inversion better. The numbers below are percent ISO, CLVD, DC, moment magnitude of the mean MT, and the variance reduction of waveform fit for observed and predicted waveforms from the mean MT. (c) Waveform fit between observed (black) and predicted waveforms (red) from mean MT plot in (b). The numbers below each row are source-station distance, azimuth, recovered station-based time shift, and noise parameter, respectively.

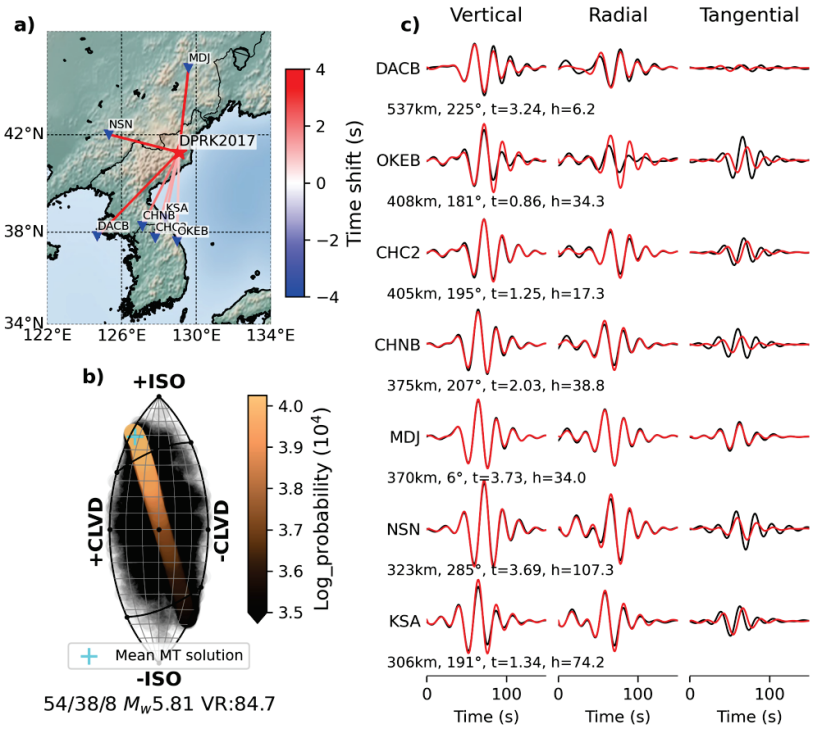

Figure 34. Same as Figure 33, except MT solutions were obtained by hierarchical Bayesian inversion considering uncorrelated noise and 2D structural error using another dataset. (a) Map of region showing seven stations and the recovered station-based time shift. (b) Lune source type diagram shows MT solutions' evolution during the inversion. (c) Waveform fit between observed and predicted waveforms from mean MT plot in (b). See the caption of Figure 33 for more details.

3.5 Discussion

3.5.1 The similarity of MT solutions

The full MTs for five DPRK explosions in 2009 – 2017 feature a significant degree of similarity (Figure 29). This has already been pointed out by Liu et al. (2018). They proposed to use the set of waveforms averaged among five events, given the high degree of similarity to invert for a representative MT solution of the explosions. Similarly, we plot the 3-component waveforms of five explosions at five same stations used in this study in Figure 35. After normalizing by each station's maximum amplitudes, the long-period (20–50 s) waveforms of five explosions show a high degree of similarity. More specifically, the vertical and radial components are almost identical. We calculate the pairwise correlation coefficient of the same components of station MDJ corresponding to five DPRK explosions in Figure 36. The observed data can be treated as the same regarding their shape, which is indicated by the large correlation coefficients (~ 1). Only the tangential component of the 2016a test shows smaller correlation coefficients (~0.7) with other events. Therefore, it is reasonable to expect that these five explosions have similar source mechanisms but different magnitudes. Indeed, the five MT solutions obtained by our Bayesian hierarchical inversions considering uncorrelated noise and 2D structural error support this hypothesis by similar source-type in a lune diagram (Figure 29). Their moment magnitudes are M_w4.69, M_w4.93, M_w5.0, M_w5.13, and M_w5.79, respectively.

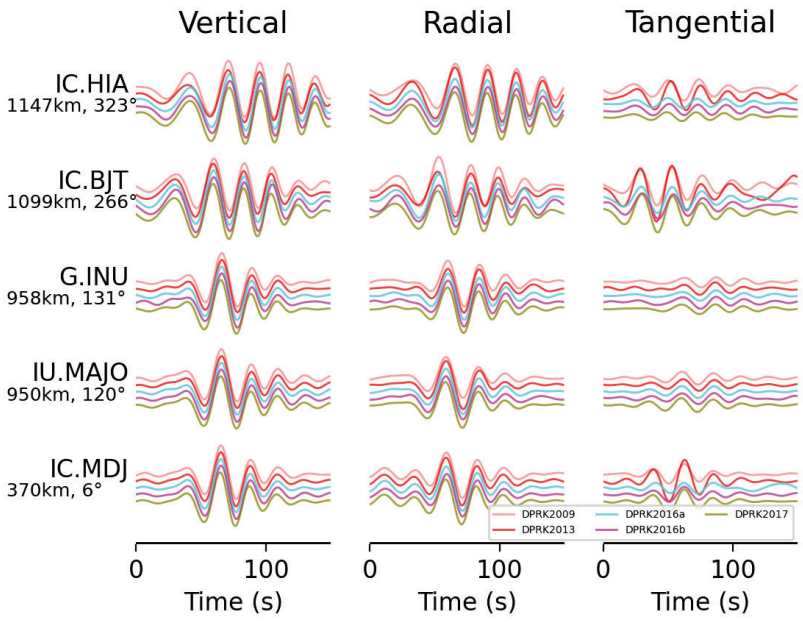

Figure 35. Comparison for 3-component waveforms (displacement) at five standard stations for five DPRK 2009-2017 tests. They are filtered in the 20–50 s period band with a 4-corner acausal Butterworth filter. Waveforms of each event are normalized separately based on the maximum amplitude of three components at each station for better comparison.

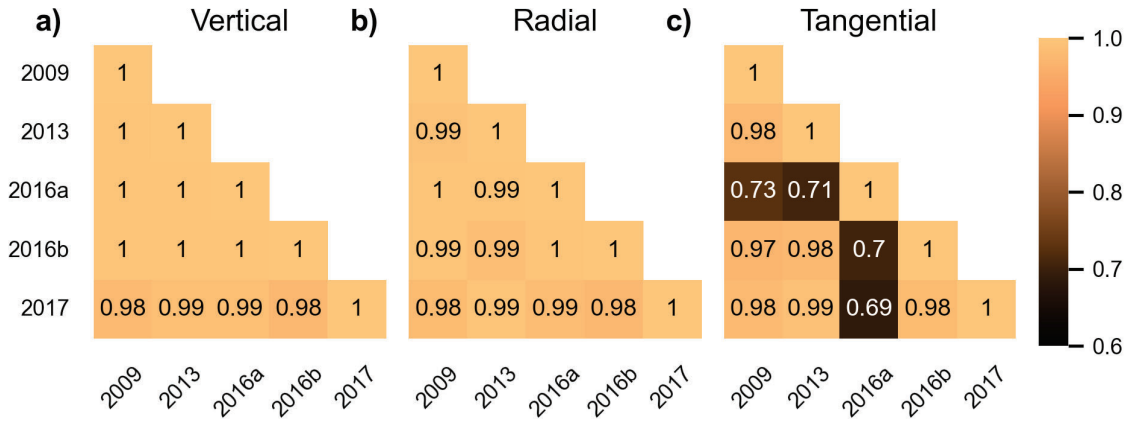

Figure 36. Pairwise correlation coefficients of the same components at MDJ station for five DPRK explosions from 2009 to 2017. (a) Vertical component; (b) Radial component; (c) Tangential component.

3.5.2 Consistency of time shifts and 2D structure.

The recovered time shifts at the same station are consistent among five explosions (Table 2). The time shifts show a consistent pattern: for the corresponding two stations in Japan, negative time shifts indicate that the MDJ2 velocity model is slower than the actual Earth structure along these paths, while for the western and northern three stations in China, and the two stations in South Korea, positive time shifts indicate that the MDJ2 velocity model is faster than the actual Earth's structure. This pattern is because the Korean Peninsula is at the margin of continental crust to the west and north and oceanic crust to the east in the Japanese Sea. Therefore, the path

of surface waves to Japanese stations in Japan (i.e., MAJO and INU) are largely sensitive to the high-speed mantle, which is exposed at shallower depths beneath thin oceanic crust. The paths of surface waves to stations in China (MDJ, BJT, and HHIA) are more sensitive to slower-velocity, thick, continental crust. Additionally, the fact that the two stations in South Korea require smaller positive time shifts compared with the three stations in China could be due to the variation of continental crust thickness along the paths. Thus, overall, the recovered time shifts are consistent with the local geological structures of the stations used.

In this study, the station-specific time shifts between observations and predictions accounts for the possible deviation of Earth structure along specific paths with respect to the assumed 1D Earth model for the entire study region. On the other hand, the error in events' origin times introduces a common time shift among all stations for a given event. This can be a component of the recovered time shifts but can be removed by removing all stations' mean of time shifts, as plotted in Figure 37. Other possible errors such as inaccurate epicenter locations and depths do not cause significant time shifts for individual stations. These errors are negligible for nuclear explosions because the location errors are just on the order of a hundred meters.

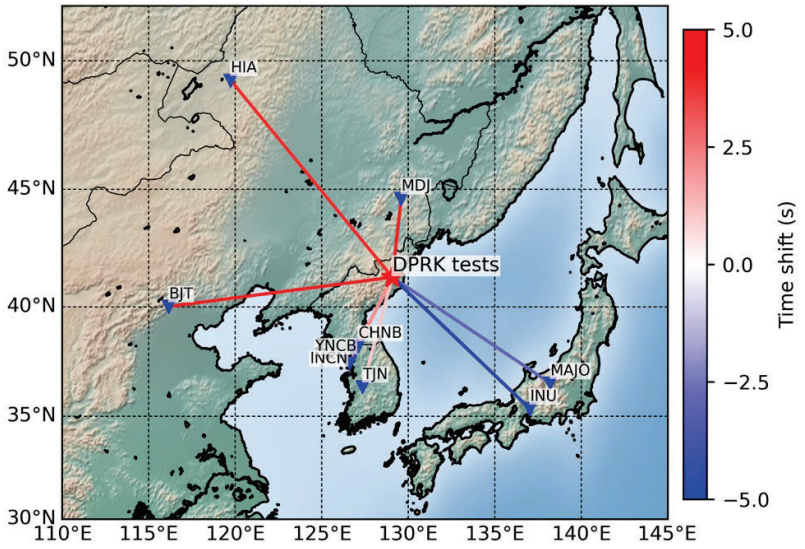

Figure 37. The mean time shift at each station. Positive time shifts (red) mean to shift the predictions forwards, while negative time shifts (blue) for backward.

3.6 The effect of the uncorrelated noise assumption

The data noise covariance matrix plays a vital role in avoiding overfitting waveforms and weighting the station's contribution through the likelihood function in the Bayesian framework. Previous studies have rigorously analyzed the importance of data noise in the source inversion (Mustać et al., 2018, 2020; Mustać & Tkalčić, 2016, 2017). The studies concluded that accounting for the correlated data noise could significantly affect the recovered source mechanism. For instance, accounting for station-specific data noise increased the ISO composition by ~5% for DPRK 2013 test (Mustać et al., 2020). Due to the band limitation, data noise shows the correlated characteristics. Treating the correlated data noise appropriately remains at the forefront of research in non-double-couple MT inversion. Several correlated noise models have been proposed, such as Gaussian or an exponentially decaying noise model (e.g., Bodin et al., 2012; Duputel et al., 2012), empirical noise model from data residuals (e.g., Dettmer et al., 2007) and from synthetic noise series (e.g., Gouveia & Scales, 1998; Piana

Agostinetti & Malinverno, 2010; Sambridge, 1999), and two-attenuated cosine functions from pre-event ambient noise (Mustać et al., 2018; Mustać & Tkalčić, 2016, 2017).

In this study, we assume uncorrelated data noise using a diagonal covariance matrix C_D and focus on another, arguably more critical uncertainty (2D structural error) as analyzed in Section 3.5.2. The uncorrelated data noise is still significant in source inversion. For comparison, we fix the noise level at each station to one instead of inverting for it. This means the noise strength is assumed to be the same as pre-event ambient noise. The MT inversions for five explosions are plotted in Figure 38. Relaxing the noise levels as free parameters increases the ISO components by ~21% for the 2009 test, ~15% for the 2013 test, ~22% for the 2016a test, and ~6% for the 2016b and 2017 tests, which are valuable for explosive source analysis. Besides, the recovered noises at stations for five explosions do not appear to have a specific pattern. This is explainable given that the ambient noise at each station is influenced by some time-dependent variation, e.g., seasonal variation. These five explosions happened at different times with large time gaps.

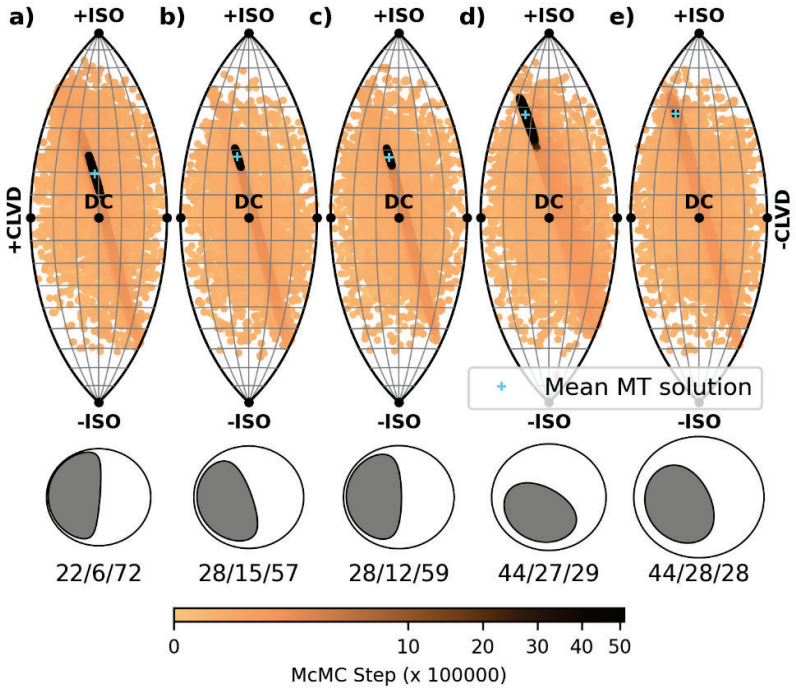

Figure 38. Lune diagram of source type of MT solutions when fixing the noise strength to the same as pre-event ambient noise for five DPRK tests from 2009-2017 with panel (a) - (e), respectively. See the caption of Figure 29 for all details.

3.6.1 Uncertainty of MTs for shallow explosions

Previous MT inversions confirmed their explosive source nature by a significant ISO component (e.g., Alvizuri & Tape, 2018; Chiang et al., 2018; Dreger et al., 2021; Liu et al., 2018; Wang et al., 2018; Xu et al., 2020). However, MT inversion suffers severe uncertainty due to several issues. Firstly, there is an ambiguity between ISO and vertical CLVD mechanisms caused by the very shallow depth of the sources (e.g., Chiang et al., 2014; Ford et al., 2012). The reason is that the long-period waveforms at regional stations are most sensitive to the radiated energy along the equator of a focal sphere, where the pure ISO and vertical CLVD emit similar surface waves at stations at regional distances. Their significant difference in radiation pattern happens only for small take-off angles, meaning teleseismic data are

required to distinguish them, as suggested by Ford et al. (2012) and Chiang et al. (2014). Secondly, the region surrounding the Punggye-ri test site exits a complex structural setting (Mustać et al., 2020), that locates at a margin from continental crust in the west to oceanic crust in the east across the Sea of Japan. Using a 1D velocity model ignoring this strong 2D structure effect may result in uncertainty to MT inversion. In this study, we use the station-based time shift between synthetics and observations to treat this 2D structure with a 1D velocity model. Thirdly, data noise can also introduce uncertainty to MT solutions which are barely considered for DPRK explosions in the previous study. As shown in Figure 39, five previous studies and this study about DPRK 2017 test gave different moment tensors even though all of them fit the observed waveform well with very high VR from 75% to 95%, which testifies to the non-uniqueness in the inversion results. All conclude with a high ISO component. This study's moment magnitude and MT results are similar to those of Liu et al. (2018) using the Crust 2.0 velocity model (Bassin et al., 2000) and an independent dataset in the 0.03-0.09 Hz band.

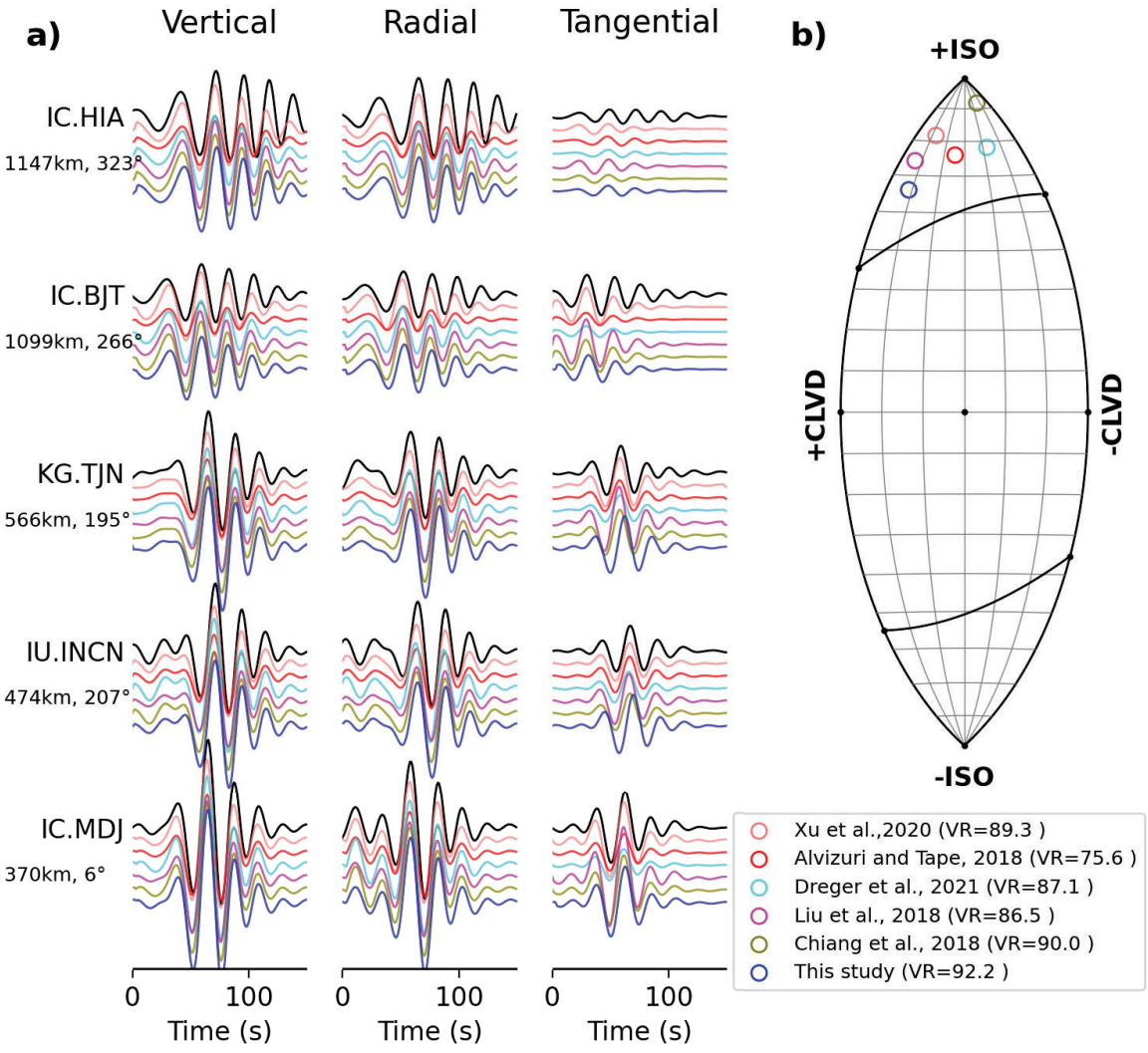

Figure 39. Fits between observed (black) and predicted waveforms (color-coded line) obtained from 5 previous studies and this study for DPRK 2017 explosion. The predicted waveforms in this study are shifted using the recovered time shifts. In contrast, the other five sets of predicted waveforms are shifted with times giving the highest cross-correlation coefficient to the observations. The fit levels (VR) are listed in panel (b) legend.

3.7 Conclusions

This study considers the uncertainty due to uncorrelated data noise involved in the data acquisition process and structural uncertainty primarily due to imperfect knowledge of Earth 2D structure for full MT inversion within the hierarchical Bayesian framework. The data noise on each component is assumed to be uncorrelated, with its standard deviation determined by allowing it to be a free parameter in the inversion. Besides, we use the station-based/path-specific time shifts between observed and predicted waveforms to address the 2D structural uncertainty. Unlike previous studies, the time shifts are relaxed as free parameters, determined simultaneously with noise and moment tensor parameters. We demonstrate the validity of this method via well-designed synthetic experiments. Then, we apply it to perform full MT inversion for five DPRK nuclear tests from 2009 to 2017.

