DISCLAIMER

SAND2025-01059

SANDIA REPORT
SAND20XX-XXXX
Printed Click to enter a date

Sandia National Laboratories

An Appraisal of the Performance and Characteristics of Summary Rays Calculated for the SALSA3D Traveltime Dataset

Anant Hariharan, Robert W. Porritt, Andrea C. Conley

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

The SALSA3D tomographic model provides a crucial community resource for improving the quality (in terms of both accuracy and precision) of predictions of the traveltimes of seismic waves and therefore improving our ability to locate anthropogenic or natural seismic events. Constructing the requisite tomographic model requires addressing the challenges implied by a massive and growing dataset of traveltime measurements. This study explores one approach to tackle this challenge: the use of summary rays, which average traveltime measurements from sources within evenly spaced cells, thereby eliminating redundant data.

We find that the application of summary rays can reduce the size of a full dataset of measurements with the GT50 designation by more than a factor of two, with a larger impact on Pn measurements than P measurements. As they are an average of an ensemble of observations, summary rays also measurably reduce noise in the individual measurements and are more consistent with predictions from the current generation of the SALSA3D model than the individual measurements. The summary ray dataset also promotes better data fit when inverted for a tomographic model than the data fit observed for a corresponding dataset with no summary rays.

The tomographic model inverted for using the summary rays has higher resolution (and lower model uncertainty) along subduction zones in the southern hemisphere, and therefore presents an opportunity to improve model resolution and traveltime predictions in these regions. Overall, the proof-of-concept tomographic inversions show that the summary rays can be successfully incorporated into inversions and generate results comparable to models solved for using standard datasets, while also generating a massive reduction in size for the datasets that they are applied to.

ACKNOWLEDGEMENTS

This Ground-based Nuclear Detonation Detection (GNDD) research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D).

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

CONTENTS

LIST OF FIGURES

This page left blank

ACRONYMS AND TERMS

Acronym/Term	Definition
SALSA3D	Sandia and Los Alamos 3D Seismic Tomography Model
GT	Ground Truth (km)
IMS	International Monitoring System
GSN	Global Seismic Network
P, P-wave	Compressional/Primary wave traveling through the mantle
Pn	Compressional wave traveling along the crust-mantle boundary
S, S-wave	Shear/secondary wave traveling through the mantle
1D, 2D, 3D	One, two, or three dimensional, respectively
RSTT	Regional Seismic Travel Times 2.5D velocity model
RMS	Root Mean Squared error
dVp	Perturbation to background P-wave isotropic velocity

This page left blank

1. INTRODUCTION

The development of accurate locations of seismic events is a fundamental aspect of the characterization and attribution of observed seismic sources. Doing so requires accurate predictions of the traveltimes used in the location problem (Downey et al., 2019) and thus necessitates a 3D Earth model designed for this purpose. Accurate uncertainty quantification of these timing predictions, while not yet standard, is essential in order to implement meaningful observation weighting in the location problem, which can in turn exert control on both the absolute value of these locations and their uncertainties. SALSA3D (Sandia and Los Alamos 3D seismic tomography model) is explicitly designed for this endeavor, improving both the accuracy of predictions of body-wave traveltimes as well as estimates of the uncertainty of these predictions. It does so by solving the inverse problem for slowness in the Earth's mantle using a formalism couched in a Bayesian approach that incorporates prior knowledge about model uncertainty as a function of depth (Ballard et al., 2016) and producing a full model covariance matrix that can be used for uncertainty estimates in traveltime predictions. The current version of SALSA3D has already been shown to improve both predictions of traveltimes relative to a state-of-the art 1D seismic model and reduce event mislocation relative to both a 1D seismic model and the RSTT model (Begnaud et al., 2021). SALSA3D continues to iterate in active development on a range of fronts (e.g. Rowe et al., 2021; Conley, 2021) to improve on its performance.

The data used in inversions for SALSA3D are measurements of body wave traveltimes. With the deployment of hundreds of regional and local temporary seismic networks, as well as the continued operation of long-term permanent seismic networks such as the GSN and IMS (Ringler et al., 2022; National Research Council, 2012), the amount of waveform data and traveltime measurements available for use has increased exponentially. Datasets of such vast size result in large matrices and data objects that present a computational challenge, particularly within an inversion framework that incorporates information from the full data covariance matrix. Furthermore, the inhomogeneous nature of the distribution of seismic sources (Ekström, 2012) and receivers means that much of these vast datasets is redundant and samples similar regions of the Earth's interior. To effectively use a full dataset of all available traveltimes within the SALSA3D inversion framework, it has now become necessary to explore approaches to truncate the traveltime dataset and eliminate measurement redundancy. In the case of the SALSA3D dataset, a large subset of measurements are classified as GT50, representing an event with an assigned epicenter within at least 50 km of the true epicenter. This designation is made based on network configuration (Bondár et al., 2004). The relatively high location uncertainties renders such events relatively unreliable compared to the rest of the dataset. Our analysis therefore focuses specifically on calculating summary rays for this subset of measurements due to their relative abundance and relatively high inherent uncertainty.

Global inversions for seismic wavespeed have had to contend with the challenges posed by massive datasets for decades and have implemented different approaches to do so. For instance, Ruan et al., (2012) proposed a generalized geographic weighting approach to eliminate the effects of uneven data coverage by parameterizing the Earth's surface using Voronoi cells and adjusting the data weighting for each measurement based on the area of the corresponding cell. Li & Romanowicz (1996) instead accounted for redundant data by explicitly altering the data covariance matrix to downweight data that had similar errors, often spanning similar raypaths

and representing nearby events. However, the most popular and long-standing method to address redundant and uneven data coverage is the use of summary rays. Using this approach, nearby events are clustered and the associated traveltimes are averaged to construct a summary ray datum, eliminating the impact of having many source-station pairs with sources from a nearly identical location. Composite rays were first implemented in the 1980s (Spakman & Nolet, 1998) in the context of body wave tomography, and served as a precursor to summary rays by implementing the idea of averaging data from certain event clusters. Later studies implemented summary rays by clustering the events within grids on the Earth's surface (e.g., Morelli & Dziewonski, 1987; Bijwaard et al., 1998; Bolton & Masters, 2001) and this approach became a workhorse of tomographic inversions in the early 2000s. Summary rays gradually accumulated alternate uses. For instance, a family of studies also used summary rays to directly infer the spectra of heterogeneity in the Earth's mantle for both P and S waves (Gudmundsson et al., 1989; Davies et al., 1992) by examining the variance in summary rays as a function of increasing ray bundle size and fitting this to predictions of traveltime scattering calculated by assuming a specific heterogeneity spectrum for the Earth's mantle.. The stable and accurate data produced by summary rays also enabled workers to identify and eliminate systematically biased data due to clock errors in the seismic dataloggers (Bolton & Masters, 2001).

With the advent of dense regional networks such as USArray, the distribution of traveltime data on the Earth's surface has dramatically changed. However, there have not been studies in recent years that assess the impact of using summary rays on contemporary datasets of body-wave traveltimes. This study addresses this challenge with particular attention to the representative and up-to-date dataset used for construction of the SALSA3D model. We explore the size, variance, and accuracy of the resulting dataset of summary rays, particularly in relation to the original dataset. Using tomographic inversions, we also explore strategies to incorporate summary rays into the SALSA3D model and analyze the impact of doing so on the resulting models. The resulting work serves as a guide for using summary ray data and can inform scientists' expectations on the benefits that can be gained from doing so.

2. DATA AND METHODS

2.1. Description of the Underlying Dataset

The dataset used in this study is an augmented version of the dataset from Ballard et al. (2016) and includes measurements from a range of temporary and permanent seismic networks that represent contributions from both regional and international data centers and experiments. Some of these sources are also described in Ballard et al. (2016), and include the permanent global IMS and GSN networks, as well as temporary networks such as the USArray Transportable Array, SKIPPY, and GLISN networks (Figure 1). All measurements used in the dataset are assigned GT (Ground Truth; Bondar et al., 2004) levels spanning discrete values between of 0 and 50. (Figure 2).