The MT inversion results indicate that the five explosions feature a high degree of similarity. Their sources are dominated by a large ISO component, i.e., 43% for DPRK 2009 and 2013 tests, and 50% for DPRK 2016a, 2016b, and 2017 tests, respectively, which confirms the nature of the explosive source. Additionally, significant CLVD components (30%, 31%, 35%, 34% and 36%) exist in the five explosion sources. But the DC components are much smaller, which are only 26%, 26%, %16, 16% and 13% for the DPRK 2009, 2013, 2016a, 2016b and 2017 tests, respectively. Relaxing the station-based data noise strength also plays an important role in full moment tensor inversion for DPRK explosions by increasing the ISO components. The likelihood function, combining the noise and waveform residuals, weights stations' contributions differently.

Moreover, the recovered station-based time shifts directly indicate the 2D structure character of the surrounding region of these nuclear tests, which demonstrates this method treats the 2D structural error well. Treating structural error, especially 3D structural error, is still leading-edge research in seismic source inversion. This work can be considered a transitional solution for uncertainty consideration between 1D and 3D Earth models.

4 Uncertainties of 3D Earth models in moment tensor inversion of shallow seismic sources

4.1 The resolvability of full moment-tensor for shallow nuclear explosions

4.1.1 Motivation

Under the assumption of no net torque, a seismic MT $\boldsymbol{m} = \{m_{ij}\}; i, j = 1, 2, 3$ is a symmetric matrix with six independent entries. Thus, an MT matrix can be described as a point in $\boldsymbol{R}^6$ space. We previously used the primitive parameterization and Tashiro's (1977) parameterization to represent a seismic MT. Here we decompose a MT as a linear combination of 6 elementary MTs (Kikuchi & Kanamori, 1982), whose focal mechanisms are shown in Figure 40a. This parameterization is particularly helpful in providing insights into the resolvability of the moment tensor as function of depths (e.g., Hejrani & Tkalčić, 2020).

DPRK nuclear explosions are shallow, mostly at 0.5 km depth (Yang et al., 2021). Hejrani & Tkalčić (2020) demonstrates that the inclusion of high-frequency signals (shorter than 6 s) is critical to resolve the full seismic MT for the shallow sources. They performed a numerical analysis to quantify the resolvability of the full MT resolvability as function of depths by analysing numerical characteristics of pre-computed Green's functions generated for AuSREM (Kennett et al., 2013; Kennett & Salmon, 2012; Salmon et al., 2013), a 3D Earth model for the Australian continent.

Here we adopt Hejrani & Tkalčić (2020)'s approach to access the resolvability of full MT solutions for shallow nuclear explosions using the source and station configuration that exists for the 2017 event. Synthetic waveforms used in this analysis are generated using a 3D Earth model of the Korean peninsula (Kim et al., 2016). We find that a shallow source at 0.5 km depth can be recovered using the least-squares inversion method at frequency contents of 6 seconds or shorter in the absence of structural uncertainty, which agrees with Hejrani & Tkalčić (2020).

However, when we adopt the MT inversion in a Bayesian framework, we find that there are other MT solutions that have completely different non-double-couple components than the least-squares MT solution, which can provide comparable fits to synthetic waveforms even at high-frequencies. This highlights the explorative nature of Bayesian approach. These findings reflect the non-uniqueness of the MT solution for shallow sources, which cannot be captured by an optimization method like least-squares. This observation motivates further work concerning the intrinsic resolvability of centroid MTs for shallow nuclear explosions prior to the consideration of structural uncertainty in the problem.

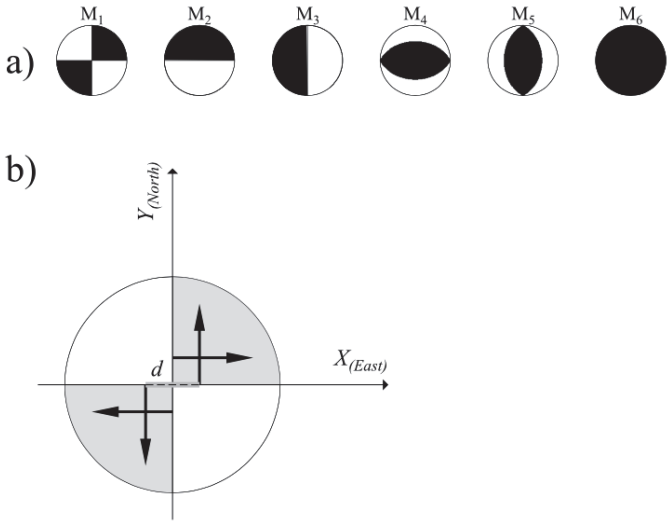

Figure 40. Six elementary basis mechanisms in Equation (34). (b) Plot of forces for the first elementary basis $\boldsymbol{M_1}$. The offset $\boldsymbol{d}$ in Equation (35) is the distance between the centre of the beach ball and each force. Figure taken from Hejrani et al. (2017).

4.1.2 Computation of 3D Green's function using the reciprocity principle

In theoretical seismology, the displacement on the Earth's surface due to a seismic source is expressed as a convolution of impulsive structure responses, known as Green's functions (GF), and a seismic-moment tensor, a mathematical representation of a seismic source under certain assumptions (Aki & Richards, 2002). Specifically, displacement u_i in the direction i, which is one of three orthogonal coordinates, can be expressed as

$$u_i = \sum_{j=1}^{3} \sum_{k=1}^{3} G_{ij,k} * M_{jk}. \tag{34}$$

M_{jk} is an element of a 3x3 moment tensor (MT) matrix. $G_{ij,k}$ is the spatial derivative (in the k-direction) of the GF recorded by the i-component of a seismometer caused by force acting along the j-direction,

$$G_{ij,k}(x_0, x_S) = \lim_{d \to 0} \frac{G_{ij}\left(x_0, x_s + d \cdot \vec{k}\right) - G_{ij}\left(x_0,\ x_s - d \cdot \vec{k}\right)}{2d} \tag{35}$$

$G_{ij}(x_0, x_s, t)$ is the Green's function or structure response of a single force placed at x_s recorded at x_0.

Pre-computed GFs for 1D structural models can be calculated using analytical methods (e.g., Herrmann, 2013; Randall, 1989) on a personal workstation. When it comes to 3D structural models, GFs must be computed using numerical methods, such as the spectral element method (SEM) (Fichtner et al., 2009; Komatitsch et al., 2002; Komatitsch & Tromp, 1999, 2002a, 2002b). Although the performance of SEM has significantly increased in the recent past and is widely used in many full-waveform simulation applications, it remains a demanding HPC task. In this project, the spectra-element code SES3D (Fichtner et al., 2009; Gokhberg & Fichtner, 2016) is used to calculate a library of synthetic GFs corresponding to elementary MTs. Previous studies have demonstrated the reliability of SES3D to generate GFs in studies of earthquakes in Papua New Guinea and the Solomon Islands (Hejrani et al., 2017; Hingee et al., 2011), and the 2016 Peterman Range earthquake (Hejrani & Tkalčić, 2018) using AusREM,

the 3D reference velocity model for Australia and surrounding regions (Kennett et al., 2013; Kennett & Salmon, 2012; Salmon et al., 2013).

In order to access the resolvability of the MT solution as a function of depth and lateral coordinates, we need to calculate structural Green's functions at a large number of mesh points surrounding the preliminarily estimated source location. This is prohibitively computationally expensive when repeating expensive 3D waveform simulations. Thus, we follow Hejrani et al. (2017) to use the reciprocal principle to generate pre-computed Green's functions for sources on a dense mesh grid to be recorded by a number of selected stations (Figure 41). In particular, in order to generate seismic wavefields from impulsive sources located on a grid (red points in the left panel of Figure 41) to be recorded by regional stations (blue triangles), we indeed place sources at the station location (red dots on the right panel of Figure 41) and record the wavefield at grid points.

Figure 42, Figure 43, and Figure 44 benchmark the results of the reciprocal simulation method by comparison with waveforms generated in a forward setting. These figures also highlight the fact that at shallow depth (0.5 km), ground motions are insensitive to dip-slip mechanisms (M_2 and M_3) at all filter bands under consideration. Therefore, great care must be taken to interpret the MT and its corresponding non-double couple components recovered at this depth.

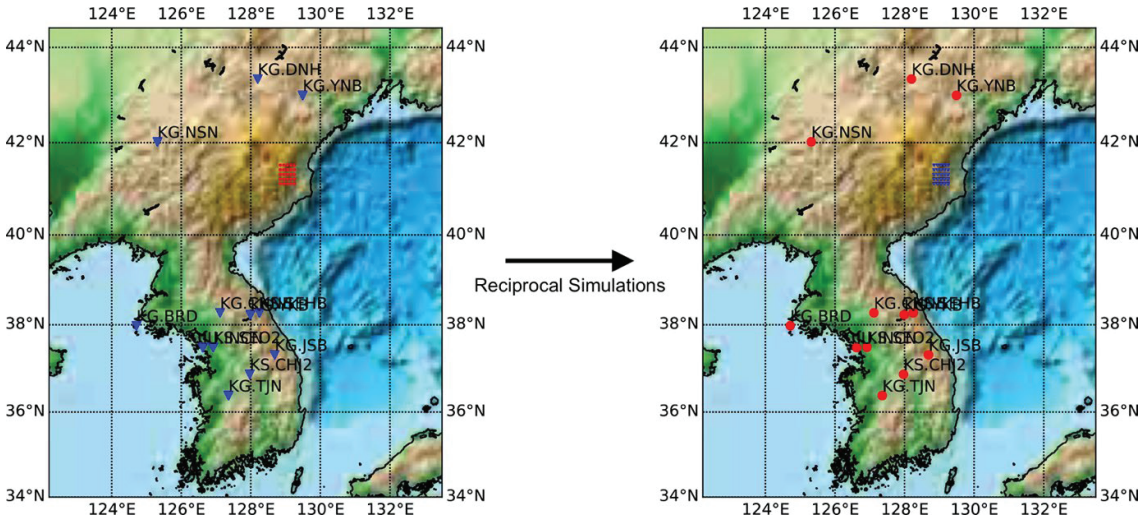

Figure 41. Principle of reciprocity. A) Real source-station settings of DPRK explosions in the Korean Peninsula. B) Reciprocal source-station settings used in the numerical simulations.

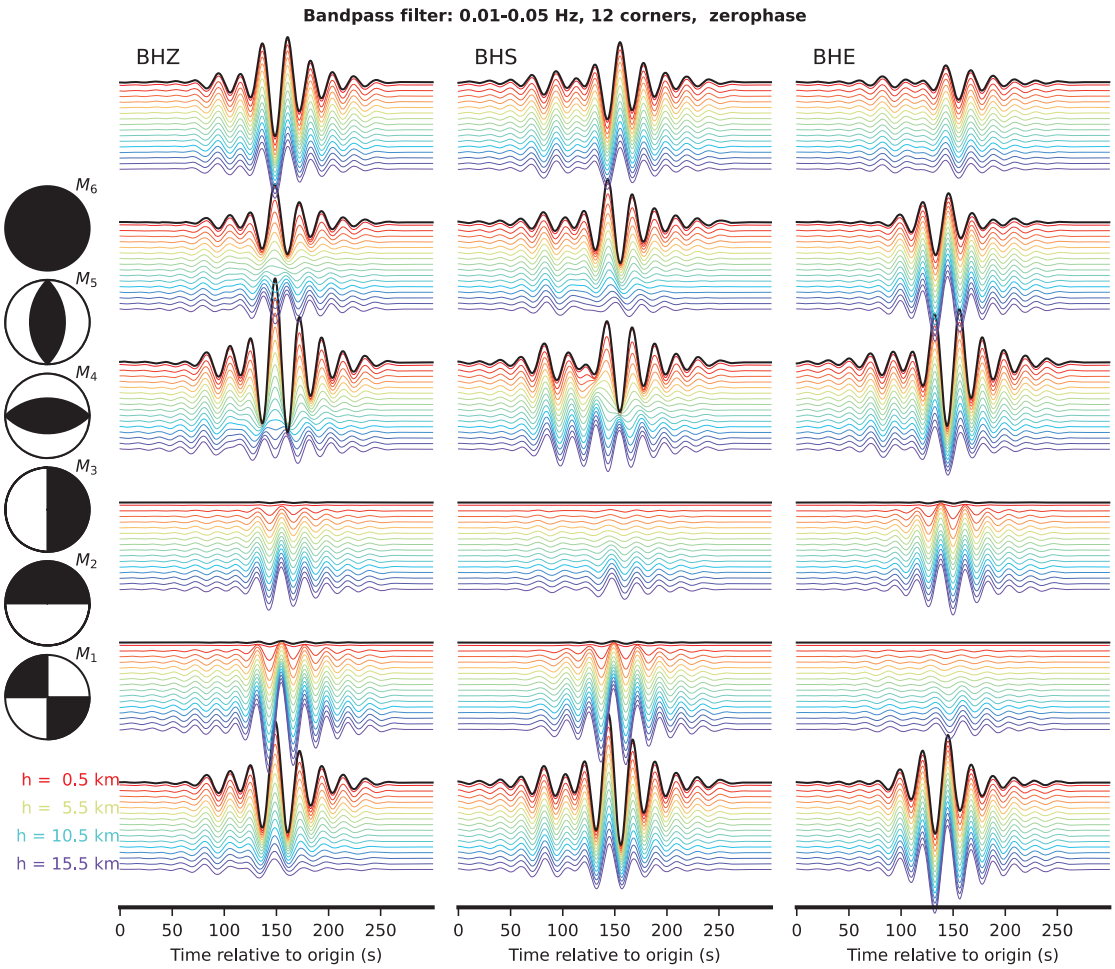

Figure 42. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.05 Hz. Black waveforms corresponding to six elementary MTs (left panel) are simulated by the actual source-location setting (Figure 41, left). The source depth is fixed at 0.5 km. Colour waveforms are generated using the reciprocal principle for 16 depth values ranging from (0.5 to 15 km).

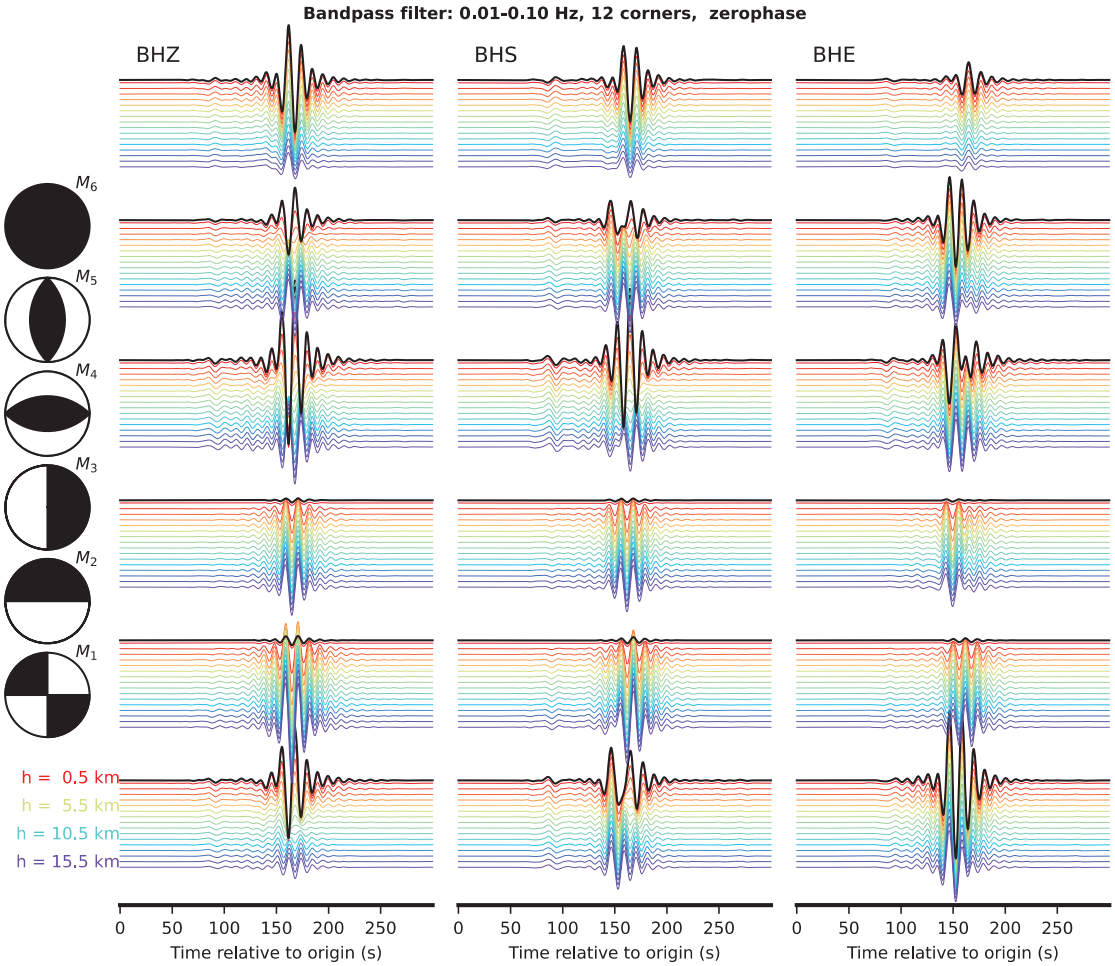

Figure 43. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.10 Hz. Black waveforms corresponding to six elementary MTs (left panel) are simulated by the actual source-location setting (Figure 41, left). The source depth is fixed at 0.5 km. Colour waveforms are generated using the reciprocal principle for 16 depth values ranging from (0.5 to 15 km).

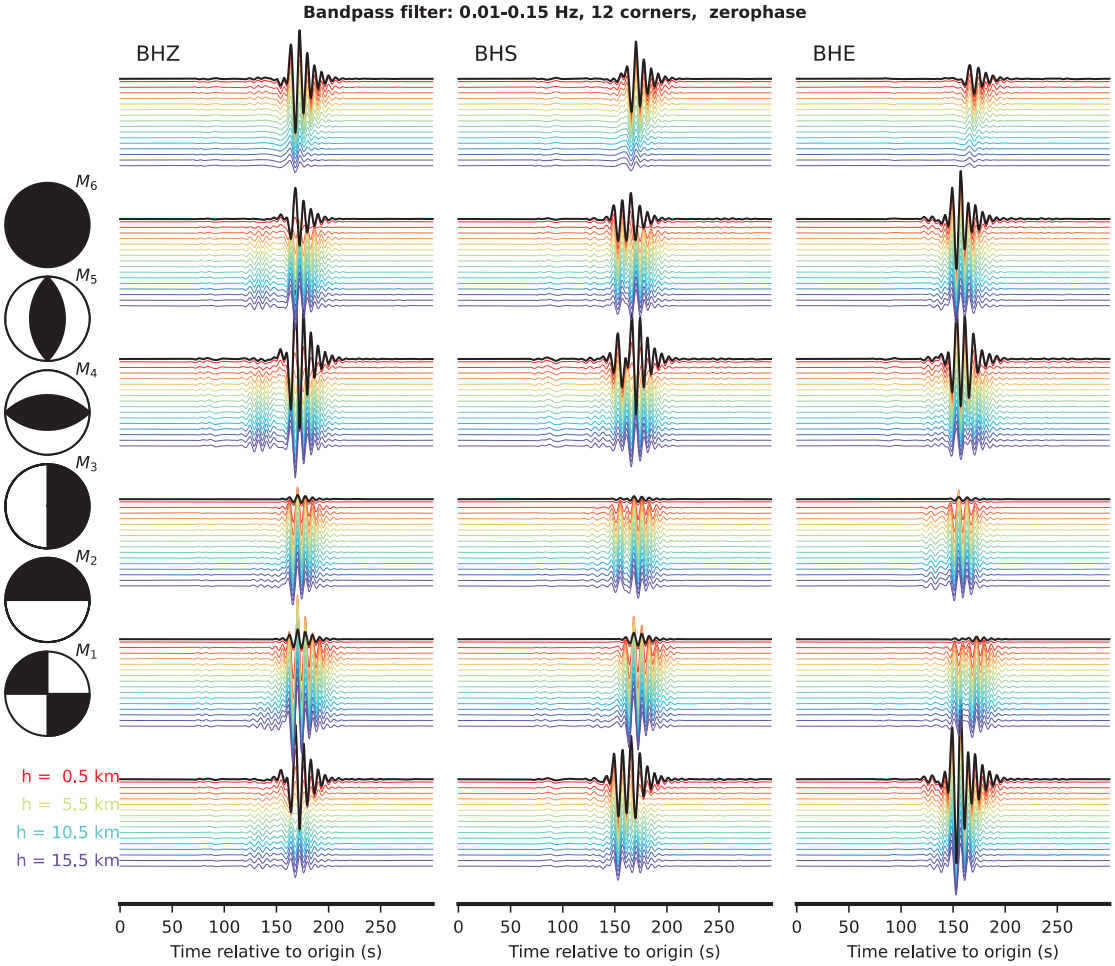

Figure 44. Benchmark of the reciprocal simulations; waveforms are filtered between 0.01–0.15 Hz. Black waveforms corresponding to six elementary MTs (left panel) are simulated by the actual source-location setting (Figure 41, left). The source depth is fixed at 0.5 km. Colour waveforms are generated using the reciprocal principle for 16 depth values ranging from (0.5 to 15 km).

4.1.3 Resolvability of shallow seismic moment tensors

Here we follow analyses in Hejrani and Tkalčić (2020) to access the resolvability of MT solutions using the least-squares estimates.