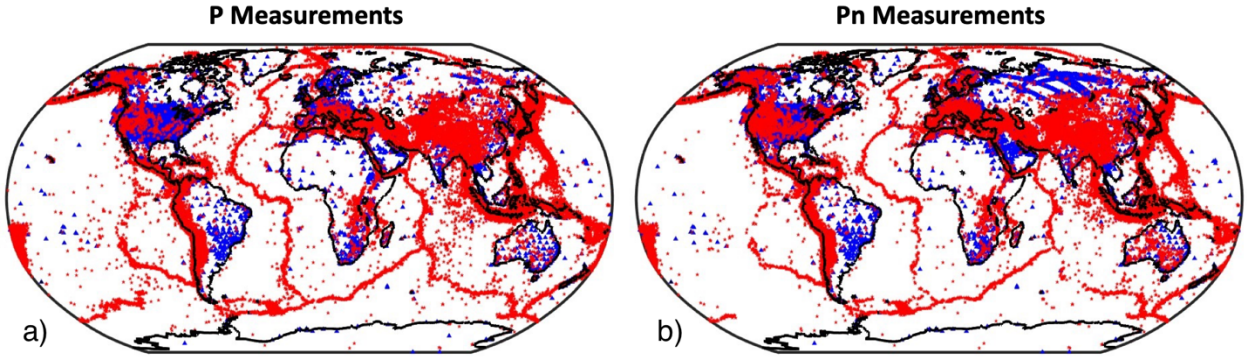

Figure 1: Locations of earthquakes (red) and stations (blue) that generate and record P (a) and Pn (b) measurements.

Both P and Pn measurements originate from events and stations across the globe, though with most stations placed in the northern hemisphere. The total number of Pn measurements is (9,803,707) and the number of P measurements is (23,074,412).

The majority of the measurements in our dataset are designated as GT=50 (Figure 2). For both P and Pn measurements, the number of these events is more than an order of magnitude larger than the other measurements in the dataset.

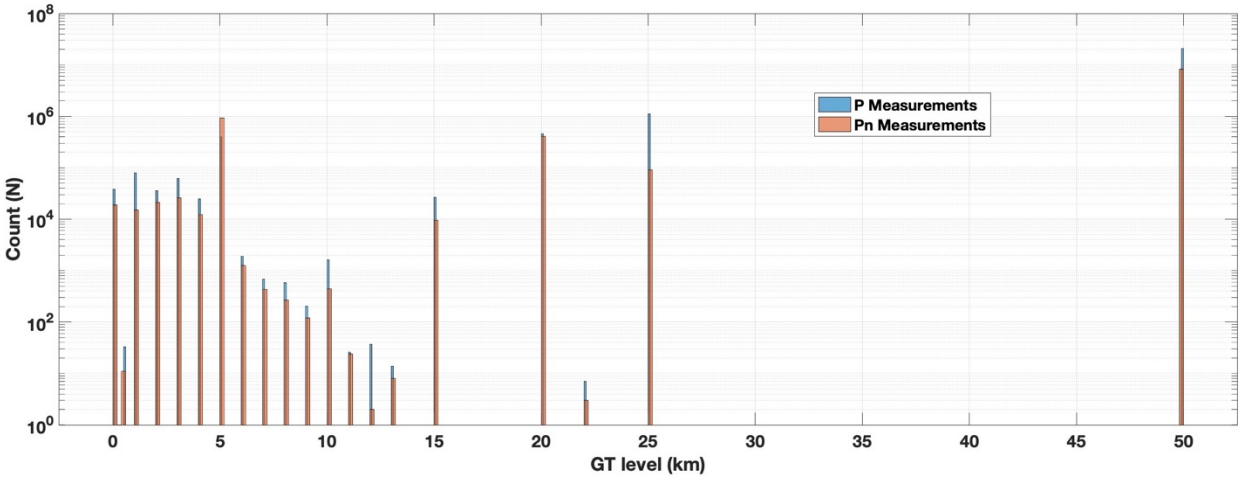

Figure 2: Histogram of the GT designation of measurements in the P (blue) and Pn (orange) datasets explore in this study. Note the logarithmic y-axis.

The spatial distribution of GT50 measurements is broadly similar to the distribution of GT25 measurements, but there are notable differences (Figure 3). GT50 data often represents the only contributing earthquakes in certain regions; for instance, the GT50 dataset provides the only events with recorded P waves located at much of the Cocos, Rivera, and North Andean subduction zones, as well as the only events located along New Zealand and the ridge dividing the Australian and Antarctic plates. Thus, they provide an important complement to the information from the GT < 50 dataset in these regions.

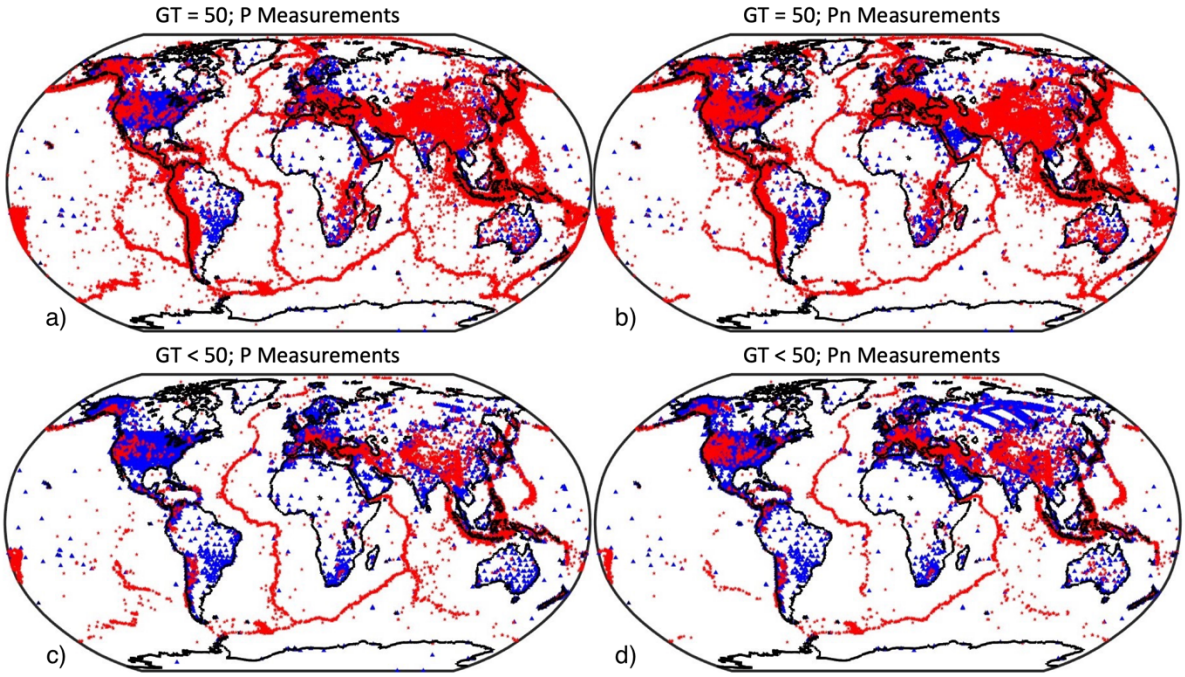

Figure 3: Locations of earthquakes (red) and stations (blue) that generate and record at least a single P or Pn measurement, shown separately for measurements with GT = 50 (a,b) and measurements with GT < 50 (c,d).

2.2. Calculation of Summary Rays

We calculate summary rays for every station by taking the mean of all GT50 traveltime measurements from earthquakes with hypocenters that fall within voxels defined by an evenly spaced grid of 2D triangular tessellations of a unit sphere with a depth extent of 15 km. A new summary ray is assigned an event location with the source depth calculated as the average of all the source depths from the contributing measurements. and the event hypocenter calculated as the average of all the contributing events' hypocenters.

In order to ensure an evenly spaced grid, the grid is generated using the GeoTess software package (Ballard et al., 2016). Such a choice improves over summary ray calculations in other tomographic studies which used a grid of constant latitude/longitude spacing and thus promoted smaller ray bundles at higher latitudes. We experiment with edge widths of 1,2,4, and 8 degrees (Figure 4). In this report, we focus on summary rays calculated using one edge width, and the following section presents a justification for this choice of edge width. The summary rays are calculated separately for P and Pn measurements.