At a specific spatial location, the forward waveform prediction can be written in linear form,

$$\boldsymbol{u} = \boldsymbol{G}\boldsymbol{m} \tag{36}$$

Equation (36) is a matrix form of Equation (34), where $\boldsymbol{u}$ represents the observed displacement vector, $\boldsymbol{G}$ is a transfer matrix constructed by pre-computed Green's functions corresponding to the seismic source at the spatial location, and $\boldsymbol{m}$ represents the seismic MT. The least-squares solution to the forward problem in Equation (36) is

$$\boldsymbol{m^{est}} = (\boldsymbol{G^t G})^{-1} \cdot \boldsymbol{G^t u} \tag{37}$$

The condition number, defined as

$$cn = \frac{\max eig(G^t G)}{\min eig\ (G^t G)} \tag{38}$$

represents the stability of the matrix inversion operator in Equation (37). A large condition number could lead to an unstable inverted matrix, meaning that a small perturbation of the observation $\boldsymbol{u}$ could lead to a large perturbation in the estimated $\boldsymbol{m}$.

The balance of amplitudes (Hejrani and Tkalčić, 2020) is a measure of relative waveform amplitudes of M_2 and M_3 in comparison with other elementary MTs. A high balance of amplitude means that the elementary seismograms corresponding to the dip-slip moment tensors are comparable to other components. This in turn enhances the possibility for the components can be correctly recovered in the inversion.

Our analyses customised for Korean peninsula settings (Figure 41) show similar results to the previous finding by Hejrani and Tkalčić (2020), where they conclude that high-frequency content (of 6 s or shorter) is required to resolve MTs for seismic sources as shallow as 1 km. Therefore, we will only present results for this frequency band in the rest of the report.

In Figure 46 and Figure 47, we adopt the least-squares method to recover input MTs from a synthetic experiment (Hejrani et al., 2017). A seismic source is placed at 0.5 km (Figure 46) and 3.5 km (Figure 47). The source depth and hypocentral coordinates are not known in the inversion. So, at each mesh point (red crosses in top panels of Figure 46 and Figure 47), we calculate the least-squares estimate of the MT using Equation (37). The goodness of fit between the observation and the prediction corresponding to the least-squares solution is defined as the variance reduction

$$VR = 100\left(1 - \frac{\sum_{i=1,N}(p_i - o_i)^2}{\sum_{i=1,N} o_i^2}\right). \tag{39}$$

The inverted solution is the MT at the mesh points that yields the greatest variance reduction. In these experiments MT solutions for sources at depths 0.5 and 3.5 kms can be correctly recovered in this method when the filter band 0.01–0.15 is considered.

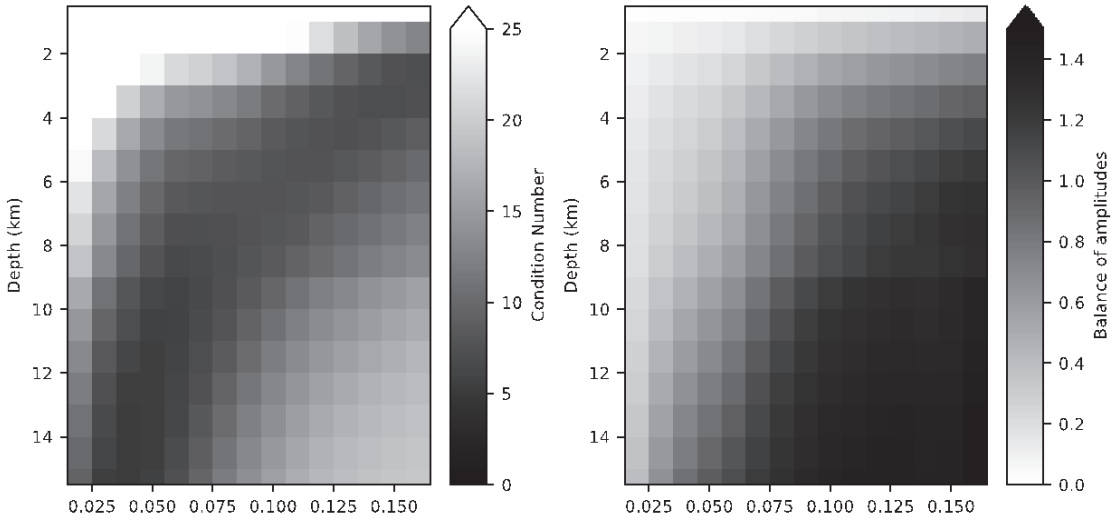

Figure 45. (left) Condition number of G^tG as function of depth and frequency (Equation (38)). (right) Balance of amplitudes among elementary Green's Functions.

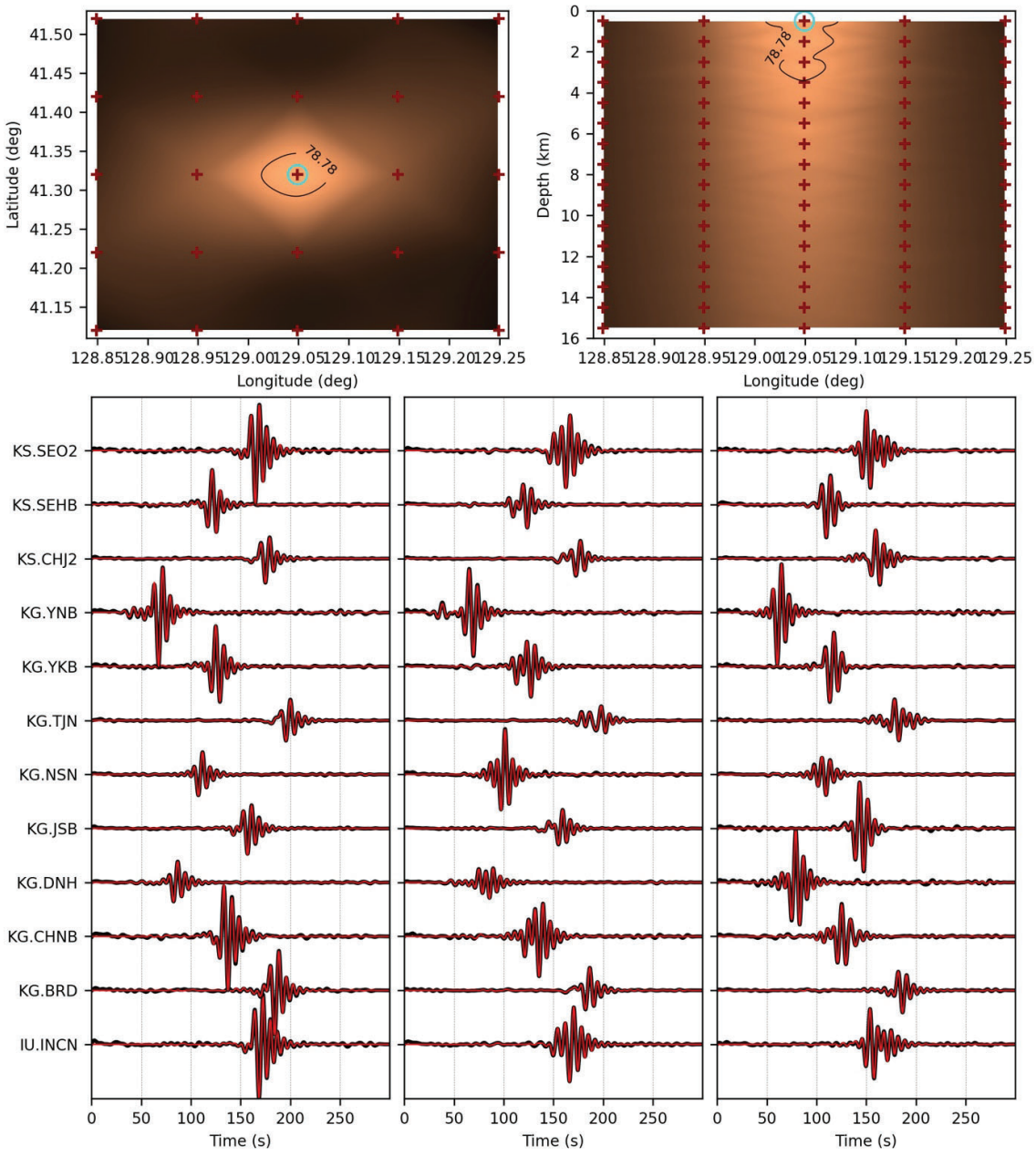

Figure 46. Least-squares solutions for the seismic source at 0.5 km. (top) Horizontal and vertical slices of the mesh grid surrounding a seismic source. Green's functions are calculated for grid points marked by red crosses. The true solution location is circled in cyan. Backgrounds show the interpolated VR (Equation (39)) among grid points. The contours show the regions having VR larger than 90% of the maximum VR in the entire mesh. (bottom) Synthetic data with added noise are in black. Recovered waveforms corresponding to the MT solution are in red.

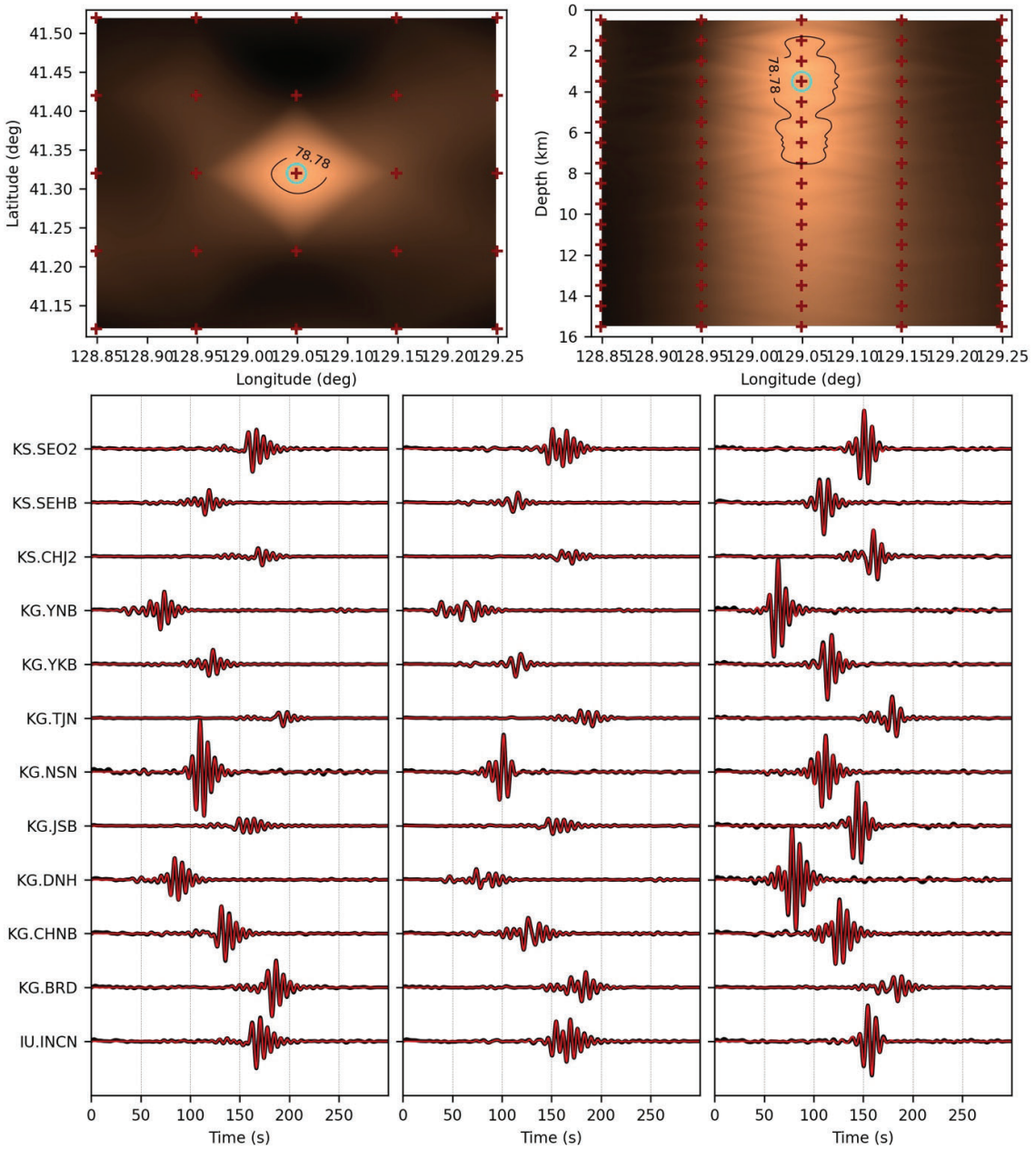

Figure 47. Least-squares solutions for the seismic source at 3.5 km. Other details are similar to Figure 46.

4.2 An approach to perturb 3D Earth structure

We propose an empirical Monte-Carlo method to perturb an original 3D earth model to obtain an ensemble of slightly perturbed 3D models (Figure 48). Green's functions corresponding to the perturbed Earth models are computed using the SES3D method (Gokhberg & Fichtner, 2016) (more details in Section 4.1.2) and shown in Figure 49. We are currently at the experimental stage of incorporating 3D structural uncertainty characterised by an ensemble of GFs function in the MT inversion scheme, similar to how we treat 1D earth model uncertainty.

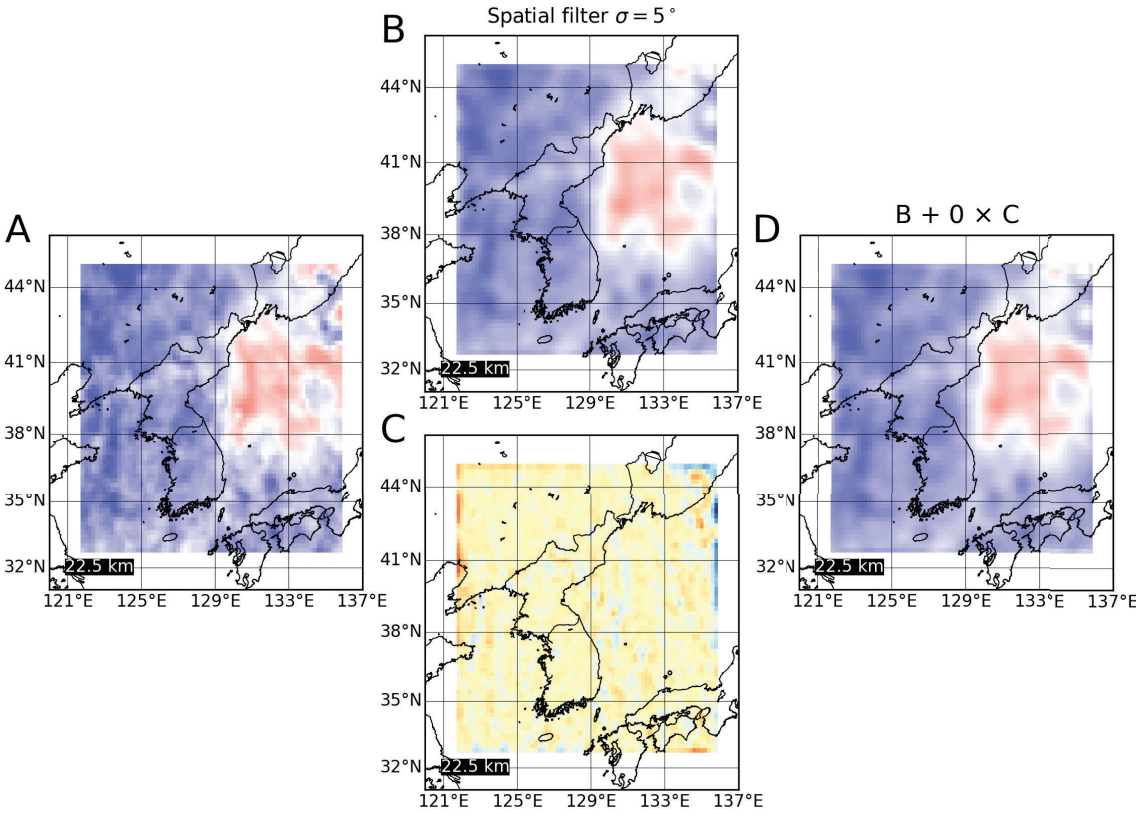

Figure 48. A method to perturb 3D Earth models. (A) A depth slice at 20 km from the 3D Earth model surrounding the Korean peninsula from ambient noise tomography and trans-dimensional Bayesian inference (Kim et al., 2016). This slice shows the presence of small-scale features (C) embedded over large-scale geological features (D). We use a Gaussian spatial filter to obtain (B) from (A); (C) is the difference of (A) and (B). The amplitudes of small-scale features are treated as unknowns in the perturbed 3D earth model as shown (D).

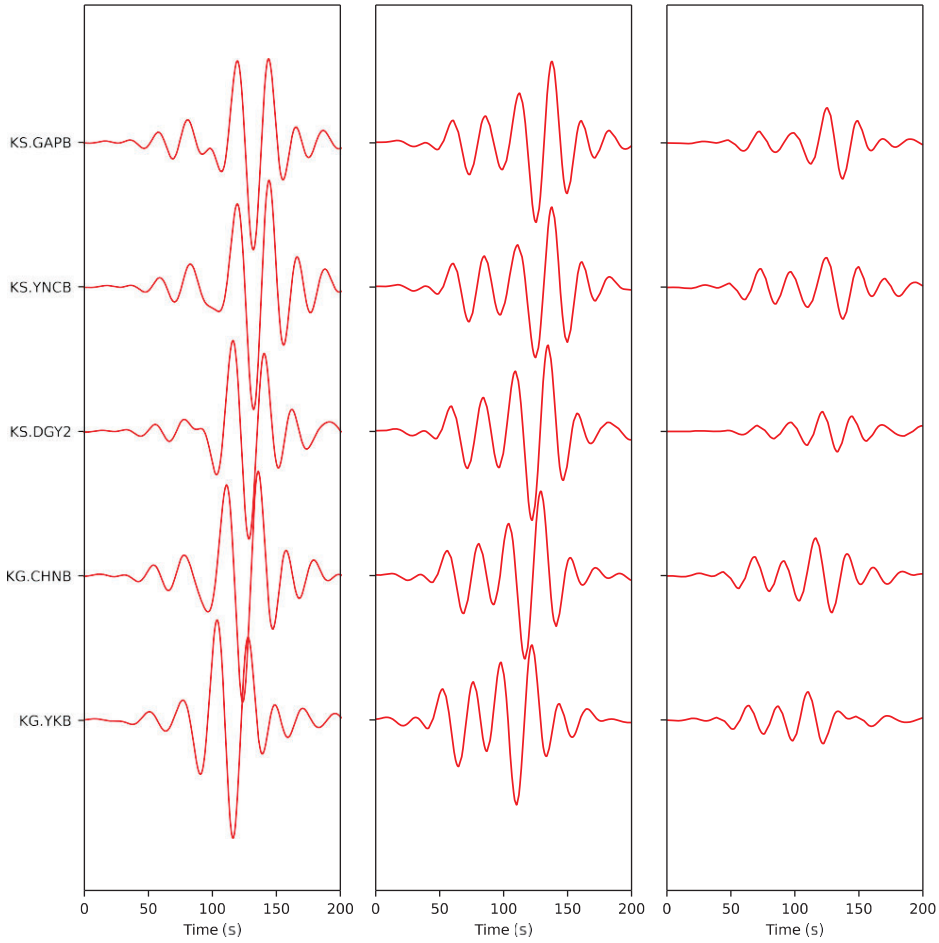

Figure 49. Green's functions corresponding to Monte-Carlo perturbed 3D Earth models using the method presented in Figure 48. Grey lines are individual Green's functions for perturbed Earth models while red lines are the mean of the individual Green's functions.

A Supporting information for Section 2

This appendix chapter includes sections A.1 to A.3, Table 3 and Figure 50 to Figure 69, which expand the brief discussion in the main text.

A.1 Derivation of combined covariance matrix

In this supplementary section, we establish the likelihood function of the observation, $\boldsymbol{d}_{obs}$, with prediction, $\boldsymbol{d}_{pred} = g(\widehat{\Omega}, \boldsymbol{m})$, corresponding to the reference Earth model, $\widehat{\Omega}$, of a MT, $\boldsymbol{m}$, in consideration. The likelihood function incorporates the structural uncertainty and the presence of data noise. This supports the abbreviated derivation presented in Section 2.2 of the main text.

The data noise is assumed due to the background noise at the recording sites, which is independent of actual ground motion induced by seismic waves, $\boldsymbol{d}$, which is the hypothetical ground motion recorded in the absence of observational noise. Therefore, the observed data, $\boldsymbol{d}_{obs}$ can be thought of as the superposition of data noise, $\boldsymbol{\epsilon}_d$, and the actual ground motion, $\boldsymbol{d}$

$$\boldsymbol{d}_{obs} = \boldsymbol{d} + \boldsymbol{\epsilon}_d. \tag{40}$$

The stochastic characteristics of the data-noise can be described by the data covariance matrix, $\boldsymbol{C}_d$, which governs the stochastic behavior of each sampling point in time and possible correlation between them. Let's assume that the data noise is Gaussian, the likelihood of observing $\boldsymbol{d}_{obs}$ given the actual ground motion $\boldsymbol{d}$ is expressed by

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{d}) = \frac{1}{\sqrt{(2\pi)^N \, |\boldsymbol{C}_d|}} \exp\left(-\frac{1}{2}[\boldsymbol{d}_{obs} - \boldsymbol{d}]^T \, \boldsymbol{C}_d^{-1} \, [\boldsymbol{d}_{obs} - \boldsymbol{d}]\right). \tag{41}$$

Operators $|\boldsymbol{C}_d|$ and $\boldsymbol{C}_d^{-1}$ are the determinant and inversion of the data noise covariance matrix and the superscript T denotes the transposed matrix/vector.