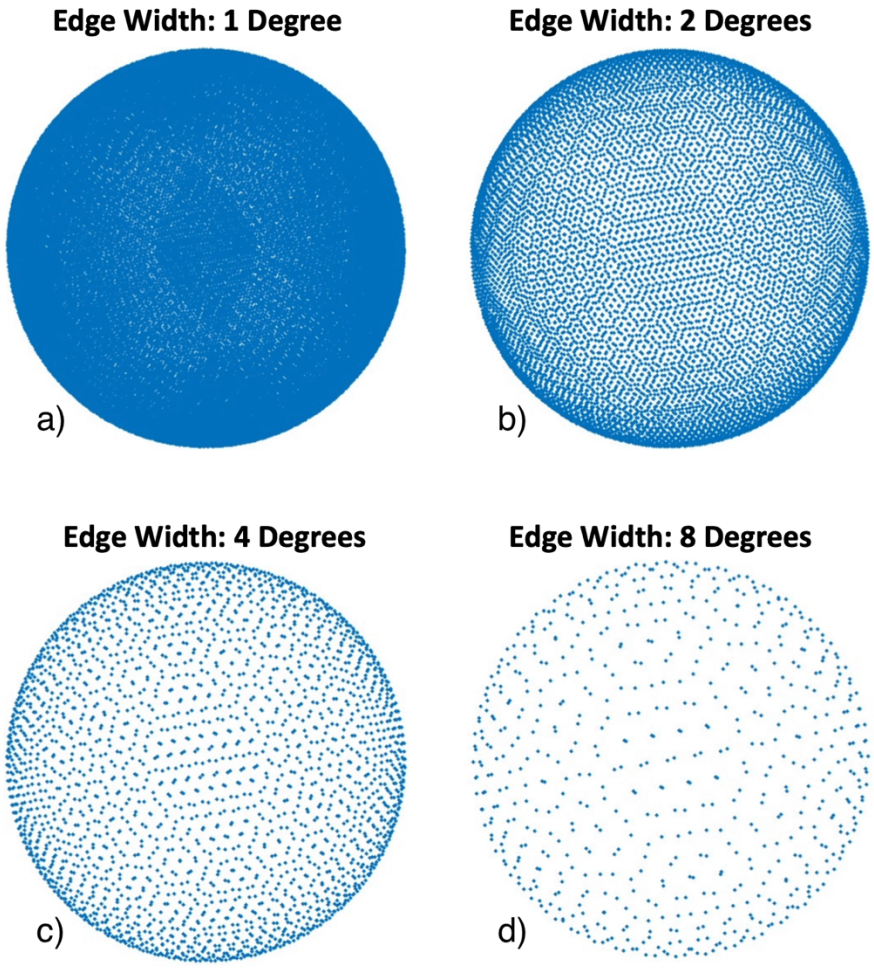

Figure 4: Figure showing the location of grid nodes calculated for a GeoTess grid with edge lengths spanning (a) 1,(b) 2, (c) 4, and (d) 8 degrees.

This page left blank

3. CHARACTERISTICS OF SUMMARY RAY CALCULATIONS ON THE SALSA3D DATASET

In this section, we describe the characteristics of the summary rays that are calculated on the SALSA3D dataset. First, we quantify the impact of the edge width of the underlying GeoTess grid that is used to calculate the summary rays on (1) the extent to which we can truncate the original datasets and (2) the variance in the summary rays. Using this, we justify our choice of edge width for the GeoTess grid. We then perform a more detailed examination of the variance and accuracy of the summary rays. Using synthetic predictions of noise-free traveltimes through a 3D Earth model, we perform a suite of synthetic tests (described in sections 3.3 and 3.4) that allow us to both examine the different sources of noise in the ray bundles and assess how accurate the summary ray calculations are when compared to the underlying dataset of GT50 measurements, and a representative sample of GT < 50 measurements. We also present an exploration of different methods to quantify uncertainty in summary rays.