The theory uncertainty is the possible difference in the waveforms predicted by the reference model, $\widehat{\Omega}$, and the actual ground motion,

$$\boldsymbol{\epsilon}_t(\boldsymbol{m}) = \boldsymbol{d} - g(\widehat{\Omega}, \boldsymbol{m}). \tag{42}$$

As noted above, in this study, we assume the theory uncertainty stems only from the uncertain knowledge of the Earth structure, and the actual ground motion can be predicted accurately if the "true Earth structure" is known precisely, $\boldsymbol{d} = g(\widetilde{\Omega}, \boldsymbol{m})$. Thus, Equation (42) can be re-written as,

$$\boldsymbol{\epsilon}_t(\boldsymbol{m}) = g(\widetilde{\Omega}, \boldsymbol{m}) - g(\widehat{\Omega}, \boldsymbol{m}). \tag{43}$$

From Equation (43), we can see that the uncertain knowledge of Earth structure is propagated to the waveforms predicted theoretically. Therefore, in the scope of this study, we refer to the theory uncertainty interchangeably as the structural uncertainty. It is also important to note that the theory uncertainty is a function of the source moment tensor (Duputel et al., 2014). It is intuitive to see the relation when the prediction theory is linear (Equation (3) in the main text),

$$\boldsymbol{\epsilon}_t(\boldsymbol{m}) = \left(\boldsymbol{G}^{\widetilde{\Omega}} - \boldsymbol{G}^{\widehat{\Omega}}\right) \cdot \overline{\boldsymbol{m}} \tag{44}$$

Unlike the data noise, the theory uncertainty is proportional to the event magnitude, and it depends on the MT solution involved.

Like the data noise, we assume the theory uncertainty, $\boldsymbol{\epsilon}_t(\boldsymbol{m})$, is Gaussian. Furthermore, given that the "true Earth" is within the uncertainty representation of the reference model, which is $\widetilde{\Omega} \in \{\Omega_i\}$, the theory noise covariance matrix can be estimated from the collection of structural models, $\{\Omega_i\}$,

$$\boldsymbol{C}_t(m) = \frac{1}{N_m - 1}\sum_{i=1}^{N_m}([g(\Omega_i, \boldsymbol{m}) - \bar{g}(\boldsymbol{m})]^T \cdot [g(\Omega_i, \boldsymbol{m}) - \bar{g}(\boldsymbol{m})]), \tag{45}$$

where $\bar{g}(\boldsymbol{m}) = \frac{1}{N_m}\sum_i^{N_m} g(\Omega_i, \boldsymbol{m})$ is the mean Green's function. Ideally, the theory covariance matrix corresponding to the unknown MT, $\boldsymbol{m}$, is correctly estimated with the true MT, $\widetilde{\boldsymbol{m}}$. However, because the true solution is unknown, we can instead estimate the structural uncertainty covariance matrix with a prior estimate of moment tensor, $\boldsymbol{m}^p$, which is expected to be close to $\widetilde{\boldsymbol{m}}$. Given the definition of the theory noise covariance matrix, the likelihood function of the actual ground motion $\boldsymbol{d}$, given the MT, $\boldsymbol{m}$ is expressed as,

$$p(\boldsymbol{d} \mid \boldsymbol{m}) = \frac{1}{\sqrt{(2\pi)^N\,|\boldsymbol{C}_t(\boldsymbol{m})|}}\exp\left(-\frac{1}{2}\left[\boldsymbol{d} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T \boldsymbol{C}_t^{-1}(\boldsymbol{m}^p)\,[\boldsymbol{d} - g(\widehat{\Omega}, \boldsymbol{m})]\right) \tag{46}$$

Ideally, the uncertainty is correctly represented with the true MT solution, that is, $\boldsymbol{m}^p = \widetilde{\boldsymbol{m}}$, but $\widetilde{\boldsymbol{m}}$ is being sought for.

The likelihood, $p(\boldsymbol{d}_{obs} \mid \boldsymbol{m})$, is the probability that the observation is true given that $\boldsymbol{m}$ is the seismic source responsible for the recorded seismic waves. It can be obtained by applying the law of total probability for Equations (41) and (46),

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \sum_{all\ possible\ \boldsymbol{d}} p(\boldsymbol{d}_{obs} \mid \boldsymbol{d}) \cdot p(\boldsymbol{d} \mid \boldsymbol{m}) \tag{47}$$

Equation (47) can be algebraically evaluated as,

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \eta(\boldsymbol{m}^p)\exp\left(-\frac{1}{2}\left[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T [\boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^p)]^{-1}\,[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)]\right) \tag{48}$$

and the normalization factor is

$$\eta(\boldsymbol{m}^p) = (2\pi)^{-N}\,|\boldsymbol{C}_d|^{-1/2}\,|\boldsymbol{C}_t(\boldsymbol{m}^p)|^{-1/2}\,|\boldsymbol{C}_d^{-1} + \boldsymbol{C}_t(\boldsymbol{m}^p)^{-1}|^{-1/2} \tag{49}$$

We refer to Tarantola and Valette (1982), Tarantola (2005) and Duputel et al. (2014) for a complete algebraic derivation and proof. The normalization factor, $\eta(\boldsymbol{m}^p)$, can be further reduced due to the fact that the determinant operator is multiplicative and the following matrix equality, $\boldsymbol{C}_d \cdot \boldsymbol{C}_t(\boldsymbol{m}^p) \cdot (\boldsymbol{C}_d^{-1} + \boldsymbol{C}_t(m^p)^{-1}) = \boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^p)$, holds. Thus, the integrated likelihood in Equation (47) has a simple form (equation 6.248, Tarantola, 2005),

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \frac{1}{(2\pi)^N\sqrt{|\boldsymbol{C}_D|}}\exp\left(-\frac{1}{2}\left[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T \boldsymbol{C}_D^{-1}\,[\boldsymbol{d}_{obs} - g\left(\widehat{\Omega}, \boldsymbol{m}\right)]\right) \tag{50}$$

and the combined covariance matrix is

$$\boldsymbol{C}_D(\boldsymbol{m}^p) = \boldsymbol{C}_d + \boldsymbol{C}_t(\boldsymbol{m}^p) \tag{51}$$

Equations (50) and (51) here are equivalent to Equations (5) and (6) in the main text.

A.2 The least-squares solution

When the forward problem is linear (Equation (3) in the main text) and the covariance matrix (Equation (6) in the main text) is determined prior to the inversion, an analytical solution minimizes the quadratic misfit function appearing in the exponential term of Equation (5) in the main text. The former condition of the linear forward problem holds because we consider event location and origin time fixed. The latter condition is satisfied if the noise level at all stations and the prior MT, $\boldsymbol{m}^p$, are kept fixed by prior knowledge. The least-squares solution in the primitive parameterization is (page 67, Tarantola, 2005),

$$\overline{\boldsymbol{m}} = \frac{1}{3N_k}\sum_{k=1}^{N_k}\sum_{j=1}^{3}\left[\left(\boldsymbol{G}_{kj}\right)^T\left(\boldsymbol{C}_D^{kj}\right)^{-1}\left(\boldsymbol{G}_{kj}\right)\right]\cdot\left[\left(\boldsymbol{G}_{kj}\right)^T\left(\boldsymbol{C}_D^{kj}\right)^{-1}\boldsymbol{d}_{obs}\right] \tag{52}$$

As the solution in Equation (52) minimizes the exponential term globally, it maximizes the likelihood function (Equation (6) in the main text). Therefore, it represents the estimate of the maximum posterior probability solution of the inverse problem.

The exponential term of the likelihood function defines the Mahalanobis distance from the prediction, $g(\Omega, \boldsymbol{m})$, to the observation, $\boldsymbol{d}_{obs}$, with uncertainty defined by the covariance matrix, C_D. The covariance matrix is positive-definite so that it can be decomposed by Cholesky's method,

$$\boldsymbol{C}_D = \boldsymbol{L}_D(\boldsymbol{L}_D)^T \tag{53}$$

where $\boldsymbol{L}_D$ is a lower triangular matrix. Therefore, the likelihood can be re-written as,

$$p(\boldsymbol{d}_{obs} \mid \boldsymbol{m}) = \frac{1}{(2\pi)^N\sqrt{|\boldsymbol{C}_D|}}\exp\left(-\frac{1}{2}\left[\boldsymbol{L}_D^{-1}\boldsymbol{d}_{obs} - \boldsymbol{L}_D^{-1}g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]^T\left[\boldsymbol{L}_D^{-1}\boldsymbol{d}_{obs} - \boldsymbol{L}_D^{-1}g\left(\widehat{\Omega}, \boldsymbol{m}\right)\right]\right). \tag{54}$$

The lower triangular matrix, $\boldsymbol{L}_D$, defines a linear transformation, such that in the transformed space, the least-squares solution (Equation 52) minimizes the Euclidean distance between transformed observed and predicted waveforms. Thus, it is more intuitive to compare recovered waveforms in the linearly transformed space (Vackář et al., 2017) rather than in the original space with the Mahalanobis distance,

$$\begin{aligned}\boldsymbol{d}'_{obs} &= \boldsymbol{L}_D^{-1}\boldsymbol{d}_{obs}\\ \boldsymbol{d}'_{pred} &= \boldsymbol{L}_D^{-1}g(\widehat{\Omega}, \boldsymbol{m})\end{aligned} \tag{55}$$

They are then called the uncertainty-weighted prediction and observation. We also define the weighted variance reduction (WVR) to measure the fit in the presence of the uncertainty covariance matrix,

$$WVR \quad = 1 - \frac{\left|\left|\boldsymbol{L}_D^{-1}\boldsymbol{d}_{obs} - \boldsymbol{L}_D^{-1}\boldsymbol{g}(\widehat{\Omega}, \mathbf{m})\right|\right|_2}{\left|\left|\boldsymbol{L}_D^{-1}\boldsymbol{d}_{obs}\right|\right|_2}. \tag{56-a}$$

The 2-norm $||.||_2$ represents the sum of squares of the traces. This is comparable to the normal standard variation reduction ignoring the presence of relevant uncertainties,

$$VR \quad = 1 - \frac{\left|\left|\boldsymbol{d}_{obs} - \boldsymbol{g}(\widehat{\Omega}, \mathbf{m})\right|\right|_2}{\left|\left|\boldsymbol{d}_{obs}\right|\right|_2}. \tag{56-b}$$

A.3 Bayesian inference for hierarchical inversion

This section elaborates on methodological details to implement the theory presented in the previous section within a probabilistic framework inference. Bayesian inference relies on the Bayes' theorem of probability

$$p(\boldsymbol{x}|\boldsymbol{d}_{obs}) = \gamma\, p(\boldsymbol{d}_{obs}|\boldsymbol{x})\, p(\boldsymbol{x}) \tag{57}$$

A model $\boldsymbol{x}$ comprises all unknowns in the physical problem we intend to infer from the observed data, $\boldsymbol{d}_{obs}$. An unknown model $\boldsymbol{x}$ in the model space $\boldsymbol{X} = \boldsymbol{M} \times \boldsymbol{H}$ has $6 + N_k$ coordinates comprising six parameters for the MT and N_k hyperparameters of noise strengths. To avoid possible confusion, it is essential to differentiate the unknown model mentioned in this context with velocity structure models. In Equation (57), the posterior probability, $p(\boldsymbol{x}|\boldsymbol{d}_{obs})$, is the probability that an unknown model $\boldsymbol{x}$ is true given the observed data, $\boldsymbol{d}_{obs}$. The prior $p(\boldsymbol{x})$ represents the prior knowledge of the distribution in model space before acquiring any actual data. The likelihood $p(\boldsymbol{d}_{obs}\,|\boldsymbol{x})$ measures the goodness of fit between the observation and the prediction made by the unknown model through a physical forward problem. Lastly, γ is a normalization factor to unify the integration of the posterior probability over the entire model space. Equation (57) can be understood as a learning process in which prior knowledge of the unknown model space is improved by the observed data (via the likelihood) to yield updated knowledge of the model space.

The Metropolis-Hasting (M-H) method (Hastings, 1970; Metropolis et al., 1953) is widely used to evaluate an arbitrary probability distribution because of its efficacy, which utilizes a finite Markov chain to sample the most likely region of the probability distribution intentionally. Here we use the M-H method to sample the posterior probability distribution Equation (57) over the unknown model space. The chain starts at an initial model, which can be either picked randomly or as the output from another crude inversion method. The initial model is accepted unconditionally, and the chain moves to the next iteration. From the second iteration of the chain, a new model is proposed as a next move by deviating the current model along a random coordinate, say x_i,

$$x_i' = x_i + \delta_i, \tag{58}$$

in which δ_i is a random value drawn from a proposed distribution. The proposal is deemed valid if x_i' falls within the pre-set bounds of the present coordinate. The proposed model will be accepted and added to the chain with probability

$$\alpha = \min\left(\frac{p(\boldsymbol{x}') \cdot p(\boldsymbol{d}_{obs} \mid \boldsymbol{x}') \cdot p(\boldsymbol{x}' \mid \boldsymbol{x})}{p(\boldsymbol{x}) \cdot p(\boldsymbol{d}_{obs} \mid \boldsymbol{x}) \cdot p(\boldsymbol{x} \mid \boldsymbol{x}')}, 1\right). \tag{59}$$

$p(\boldsymbol{x} \mid \boldsymbol{x}')$ and $p(\boldsymbol{x}'|\boldsymbol{x})$ are transition probabilities from the current model to the proposal and vice versa. If the proposed distribution selected is symmetric, such as the Gaussian or Cauchy distribution, the transition probabilities are equal, and their term can be dropped from the acceptance probability Equation (59). Furthermore, the model's prior probability is defined as homogeneous, so we have $p(\boldsymbol{x}) = p(\boldsymbol{x}')$. Thus, the acceptance probability can be reduced as,

$$\alpha = \min(\frac{p(\boldsymbol{d}_{obs} \mid \boldsymbol{x}')}{p(\boldsymbol{d}_{obs} \mid \boldsymbol{x})}, 1) \tag{60}$$

The acceptance probability α only involves the likelihood term of the current and proposed models as defined in Equation (60).

In this study, we propose a new model using the Cauchy distribution,

$$p(\boldsymbol{x} \mid \boldsymbol{x}') = p(\boldsymbol{x}' \mid \boldsymbol{x}) = \frac{1}{\pi(1 + \delta_i^2)}, \tag{61}$$

where p_γ is the density probability function of the standard Cauchy distribution. Because the Cauchy distribution has "thick" tails, this helps the sampler become relatively exploratory and efficiently avoid being trapped in local maxima.

It is equivalent but markedly more convenient to compute the natural logarithm of the likelihood (also known as the log-likelihood), as it avoids the underflow or overflow errors when computing large exponential factors,

$$\begin{aligned} \log \alpha = & -\frac{1}{2}(\log|C_D| - [d_{obs} - g(m)]^T C_D^{-1}[d_{obs} - g(m)]) \\ & + \frac{1}{2}(\log|C_D| - [d_{obs} - g(m)]^T C_D^{-1}[d - g(m)]). \end{aligned} \tag{62}$$

With the log-likelihood, the acceptance condition defined in Equation (62) now becomes,

$$\log r \leq \log \alpha. \tag{63}$$

At the beginning of the chain, the theoretical covariance matrix is pre-computed with a prior MT, $\boldsymbol{m}^p$ without updating. It will only get updated through major phases.

Table 3. Glossary of Symbols

Symbols	*Verbal Definitions*
N_k	The number of seismic stations being used.
N_m	The number of randomly perturbed velocity structures.
N	The number of points in each seismogram.
m	A generic moment tensor (MT).
$\overline{m}$	MT parameterized in the primitive coordinates.
m^p	Prior MT to define the structure-uncertainty covariance matrix.
m_0	MT solution estimated preliminarily to determine for time shifts of observed waveforms.
ϵ_d	Observational (data) error.
C_d	Data noise covariance matrix.
L_d	Lower Cholesky decomposition of C_d
ϵ_t	Structural uncertainty error.
C_t	Structural uncertainty covariance matrix.
C_D	The combined covariance matrix of data and structural uncertainty.
L_D	Lower Cholesky decomposition of C_D
Ω	A generic velocity structure of the Earth.
$\widetilde{\Omega}$	Actual Earth structure, which is generally unknown.
$\widehat{\Omega}$	Reference structure model representing subjective knowledge of the Earth structure.
Ω_i	A randomly perturbed structure model from the reference model.

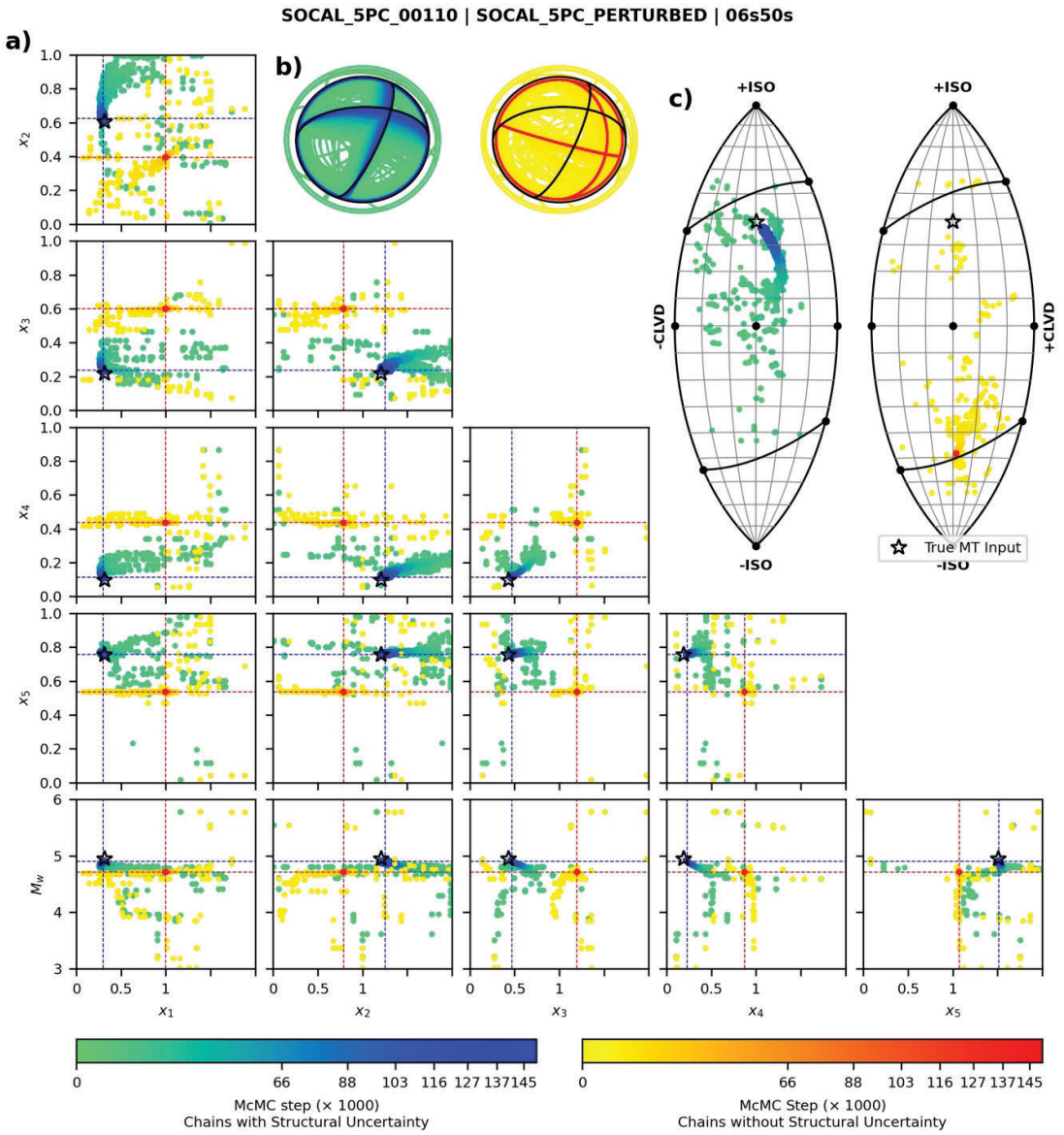

Figure 50. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels. (See further description of the scenarios in Section 4.2). Synthetic data are filtered in the 6–50 s period band. Winter color scheme (transitioning from green to blue) is used for McMC steps in the MT inversion incorporating structural uncertainty, which is computed hypothetically with the true MT input. The autumn color scheme is used for McMC steps in the MT inversion that retains only the data noise components and ignores structural uncertainty. We use the power-law of 2/5 to normalize McMC steps in the color schemes to better visualize the variation at the early stage of the inversion chains, which are regarded as the burning stage. The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band considered. *(a)* Each sub-panel shows a pair of parameters in Tashiro's (1977) parameterization of MTs (see definition in Equation (15)). Open black triangles show the true input MT. The red and blue dashed lines indicate the position of the mean MT solutions of the post-burning stages of the chains. *(b)*

Beach balls show the fault planes corresponding to the double couple part of the MT solutions in the inversion chains of the two scenarios. Fault planes corresponding to the true MT solution are in black. *(c)* In lune diagrams of source types (Tape & Tape, 2012), we compare the evolution of MT solution in two scenarios. Open black stars mark the source type of the true MT input in both diagrams.

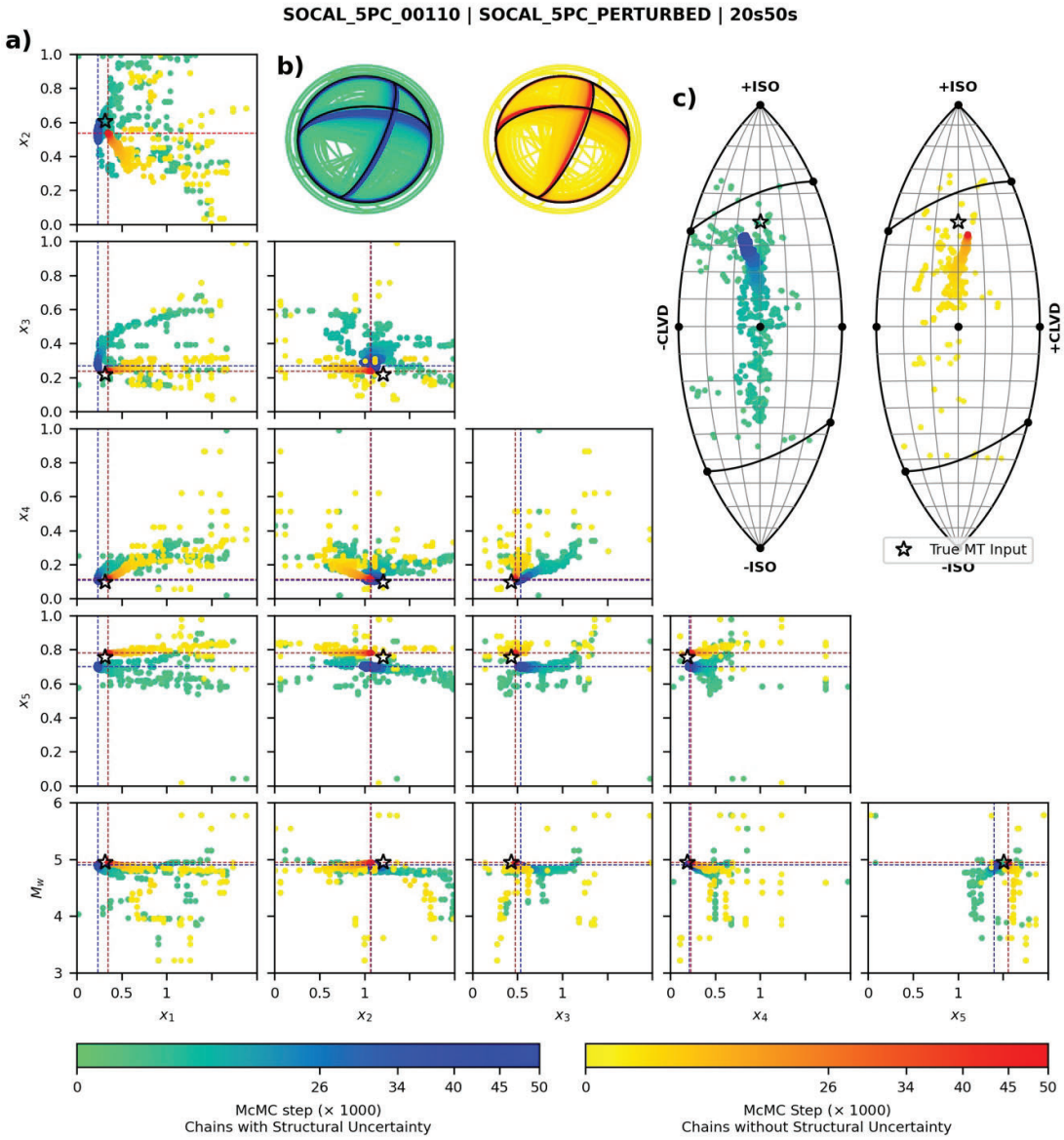

Figure 51. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structural uncertainty, given true input data noise levels. (See further description of the scenarios in Section 4.2). Synthetic data are filtered between 20–50 s. See the caption of Figure 60 for other details.

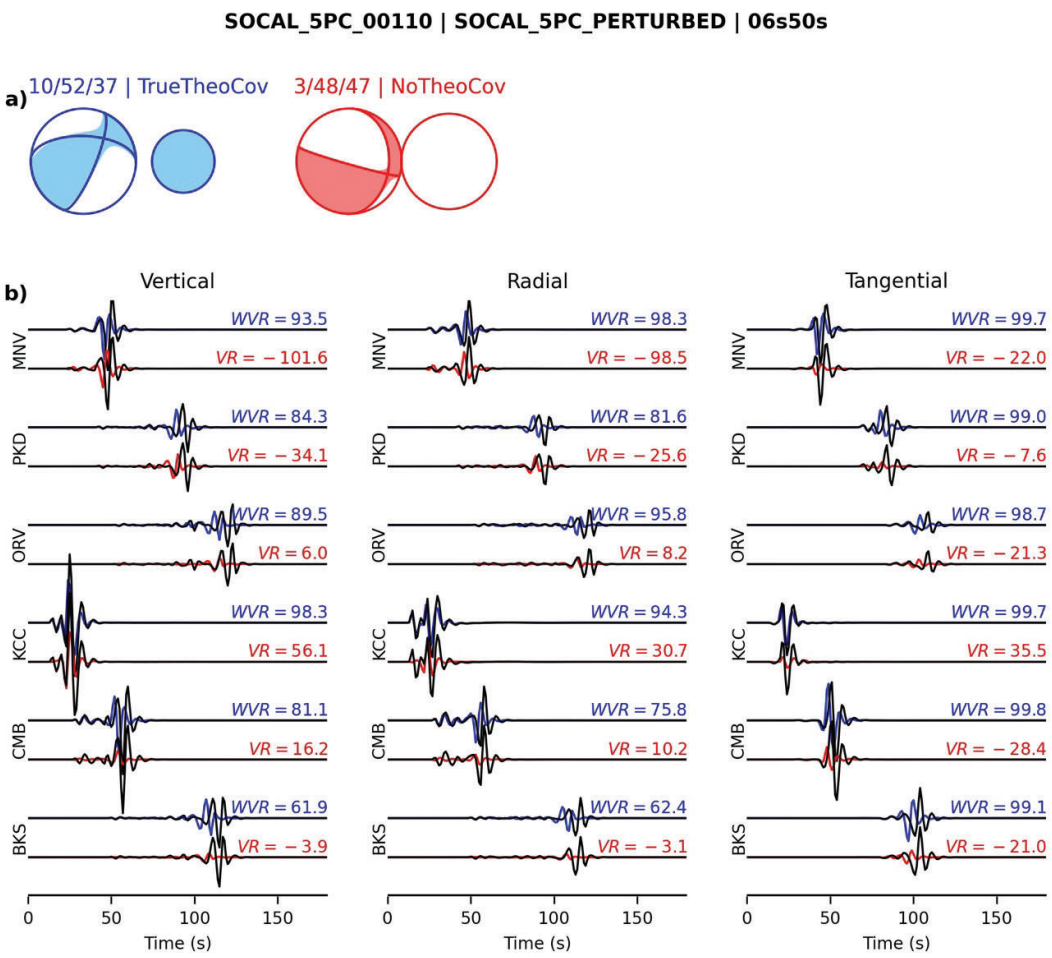

Figure 52. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels. (See further description of the scenarios in Section 4.2). Synthetic data are bandpass filtered in the 6–50 s period band. In all sub-panels, the red colour corresponds to the mean MT solution obtained when ignoring structural uncertainty. In contrast, blue corresponds to the MT solution inferred with structural uncertainty covariance accounted for (see the labels on top of the beach balls). The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$ used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band considered. *(a)* Decomposition of mean MT solutions shown in Figure 4 into deviatoric and isotropic parts. The beachball sizes are proportional to the percentages. Percentages of CLVD, DC, ISO parts in the full MT are displayed on top of the deviatoric beachballs. A filled isotropic ball means an explosion, while a non-filled isotropic ball means an implosion. *b)* Waveform fit corresponding to the mean MT solutions of the two numerical scenarios. Input synthetic seismograms are black traces, while the recovered waveforms are in red and blue colors. On the right-hand side of each sub-panel is the weighted variance reduction (WVR – defined in Equation 56-a) for waveforms recovered with structural uncertainty and the variance reduction (VR – defined in Equation 56-b) for waveforms recovered without structural uncertainty.

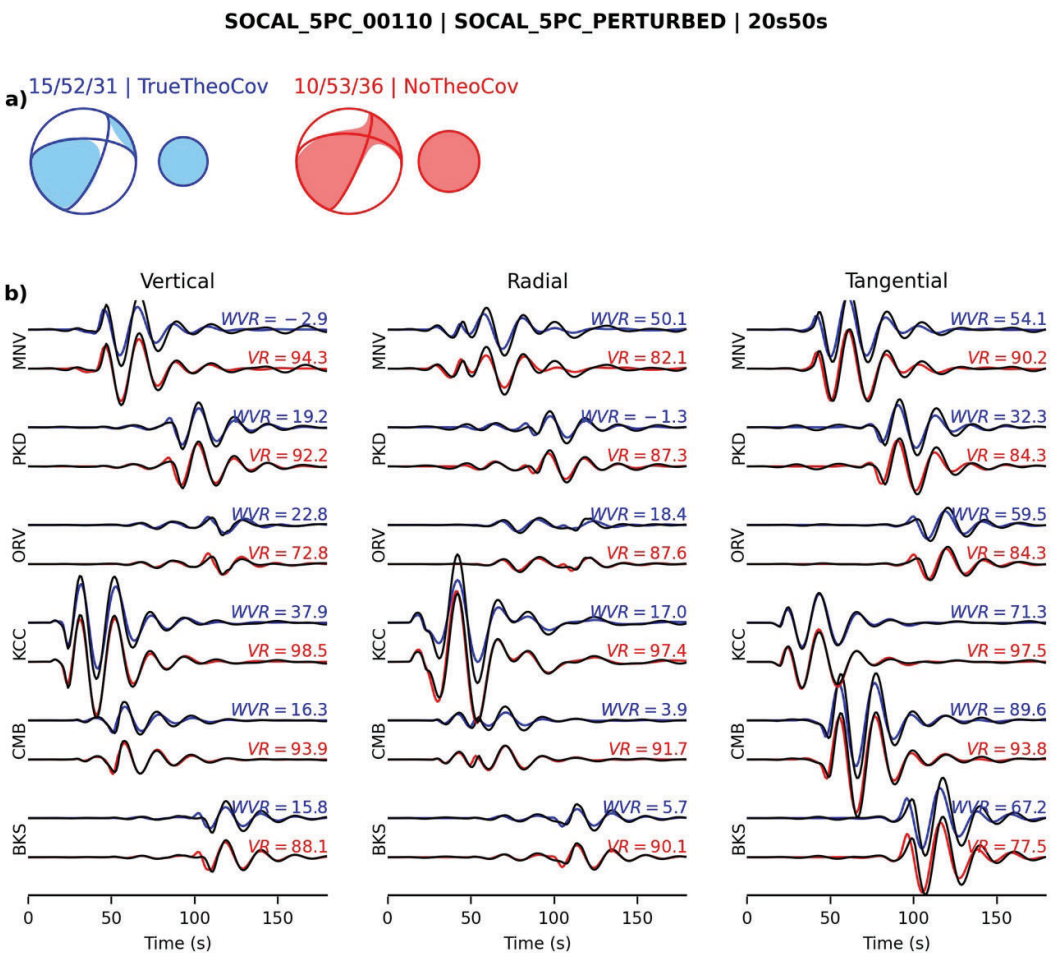

Figure 53. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty, given true input data noise levels. (See further description of the scenarios in Section 4.2). Synthetic data are bandpass filtered between 20–50 s. See the caption of Figure 52 for other details.

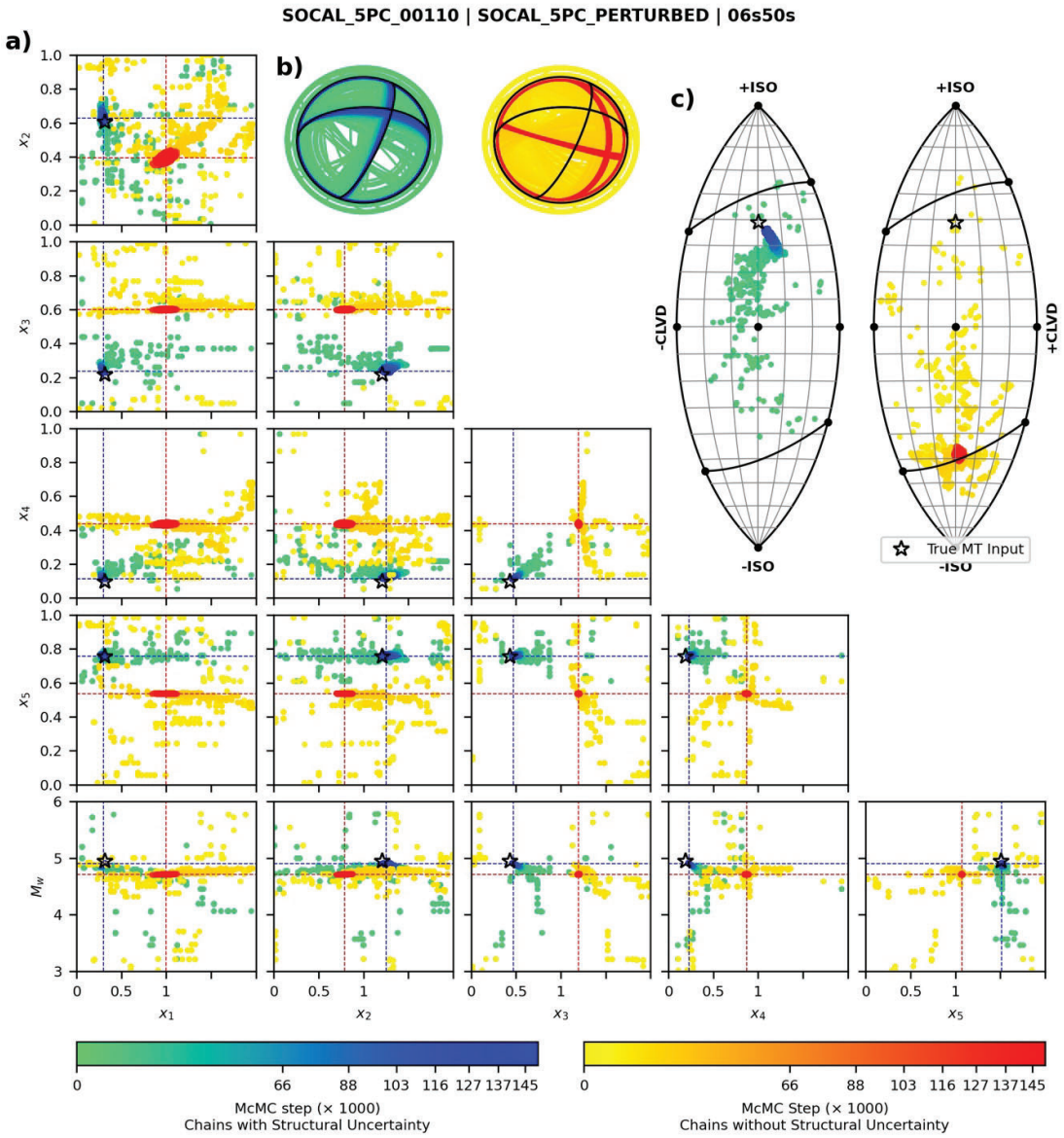

Figure 54. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, *h*, are hyper-parameters to be recovered in the inversion. (See further description of the scenarios in Section 4.3). Synthetic data are filtered between 6–50 s. See the caption of Figure 50 for all other details.

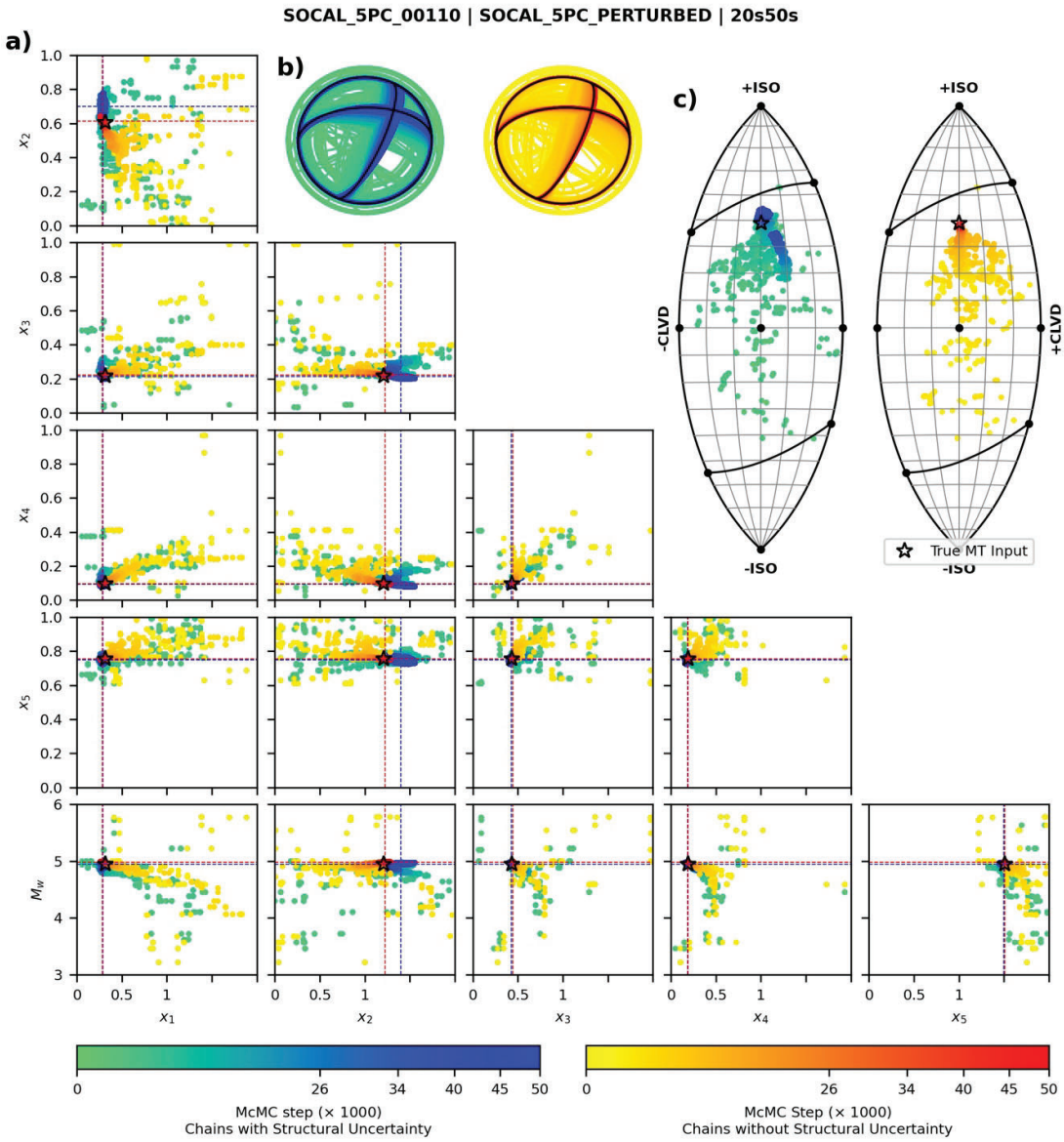

Figure 55. Comparison of inversion results in the two synthetic scenarios when accounting for and not accounting for structure uncertainty, given station-specific data noise scaling factors, *h*, are hyper-parameters to be recovered in the inversion. (See further description of the scenarios in Section 4.3). Synthetic data are filtered between 20–50 s. See the caption of Figure 50 for all other details.

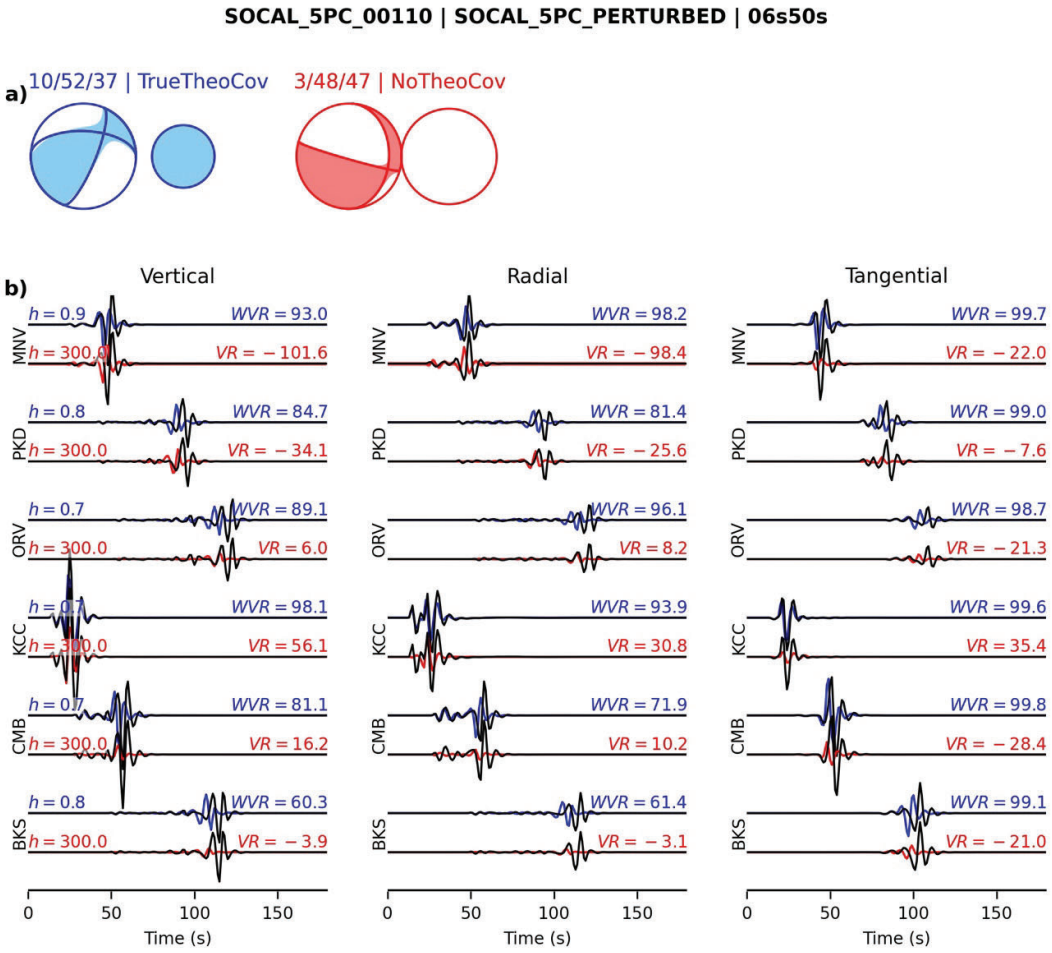

Figure 56. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 4.3). Synthetic data are filtered between 6–50 s. See the caption of Figure 5 for other details. In addition to details in Figure 52, the numbers on left of the vertical traces are recovered noise scaling factors, $\boldsymbol{h_k}$, for the stations.

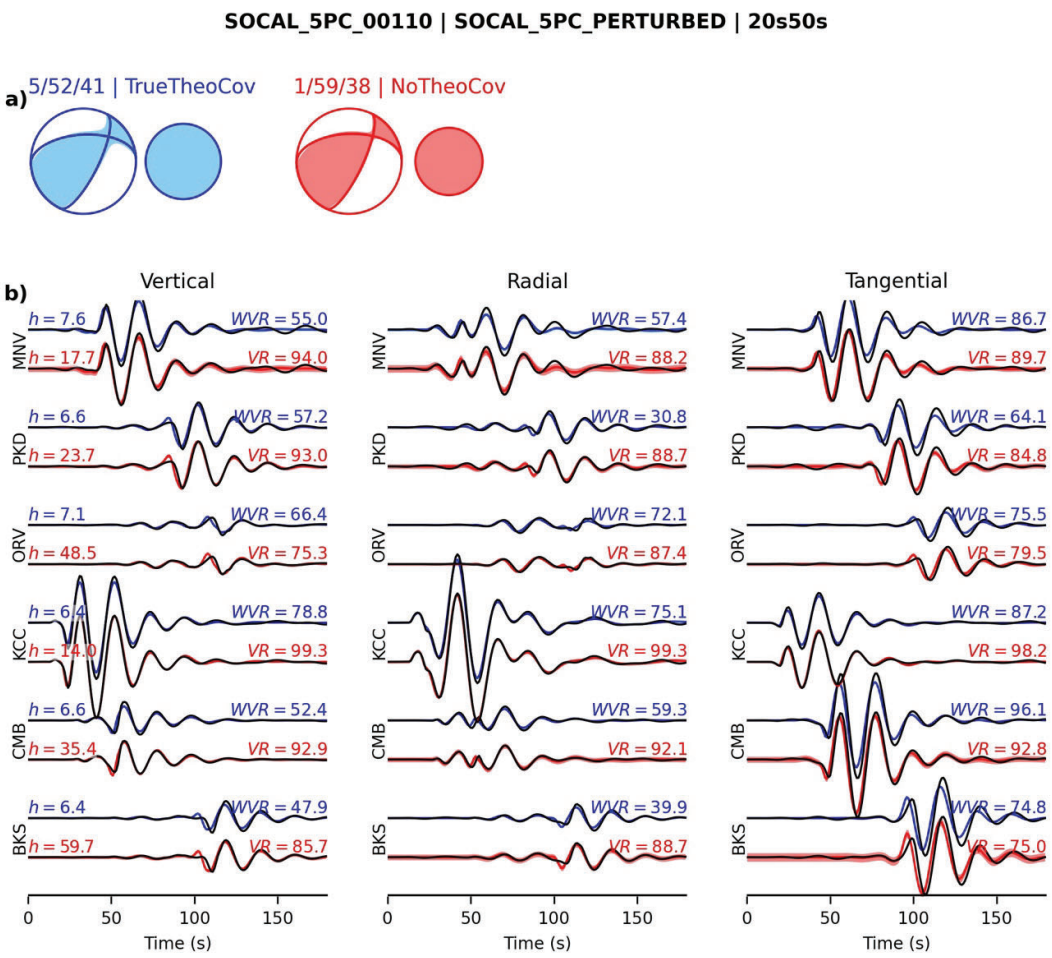

Figure 57. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 4.3). Synthetic data are filtered between 20–50 s. See the caption of Figure 52 for other details. In addition to details in Figure 52, the numbers on left of the vertical traces are recovered noise scaling factors, $\boldsymbol{h_k}$, for the stations.

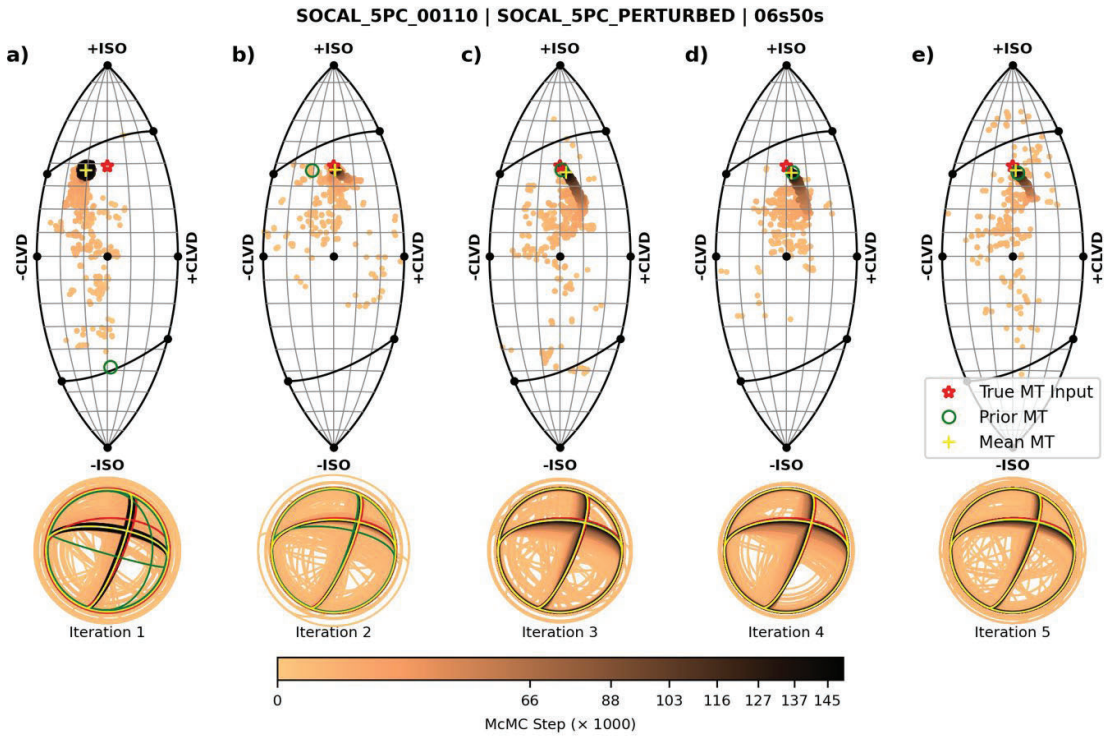

Figure 58. Evolution of McMC chains the iterative inversion scheme. (See description of the scenario in Section 4.4). Synthetic data are filtered in the 6–50 s period band. Progression of MT solutions in the chains is shown by their locations in the lune diagrams representing source types and by double-couple parts via beach balls representing their fault planes. The title indicates the "true" Earth model, $\tilde{\boldsymbol{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\boldsymbol{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band being considered. Five iterations involved in the scheme are plotted in five sub-figures from *a)* to *e)*. In each iteration, the prior MT is required to compute the structural uncertainty covariance matrix, $\boldsymbol{m^p}$ – see Table 3. It is denoted by an cyan open circle in the lune diagram and is plotted in cyan in the beachball diagram. Similarly, the mean MT solution of the post-burning chain is shown by a yellow plus sign and plotted in yellow in the beachball, and the true input MT is shown by an open red star and plotted in red the beachball. Like Figure 50 and Figure 51, we use the power-law of 2/5 to normalize the steps in the reversed copper scale to visualize the evolution of the chains during the burning stage.

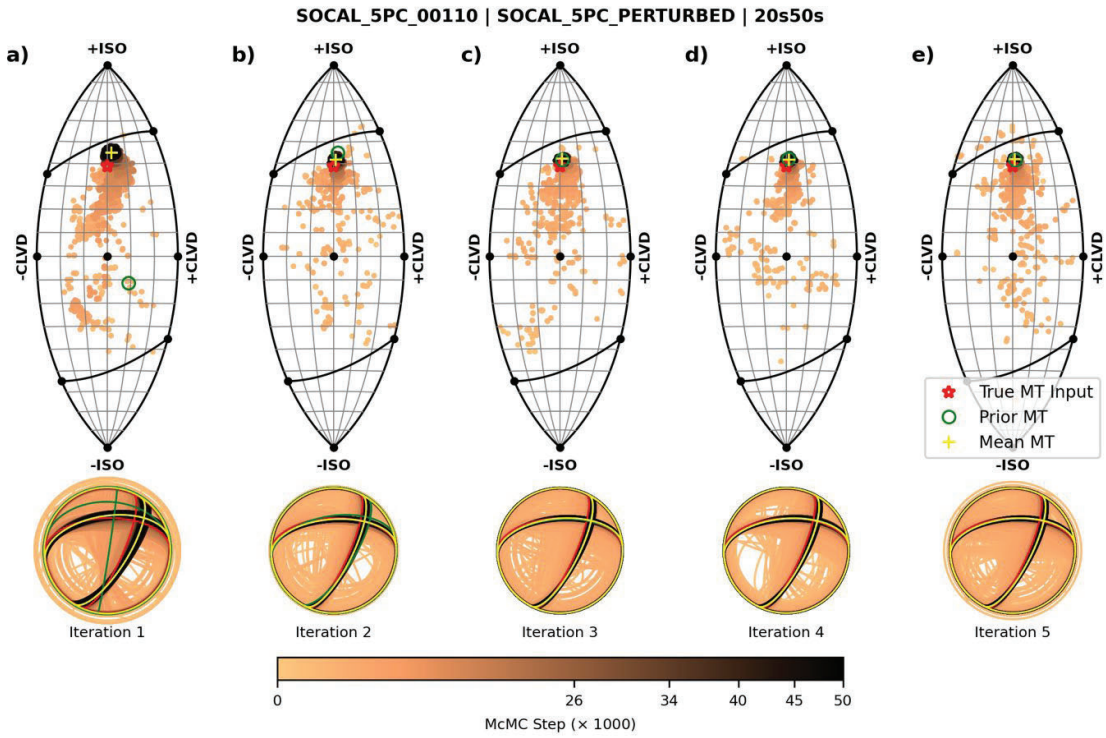

Figure 59. Evolution of McMC chains of the iterative inversion scheme. (See description of the scenario in Section 4.4). Synthetic data are filtered between 20–50 s. See the caption of Figure 58 for other details.

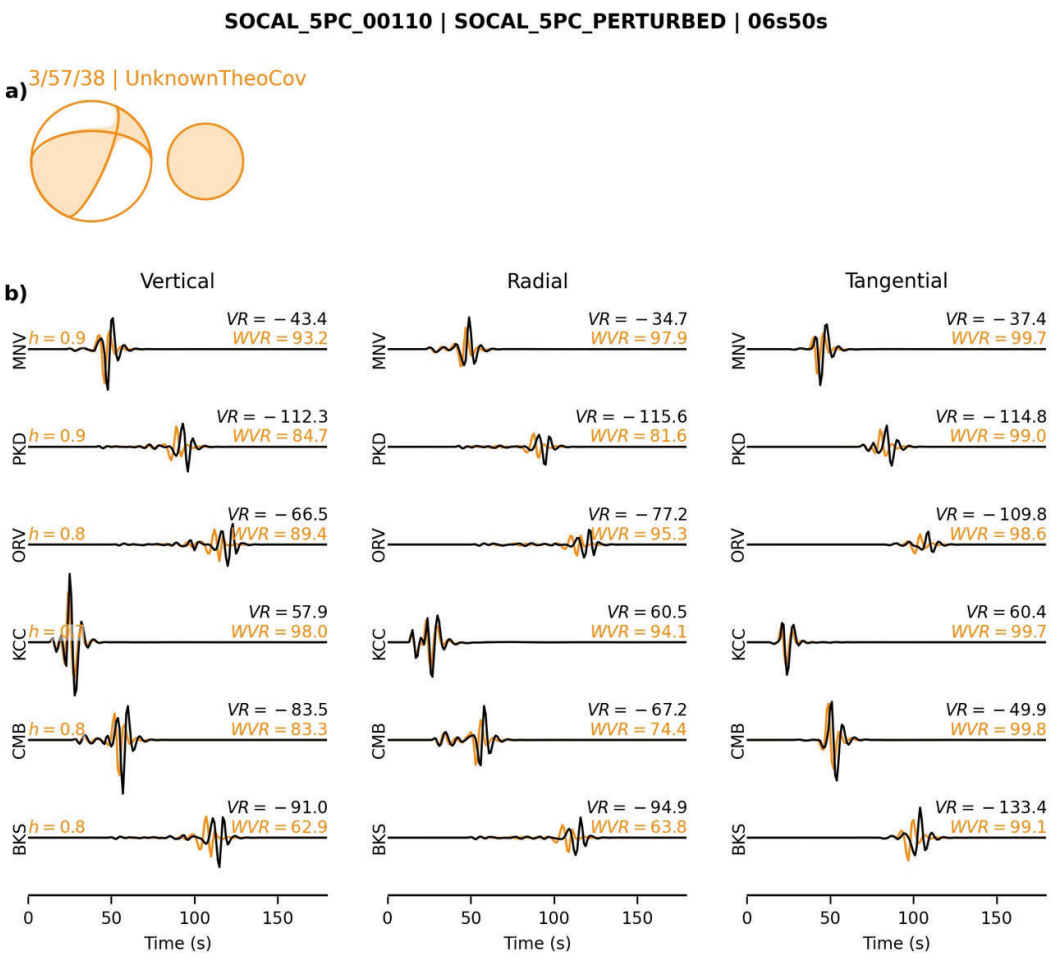

Figure 60. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. Synthetic data are bandpass filtered in the 6–50 s period band. The figure title consists of the "true" Earth model, $\widetilde{\mathbf{\Omega}}$, used to generate "fake" data, the reference Earth model, $\widehat{\mathbf{\Omega}}$, used to generate predicted waveforms and its assumed level of uncertainty (e.g., 5% in this experiment), and the period band being considered. *a)* Decomposition of the mean MT solution into deviatoric and isotropic parts. The beachball sizes are proportional to their percentages in the full MT solution. Percentages of CLVD, DC, ISO parts of the full MT are displayed on top of the deviatoric beachballs. A filled isotropic ball signifies an explosion, while a non-filled isotropic ball signifies an implosion. *b)* Waveform fit corresponding to the mean MT solutions. Input synthetic seismograms are black traces, while the recovered waveforms are copper colored. On the right-hand side of each sub-panel are the weighted variance reduction (WVR – defined in Equation 56-a) for waveforms recovered with structural uncertainty and the variance reduction (VR – defined in Equation 56-b) between the recovered waveforms and the "fake" data. On the left are recovered data noise scaling factors, $\boldsymbol{h_k}$, for the stations.

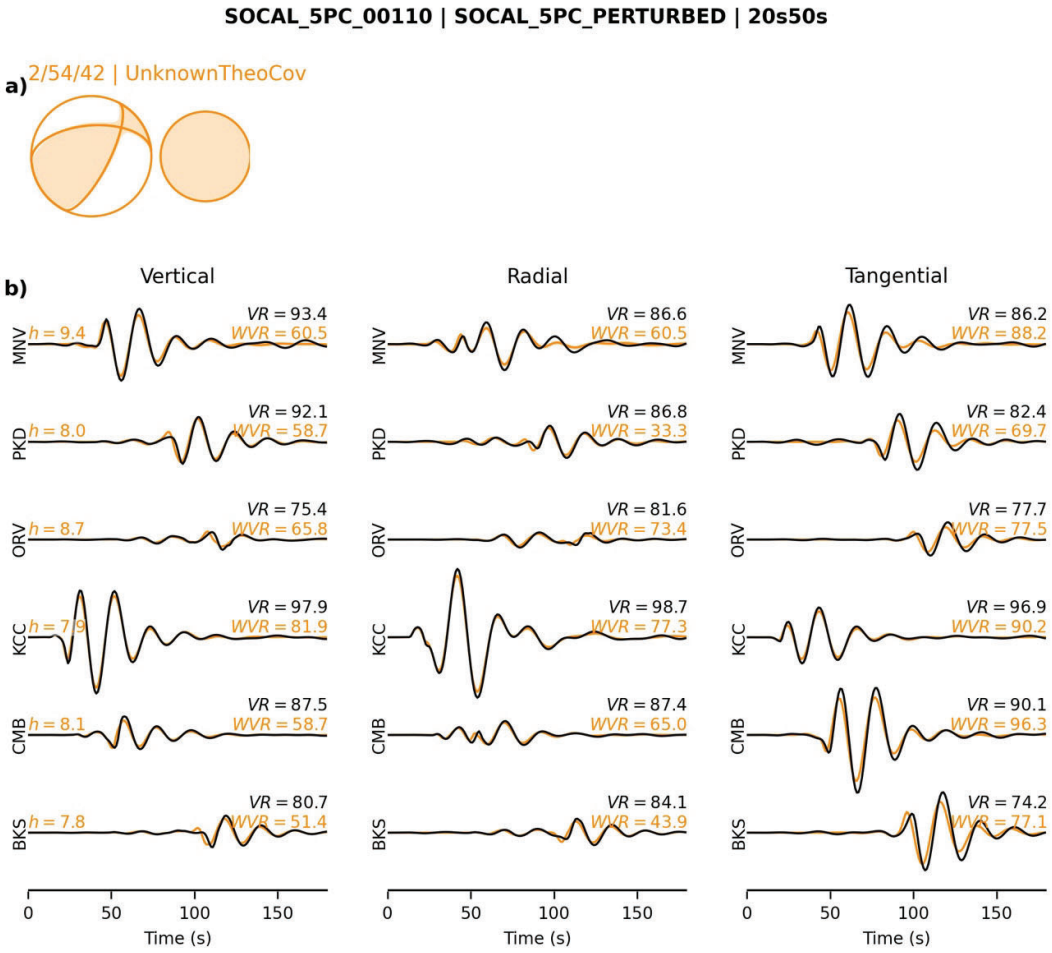

Figure 61. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. Synthetic data are bandpass filtered between 20–50 s. See the caption of Figure 60 for other details.

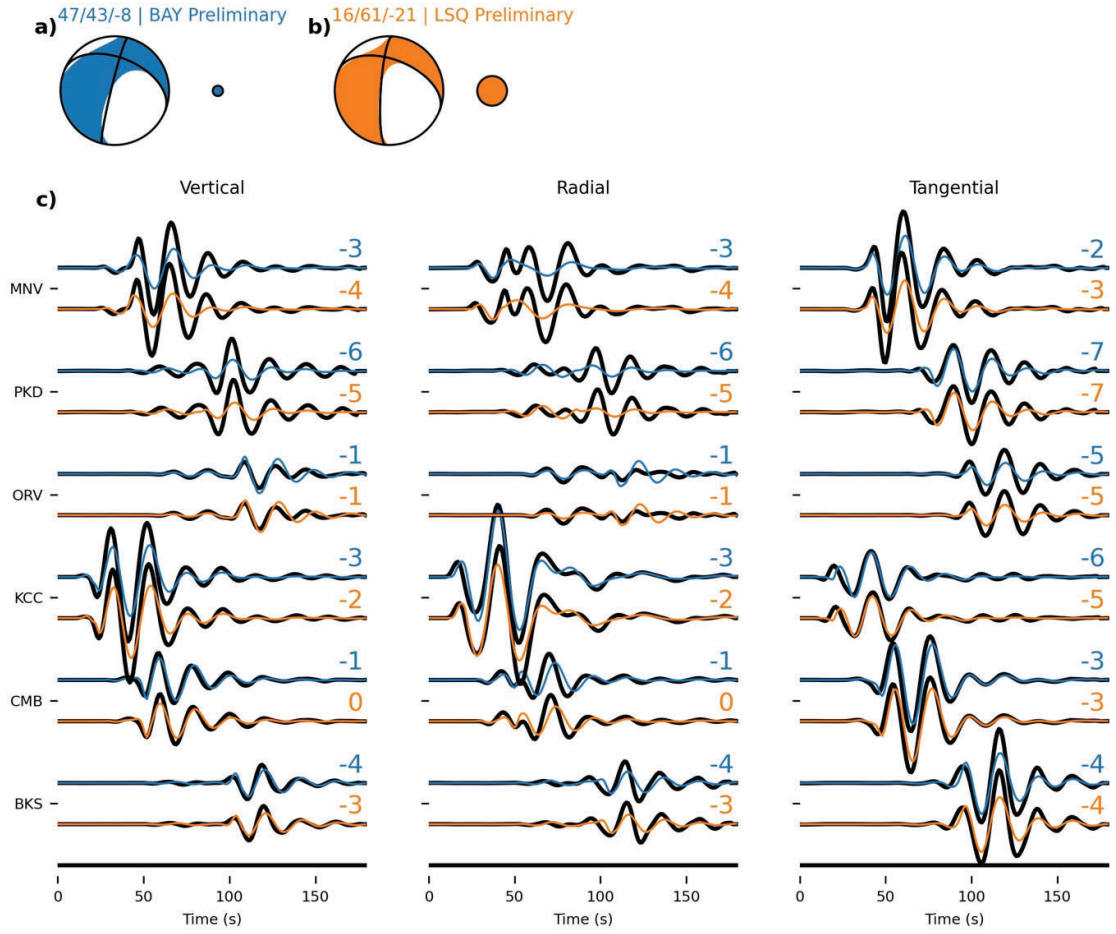

Figure 62. Preliminary MT solutions by hierarchical Bayesian inversion and least-squares inversion. *a)* and *b)* Deviatoric and isotropic parts of the full MTs of the two solutions. On top of the deviatoric beachball are percentages of CLVD, DC and ISO parts. Beachball sizes are scaled by their percentages. *c)* Observed waveforms are shown in black traces while predictions corresponding to the preliminary solutions by Bayesian and least squares inversion are in blue and orange respectively. The observed waveforms are shifted to align with the synthetic waveform calculated with the preliminary MT solutions. Colour numbers over corresponding waveforms are time shifts to match the (black) observed waveforms.

Figure 63. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 5.2). Observed data are filtered between 20–50 s. The initial time shifts for the observed waveform are determined from the preliminary solution that was obtained by the hierarchical inversion (see more details in Section 5.1). See the caption of Figure 52 for other details. In addition to details in Figure 52, the numbers left of the vertical traces are recovered noise scaling factors, $\boldsymbol{h_k}$, for the stations.

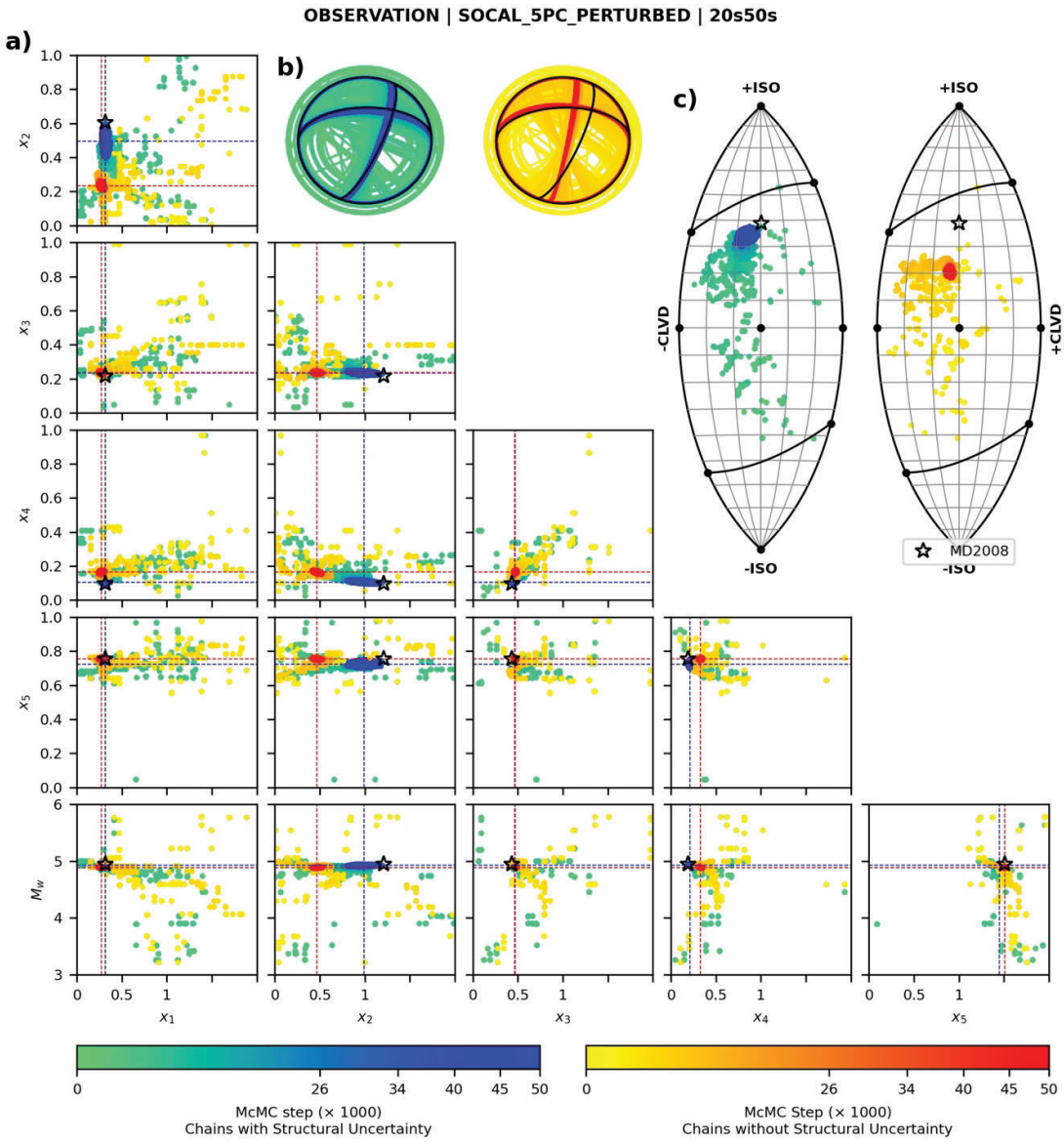

Figure 64. Comparison of inversion results in the two scenarios for real data inversion when accounting for and not accounting for structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 4.2). Observed data are filtered between 20–50 s. The initial time shifts for the observed waveform are determined from the preliminary solution that was obtained by the least squares method (shown in Figure 63; see more details in Section 5.2). See the caption of Figure 50 for all details, except we use the open black stars for the MD2008 solution (Minson & Dreger, 2008) rather than being the true MT input in the real application.

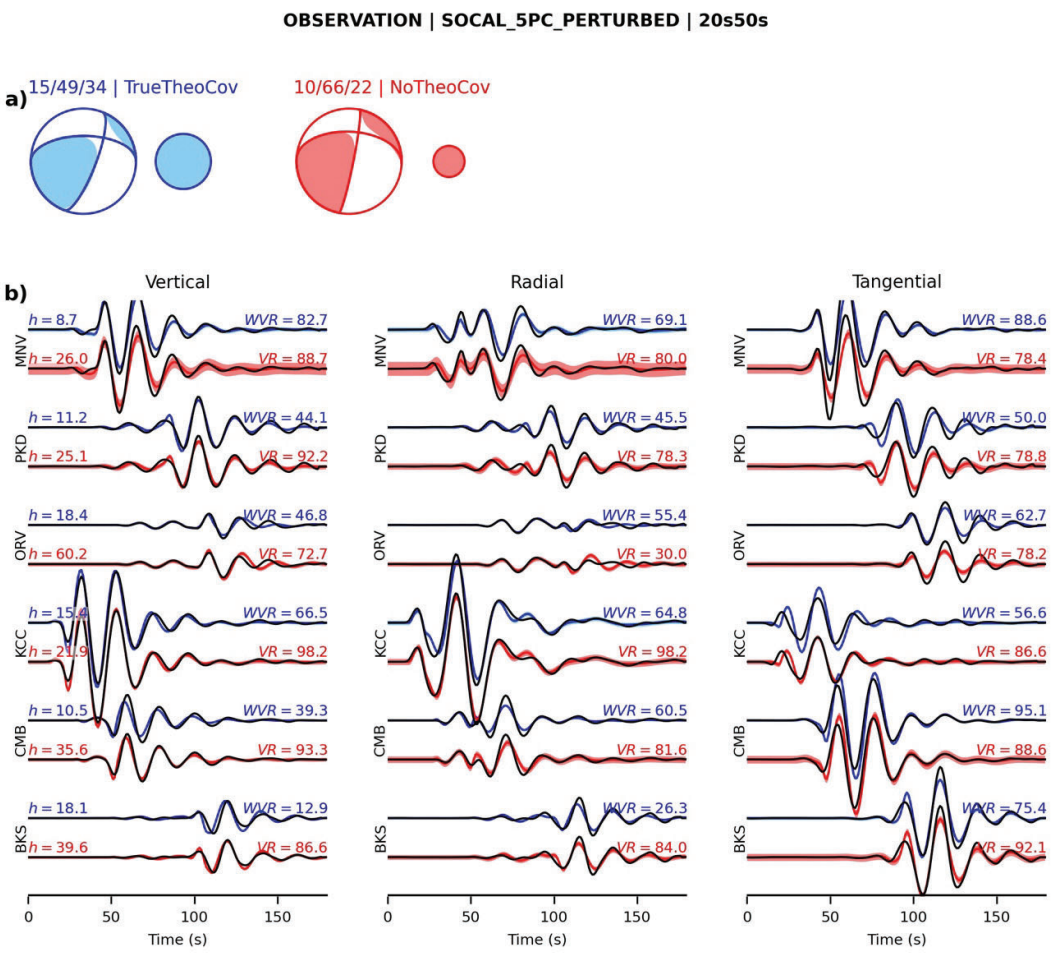

Figure 65. Comparison of MT decomposition and waveform fit for two numerical scenarios when accounting or not accounting for the structural uncertainty and unknown data noise levels. (See further description of the scenarios in Section 5.2). Observed data are filtered between 20–50 s. The initial time shifts for the observed waveform are determined from the preliminary solution that was obtained by the least squares method (shown in Figure 63; see more details in Section 5.2). See the caption of Figure 52 for other details. Numbers on left of the vertical traces are recovered noise scaling factors, $\boldsymbol{h_k}$, for the stations.

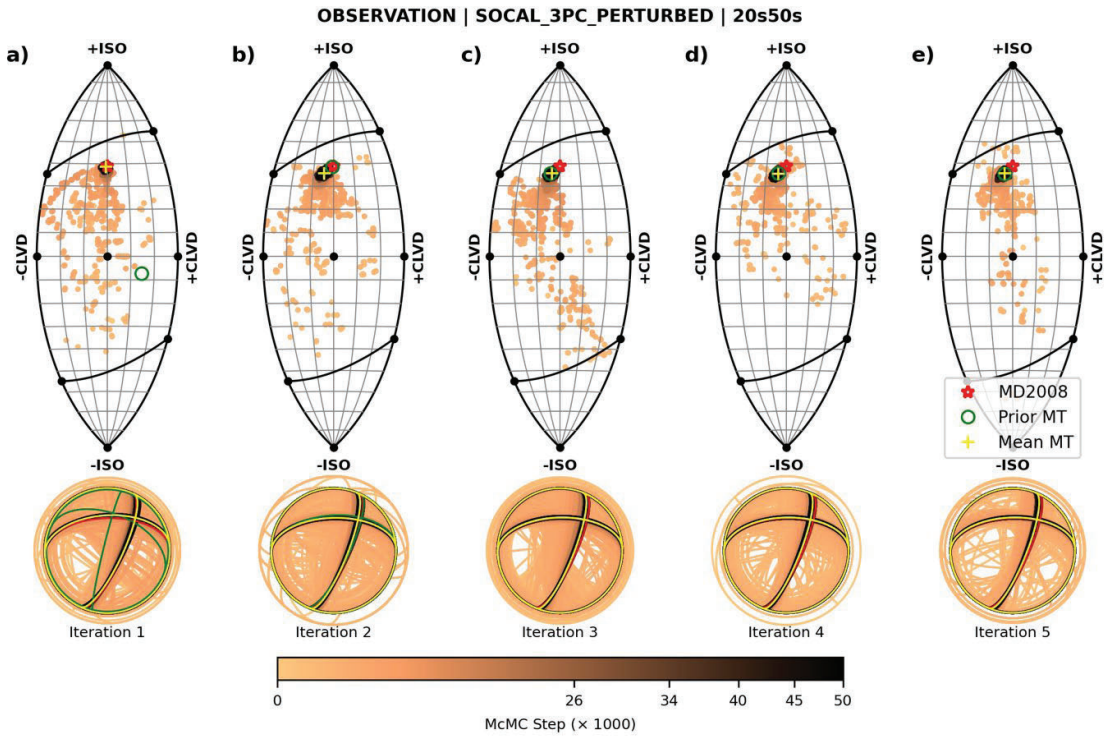

Figure 66. Evolution of McMC chains in the iterative inversion scheme. (See description of the scenario in Section 5.2). Observed data are filtered between 20–50 s. This figure is equivalent to Figure 12 in the main text, but the level of structural uncertainty is assumed to be $\boldsymbol{\kappa = 3}$% here instead of $\boldsymbol{\kappa = 5}$%. See the caption of Figure 58 for other details.

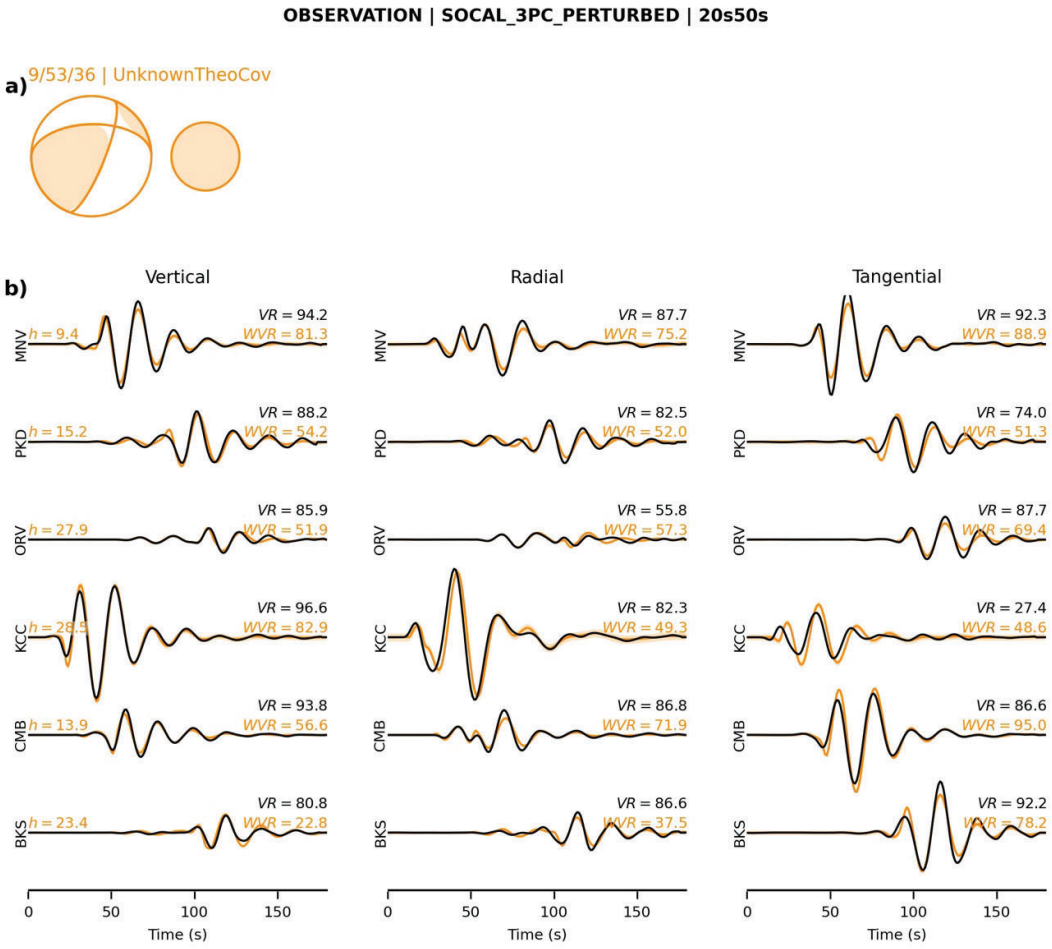

Figure 67. MT decomposition and waveform fit corresponding to the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. This figure is equivalent to Figure 13 in the main text, but the level of

structural uncertainty is assumed to $\boldsymbol{\kappa = 3\%}$ here instead of $\boldsymbol{\kappa = 5\%}$. See the caption of Figure 60 for all details.

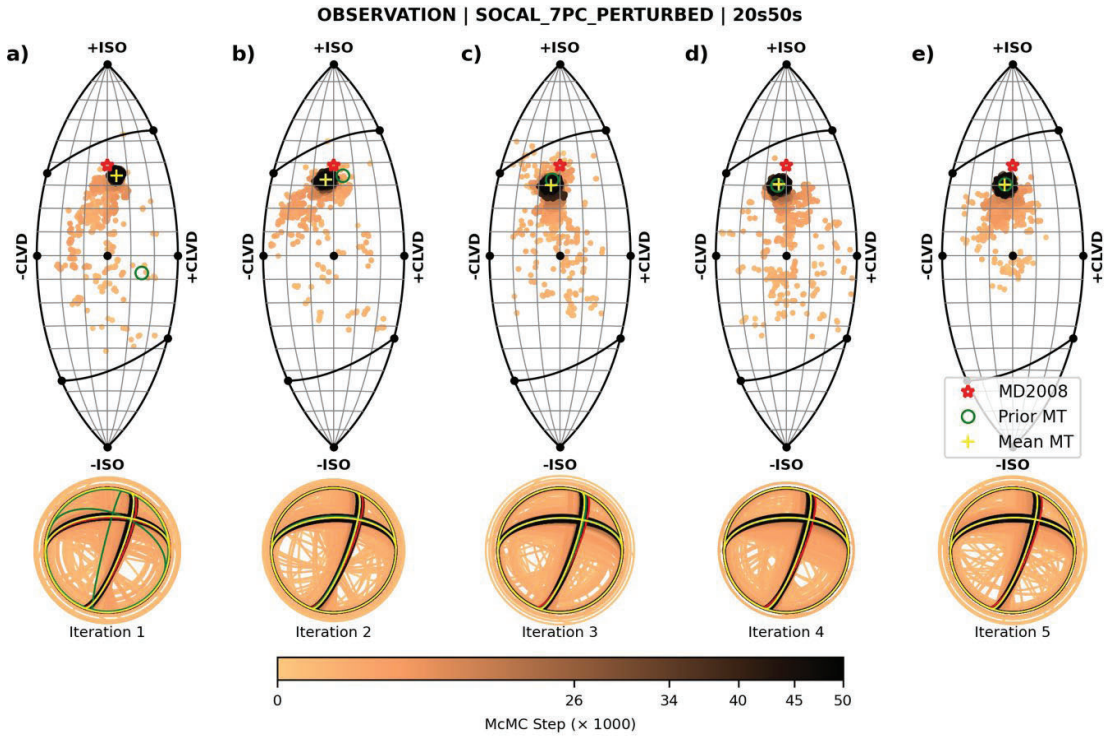

Figure 68. Evolution of McMC chains in the iterative inversion scheme. (See description of the scenario in Section 5.2). Observed data are filtered between 20–50 s. This figure is equivalent to Figure 12 in the main text, but the level of structural uncertainty is assumed to be $\boldsymbol{\kappa = 7\%}$ here instead of $\boldsymbol{\kappa = 5\%}$. See the caption of Figure 58 for other details.

Figure 69. MT decomposition and waveform fit corresponding the MT solution obtained from the iterative inversion scheme, which is the mean of the post-burning chain in the fifth iteration. This figure is equivalent to Figure 13 in the main text, but the level of structural uncertainty is assumed to $\boldsymbol{\kappa = 7}$% here instead of $\boldsymbol{\kappa = 5}$%. See the caption of Figure 60 for all details.

REFERENCES

Aki, K. and Richards, P. G. (2002), *Quantitative Seismology* (Second edition), University Science Books, Sausalito, CA.

Allen, R. M. and Melgar, D. (2019), Earthquake Early Warning: Advances, Scientific Challenges, and Societal Needs, *Annual Review of Earth and Planetary Sciences*, *47*(1), pp. 361–388, https://doi.org/10.1146/annurev-earth-053018-060457.

Alvizuri, C. and Tape, C. (2018), Full Moment Tensor Analysis of Nuclear Explosions in North Korea, *Seismological Research Letters*, *89*(6), pp. 2139–2151, https://doi.org/10.1785/0220180158.

Alvizuri, C., Silwal, V., Krischer, L., and Tape, C. (2018), Estimation of Full Moment Tensors, Including Uncertainties, for Nuclear Explosions, Volcanic Events, and Earthquakes, *Journal of Geophysical Research: Solid Earth*, *123*(6), pp. 5099–5119, https://doi.org/10.1029/2017JB015325.

Aster, R. C., Borchers, B., and Thurber, C. H. (2013), Bayesian Methods, In *Parameter Estimation and Inverse Problems,* pp. 253–280, Elsevier, https://doi.org/10.1016/B978-0-12-385048-5.00011-2.

Bassin, C., Laske, G., and Masters, G. (2000), The Current Limits of Resolution for Surface Wave Tomography in North America, EOS Trans AGU, 81, F897.

Bayes, Mr. and Price, Mr. (1763), An Essay towards Solving a Problem in the Doctrine of Chances, By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions (1683-1775)*, *53*, pp. 370–418.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N. (2012), Transdimensional inversion of receiver functions and surface wave dispersion, *Journal of Geophysical Research: Solid Earth*, *117*, B02301, https://doi.org/10.1029/2011JB008560.

Bouchon, M. (1981), A simple method to calculate Green's functions for elastic layered media, *Bulletin of the Seismological Society of America*, *71*(4), pp. 959–971, https://doi.org/10.1785/BSSA0710040959.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., et al. (2016), Global adjoint tomography: first-generation model, *Geophysical Journal International*, *207*(3), pp. 1739–1766, https://doi.org/10.1093/gji/ggw356.

Cesca, S., Heimann, S., Kriegerowski, M., Saul, J., and Dahm, T. (2017), Moment Tensor Inversion for Nuclear Explosions: What Can We Learn from the 6 January and 9 September 2016 Nuclear Tests, North Korea? *Seismological Research Letters*, *88*(2A), pp. 300–310, https://doi.org/10.1785/0220160139.

Chiang, A., Dreger, D. S., Ford, S. R., and Walter, W. R. (2014), Source Characterization of Underground Explosions from Combined Regional Moment Tensor and First-Motion Analysis, *Bulletin of the Seismological Society of America*, *104*(4), pp. 1587–1600, https://doi.org/10.1785/0120130228.

Chiang, A., Ichinose, G. A., Dreger, D. S., Ford, S. R., Matzel, E. M., Myers, S. C., and Walter, W. R. (2018), Moment Tensor Source-Type Analysis for the Democratic People's Republic of Korea–Declared Nuclear Explosions (2006–2017) and 3 September 2017 Collapse Event, *Seismological Research Letters*, *89*(6), pp. 2152–2165, https://doi.org/10.1785/0220180130.

Dettmer, J. and Dosso, S. E. (2012), Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains, *The Journal of the Acoustical Society of America*, *132*(4), pp. 2239–2250, https://doi.org/10.1121/1.4746016.

Dettmer, J., Dosso, S. E., and Holland, C. W. (2007), Uncertainty estimation in seismo-acoustic reflection travel time inversion, *The Journal of the Acoustical Society of America*, *122*(1), pp. 161–176, https://doi.org/10.1121/1.2736514.

Dreger, D. S. and Helmberger, D. V. (1990), Broadband modelling of local earthquakes, *Bulletin of the SeismologicalSocietyof America*, *80*(5), pp. 1162–1179.

Dreger, D. S., Gritto, R., and Nelson, O. (2021), Path Calibration of the Democratic People's Republic of Korea 3 September 2017 Nuclear Test, *Seismological Research Letters,* https://doi.org/10.1785/0220210105.

Dreger, D.S., Tkalčić, H., and Johnston, M. (2000), Dilational Processes Accompanying Earthquakes in the Long Valley Caldera, *Science*, *288*(5463), pp. 122–125, https://doi.org/10.1126/science.288.5463.122.

Duputel, Z., Rivera, L., Fukahata, Y., and Kanamori, H. (2012), Uncertainty estimations for seismic source inversions, *Geophysical Journal International*, *190*(2), pp. 1243–1256, https://doi.org/10.1111/j.1365-246X.2012.05554.x.

REFERENCES (Continued)

Duputel, Z., Agram, P. S., Simons, M., Minson, S. E., and Beck, J. L. (2014), Accounting for prediction uncertainty when inferring subsurface fault slip, *Geophysical Journal International*, *197*(1), pp. 464–482, https://doi.org/10.1093/gji/ggt517.

Ekström, G., Nettles, M., and Dziewoński, A. M. (2012), The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and Planetary Interiors*, pp. *200–201*, 1–9, https://doi.org/10.1016/j.pepi.2012.04.002.

Fichtner, A. and Simutė, S. (2018), Hamiltonian Monte Carlo Inversion of Seismic Sources in Complex Media, *Journal of Geophysical Research: Solid Earth*, *123*(4), pp. 2984–2999, https://doi.org/10.1002/2017JB015249.

Fichtner, A., Igel, H., Bunge, H.-P., and Kennett, B. L. N. (2009), Simulation and Inversion of Seismic Wave Propagation on Continental Scales Based on a Spectral-Element Method, 12.

Ford, S., Dreger, D., and Walter, W. (2010a), Source Analysis of the Memorial Day Explosion, Kimchaek, North Korea, *Geophysical Research Letters*, *12*, p. 15644, https://doi.org/10.1029/2009GL040003.

Ford, S., Dreger, D., and Walter, W. (2010b), Network Sensitivity Solutions for Regional Moment-Tensor Inversions, *Bulletin of the Seismological Society of America*, *100*(5A), pp. 1962–1970, https://doi.org/10.1785/0120090140.

Ford, S., Walter, W., and Dreger, D. (2012), Event Discrimination using Regional Moment Tensors with Teleseismic-P Constraints, *Bulletin of the Seismological Society of America*, *102*(2), pp. 867–872, https://doi.org/10.1785/0120110227.

Ford, S. R., Dreger, D. S., and Walter, W. R. (2009), Source analysis of the Memorial Day explosion, Kimchaek, North Korea, *Geophysical Research Letters*, *36*(21), https://doi.org/10.1029/2009GL040003.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013), emcee: The MCMC Hammer, *Publications of the Astronomical Society of the Pacific*, *125*(925), pp. 306–312, https://doi.org/10.1086/670067.

Gerstenberger, M. C., Marzocchi, W., Allen, T., Pagani, M., Adams, J., Danciu, L., et al. (2020), Probabilistic Seismic Hazard Analysis at Regional and National Scales: State of the Art and Future Challenges, *Reviews of Geophysics*, *58*(2), e2019RG000653, https://doi.org/10.1029/2019RG000653.

Gokhberg, A. and Fichtner, A. (2016), Full-waveform inversion on heterogeneous HPC systems. *Computers & Geosciences*, *89*, pp. 260–268, https://doi.org/10.1016/j.cageo.2015.12.013.

Goodman, J. and Weare, J. (2010), Ensemble samplers with affine invariance, *Communications in Applied Mathematics and Computational Science*, *5*(1), pp. 65–80, https://doi.org/10.2140/camcos.2010.5.65.

Gouveia, W. P. and Scales, J. A. (1998), Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis, *Journal of Geophysical Research: Solid Earth*, *103*(B2), pp. 2759–2779, https://doi.org/10.1029/97JB02933.

Hallo, M. and Gallovič, F. (2016), Fast and cheap approximation of Green function uncertainty for waveform-based earthquake source inversions, *Geophysical Journal International*, *207*(2), pp. 1012–1029, https://doi.org/10.1093/gji/ggw320.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*(1), pp. 97–109, https://doi.org/10.1093/biomet/57.1.97.

Hejrani, B. and Tkalčić, H. (2018), The 20 May 2016 Petermann Ranges earthquake: centroid location, magnitude and focal mechanism from full waveform modelling, *Australian Journal of Earth Sciences*, *0*(0), pp. 1–9, https://doi.org/10.1080/08120099.2018.1525783.

Hejrani, B. and Tkalčić, H. (2020), Resolvability of the Centroid-Moment-Tensors for Shallow Seismic Sources and Improvements From Modeling High-Frequency Waveforms, *Journal of Geophysical Research: Solid Earth*, *125*(7), e2020JB019643, https://doi.org/10.1029/2020JB019643.

Hejrani, B., Tkalčić, H., and Fichtner, A. (2017), Centroid moment tensor catalogue using a 3-D continental scale Earth model: Application to earthquakes in Papua New Guinea and the Solomon Islands, *Journal of Geophysical Research: Solid Earth*, *122*(7), pp. 5517–5543, https://doi.org/10.1002/2017JB014230.

Herrmann, R. B. (2013), Computer Programs in Seismology: An Evolving Tool for Instruction and Research, *Seismological Research Letters*, *84*(6), pp. 1081–1088, https://doi.org/10.1785/0220110096.

REFERENCES (Continued)

Hingee, M., Tkalčić, H., Fichtner, A., and Sambridge, M. (2011), Seismic moment tensor inversion using a 3-D structural model: applications for the Australian region: Moment tensor inversion using 3-D earth, *Geophysical Journal International*, *184*(2), pp. 949–964, https://doi.org/10.1111/j.1365-246X.2010.04897.x.

Jost, M. L. and Herrmann, R. B. (1989), A Student's Guide to and Review of Moment Tensors, *Seismological Research Letters*, *60*(2), pp. 37–57, https://doi.org/10.1785/gssrl.60.2.37.

Kanamori, H. (1977), The energy release in great earthquakes, *Journal of Geophysical Research (1896-1977)*, *82*(20), pp. 2981–2987, https://doi.org/10.1029/JB082i020p02981.

Käufl, P., Valentine, A. P., and Trampert, J. (2016), Probabilistic point source inversion of strong-motion data in 3-D media using pattern recognition: A case study for the 2008 Mw 5.4 Chino Hills earthquake, *Geophysical Research Letters*, *43*(16), pp. 8492–8498, https://doi.org/10.1002/2016GL069887.

Kawakatsu, H. (1996), Observability of the isotropic component of a moment tensor, *Geophysical Journal International*, *126*(2), pp. 525–544, https://doi.org/10.1111/j.1365-246X.1996.tb05308.x.

Kennett, B. L. N. and Salmon, M. (2012), AuSREM: Australian Seismological Reference Model, *Australian Journal of Earth Sciences*, *59*(8), pp. 1091–1103, https://doi.org/10.1080/08120099.2012.736406.

Kennett, B. L. N., Engdahl, E. R., and Buland, R. (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophysical Journal International*, *122*(1), pp. 108–124, https://doi.org/10.1111/j.1365-246X.1995.tb03540.x.

Kennett, B. L. N., Fichtner, A., Fishwick, S., and Yoshizawa, K. (2013), Australian Seismological Reference Model (AuSREM): mantle component, *Geophysical Journal International*, *192*(2), pp. 871–887, https://doi.org/10.1093/gji/ggs065.

Kikuchi, M. and Kanamori, H. (1982), Inversion of complex body waves, *Bulletin of the Seismological Society of America*, *72*(2), pp. 491–506.

Kim, S., Dettmer, J., Rhie, J., and Tkalčić, H. (2016a), Highly efficient Bayesian joint inversion for receiver-based data and its application to lithospheric structure beneath the southern Korean Peninsula, *Geophysical Journal International*, *206*(1), pp. 328–344, https://doi.org/10.1093/gji/ggw149.

Kim, S., Tkalčić, H., Rhie, J., and Chen, Y. (2016b), Intraplate volcanism controlled by back-arc and continental structures in NE Asia inferred from transdimensional Bayesian ambient noise tomography, *Geophysical Research Letters*, *43*(16), pp. 8390–8398, https://doi.org/10.1002/2016GL069483.

Komatitsch, D. and Tromp, J. (1999), Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophysical Journal International*, *139*(3), pp. 806–822, https://doi.org/10.1046/j.1365-246x.1999.00967.x.

Komatitsch, D. and Tromp, J. (2002a), Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophysical Journal International*, *149*(2), pp. 390–412, https://doi.org/10.1046/j.1365-246X.2002.01653.x.

Komatitsch, D. and Tromp, J. (2002b), Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophysical Journal International*, *150*(1), pp. 303–318, https://doi.org/10.1046/j.1365-246X.2002.01716.x.

Komatitsch, D., Ritsema, J., and Tromp, J. (2002), The Spectral-Element Method, Beowulf Computing, and Global Seismology, *Science*, *298*(5599), pp. 1737–1742, https://doi.org/10.1126/science.1076024.

Liu, J., Li, L., Zahradník, J., Sokos, E., and Plicka, V. (2018), Generalized Source Model of the North Korea Tests 2009–2017, *Seismological Research Letters*, *89*(6), pp. 2166–2173, https://doi.org/10.1785/0220180106.

Lloyd, A. J., Wiens, D. A., Zhu, H., Tromp, J., Nyblade, A. A., Aster, R. C., et al. (2020), Seismic Structure of the Antarctic Upper Mantle Imaged with Adjoint Tomography, *Journal of Geophysical Research: Solid Earth*, *125*(3), https://doi.org/10.1029/2019JB017823.

Malinverno, A. and Briggs, V. A. (2004), Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes, *GEOPHYSICS*, *69*(4), pp. 1005–1016, https://doi.org/10.1190/1.1778243.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953), Equation of state calculations by fast computing machines, https://doi.org/10.1063/1.1699114.

REFERENCES (Continued)

Minson, S. E., and Dreger, D. S. (2008), Stable inversions for complete moment tensors, *Geophysical Journal International*, *174*(2), pp. 585–592, https://doi.org/10.1111/j.1365-246X.2008.03797.x.

Mustać, M. and Tkalčić, H. (2016), Point source moment tensor inversion through a Bayesian hierarchical model, *Geophysical Journal International*, *204*(1), pp. 311–323, https://doi.org/10.1093/gji/ggv458.

Mustać, M. and Tkalčić, H. (2017), On the Use of Data Noise as a Site-Specific Weight Parameter in a Hierarchical Bayesian Moment Tensor Inversion: The Case Study of The Geysers and Long Valley Caldera Earthquakes, *Bulletin of the Seismological Society of America,* https://doi.org/10.1785/0120160379.

Mustać, M., Tkalčić, H., and Burky, A. L. (2018), The Variability and Interpretation of Earthquake Source Mechanisms in The Geysers Geothermal Field From a Bayesian Standpoint Based on the Choice of a Noise Model, *Journal of Geophysical Research: Solid Earth*, *123*(1), pp. 513–532, https://doi.org/10.1002/2017JB014897.

Mustać, M., Hejrani, B., Tkalčić, H., Kim, S., Lee, S.-J., and Cho, C.-S. (2020), Large Isotropic Component in the Source Mechanism of the 2013 Democratic People's Republic of Korea Nuclear Test Revealed via a Hierarchical Bayesian Inversion, *Bulletin of the Seismological Society of America*, *110*, pp. 166–177, https://doi.org/10.1785/0120190062.

Myers, S. C., Ford, S. R., Mellors, R. J., Baker, S., and Ichinose, G. (2018), Absolute Locations of the North Korean Nuclear Tests Based on Differential Seismic Arrival Times and InSAR, *Seismological Research Letters*, *89*(6), pp. 2049–2058, https://doi.org/10.1785/0220180123.

Nettles, M. and Ekström, G. (1998), Faulting mechanism of anomalous earthquakes near Bárdarbunga Volcano, Iceland, *Journal of Geophysical Research: Solid Earth*, *103*(B8), pp. 17973–17983, https://doi.org/10.1029/98JB01392.

Nguyen, B. (1994), Surface-wave study of the 22 January 1992 western Yellow Sea earthquake, Eos Trans, AGU, 75, Fall Meet. Suppl., Abstract S51D-12.

Panning, M., Dreger, D., and Tkalčić, H. (2001), Near-source velocity structure and isotropic moment tensors: A case study of the Long Valley Caldera, *Geophysical Research Letters*, *28*(9), pp. 1815–1818, https://doi.org/10.1029/2000GL012389.

Pasyanos, M. E., Dreger, D. S., & Romanowicz, B. (1996). Toward real-time estimation of regional moment tensors. *Bulletin of the Seismological Society of America*, *86*(5), 1255–1269.

Phạm, T.-S. and Tkalčić, H. (2021), Toward Improving Point-Source Moment-Tensor Inference by Incorporating 1D Earth Model's Uncertainty: Implications for the Long Valley Caldera Earthquakes, *Journal of Geophysical Research: Solid Earth*, *126*(11), e2021JB022477, https://doi.org/10.1029/2021JB022477.

Piana Agostinetti, N. and Malinverno, A. (2010), Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, *181*(2), pp. 858–872, https://doi.org/10.1111/j.1365-246X.2010.04530.x.

Randall, G. E. (1989), Efficient calculation of differential seismograms for lithospheric receiver functions, *Geophysical Journal International*, *99*(3), pp. 469–481, https://doi.org/10.1111/j.1365-246X.1989.tb02033.x.

Salmon, M., Kennett, B. L. N., and Saygin, E. (2013), Australian Seismological Reference Model (AuSREM): crustal component, *Geophysical Journal International*, *192*(1), pp. 190–206, https://doi.org/10.1093/gji/ggs004.

Sambridge, M. (1999), Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble, *Geophysical Journal International*, *138*(3), pp. 727–746, https://doi.org/10.1046/j.1365-246x.1999.00900.x.

Sambridge, M. and Mosegaard, K. (2002), Monte Carlo Methods in Geophysical Inverse Problems, *Reviews of Geophysics*, *40*(3), 3-1-3–29, https://doi.org/10.1029/2000RG000089.

Shang, X. and Tkalčić, H. (2020), Point-Source Inversion of Small and Moderate Earthquakes From P-wave Polarities and P/S Amplitude Ratios Within a Hierarchical Bayesian Framework: Implications for the Geysers Earthquakes, *Journal of Geophysical Research: Solid Earth*, *125*(2), https://doi.org/10.1029/2019JB018492.

REFERENCES (Continued)

Stähler, S. C., and Sigloch, K. (2014), Fully probabilistic seismic source inversion - Part 1: Efficient parameterisation, *Solid Earth*, *5*(2), pp. 1055–1069, https://doi.org/10.5194/se-5-1055-2014.

Takei, Y. and Kumazawa, M. (1994), Why have the single force and torque been excluded from seismic source models? *Geophysical Journal International*, *118*(1), pp. 20–30, https://doi.org/10.1111/j.1365-246X.1994.tb04672.x.

Tape, W. and Tape, C. (2012), A geometric setting for moment tensors, *Geophysical Journal International*, *190*(1), pp. 476–498, https://doi.org/10.1111/j.1365-246X.2012.05491.x.

Tape, W. and Tape, C. (2015), A uniform parametrization of moment tensors, *Geophysical Journal International*, *202*(3), pp. 2074–2081, https://doi.org/10.1093/gji/ggv262.

Tarantola, A. and Valette, B. (1982), Inverse problems = Quest for information, *Journal of Geophysics*, *50*(1), pp. 159–170.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation,* Society for Industrial and Applied Mathematics, https://doi.org/10.1137/1.9780898717921.

Tashiro, Y. (1977), On methods for generating uniform random points on the surface of a sphere, *Annals of the Institute of Statistical Mathematics*, *29*(1), pp. 295–300, https://doi.org/10.1007/BF02532791.

Tkalčić, H., Dreger, D. S., Foulger G. R., and Julian B. R. (2009), The Puzzle of the 1996 Bardarbunga, Iceland, Earthquake: No Volumetric Component in the Source Mechanism, *Bulletin of the Seismological Society of America*, 99(5), pp. 3077–85, https://doi.org/10.1785/0120080361.

Tkalčić, H. and Phạm, T.-S. (2020), Excitation of the global correlation wavefield by large earthquakes, *Geophysical Journal International*, *223*(3), pp. 1769–1779, https://doi.org/10.1093/gji/ggaa369.

Vackář, J., Burjánek, J., Gallovič, F., Zahradník, J., and Clinton, J. (2017), Bayesian ISOLA: new tool for automated centroid moment tensor inversion, *Geophysical Journal International*, *210*(2), pp. 693–705, https://doi.org/10.1093/gji/ggx158.

Vasyura-Bathke, H., Dettmer, J., Dutta, R., Mai, P. M., and Jónsson, S. (2021), Accounting for theory errors with empirical Bayesian noise models in nonlinear centroid moment tensor estimation, *Geophysical Journal International*, *225*(2), pp. 1412–1431, https://doi.org/10.1093/gji/ggab034.

Vavryčuk, V. (2001), Inversion for parameters of tensile earthquakes, *Journal of Geophysical Research: Solid Earth*, *106*(B8), pp. 16339–16355, https://doi.org/10.1029/2001JB000372.

Wang, T., Shi, Q., Nikkhoo, M., Wei, S., Barbot, S., Dreger, D., et al. (2018), The rise, collapse, and compaction of Mt. Mantap from the 3 September 2017 North Korean nuclear test, *Science*, eaar7230, https://doi.org/10.1126/science.aar7230

Weston, J., Engdahl, E. R., Harris, J., Di Giacomo, D., and Storchak, D. A. (2018), ISC-EHB: reconstruction of a robust earthquake data set, *Geophysical Journal International*, *214*(1), pp. 474–484, https://doi.org/10.1093/gji/ggy155.

Xu, H., Ni, S., Liu, W., Zhu, H., and Wang, X. (2020), Focal mechanisms of the 2017 North Korean nuclear test and its early collapse event, *Geophysical Journal International*, *220*(2), pp. 737–752, https://doi.org/10.1093/gji/ggz462.

Yagi, Y. and Fukahata, Y. (2011), Introduction of uncertainty of Green's function into waveform inversion for seismic source processes: Uncertainty of Green's function in inversion, *Geophysical Journal International*, *186*(2), pp. 711–720, https://doi.org/10.1111/j.1365-246X.2011.05043.x.

Yang, G., Zhao, L. F., Xie, X. B., Zhang, L., and Yao, Z. X. (2021), High-Precision Relocation With the Burial Depths of the North Korean Underground Nuclear Explosions by Combining Pn and Pg Differential Traveltimes, *Journal of Geophysical Research: Solid Earth*, *126*(6), e2020JB020745, https://doi.org/10.1029/2020JB020745.

LIST OF SYMBOLS, ABBREVIATIONS, and ACRONYMS

CLVD	Compensated Linear Vector Dipole
DC	Double Couple
DPRK	Democratic People's Republic of Korea
ISO	Isotropic
McMC	Multi-chain Monte Carlo
MT	Moment Tensor
SNR	Signal to Noise Ratio

DISTRIBUTION LIST

DTIC/OCP
8725 John J. Kingman Rd, Suite 0944
Ft Belvoir, VA 22060-6218 1 cy

AFRL/RVIL
Kirtland AFB, NM 87117-5776 1 cy

Official Record Copy
AFRL/RVB/Dr. Megan P. Flanagan 1 cy