3.1. Impact of Different Ray Bundle Size on Summary Ray Behavior

3.1.1. Impact of Ray Bundle Size on Dataset Truncation

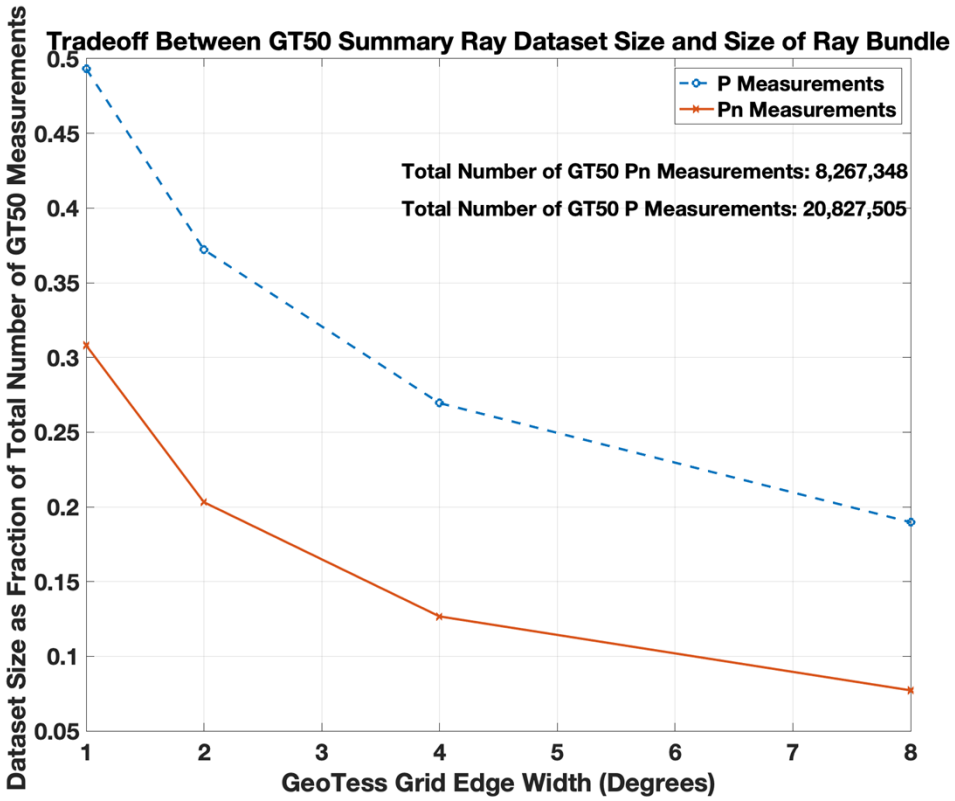

Figure 5: Impact of the use of Summary Rays on the size of the underlying GT50 dataset (as a fraction of the number of rays in original dataset), shown separately for P and Pn waves. The absolute number of rays in the original dataset is presented for reference in the top-right corner.

As expected, we find that the summary rays can substantially reduce the size of the original GT50 dataset (Figure 5). At the lowest grid size, the size of the GT50 P dataset is reduced by more than a factor of 2 (corresponding to a reduction of ~10 million measurements), and the size of the GT50 Pn dataset is reduced by more than a factor of 3. The fact that the summary rays impact the GT50 Pn dataset to a larger extent than the corresponding P-wave dataset suggests that the Pn epicenters tend to be less sparsely distributed and more clustered on average. This is consistent with the plot of Pn hypocenters shown in Figure 1, which visibly has fewer isolated seismic events than for P measurements.

3.1.2. Impact of Ray Bundle Size on Dataset Variance

The size of the voxel, within which we calculate summary rays, controls the numbers of events included in a summary ray, which may in turn impact the spread of the measurements captured within the summary rays and therefore also the precision and accuracy of the summary rays. To explore the impact of the edge width of the GeoTess grid used on this facet of measurement quality, we calculate the average standard deviation of all the measurements contributing to a summary ray and plot the dependence of the average value of these standard deviations as a function of for every grid size explored.

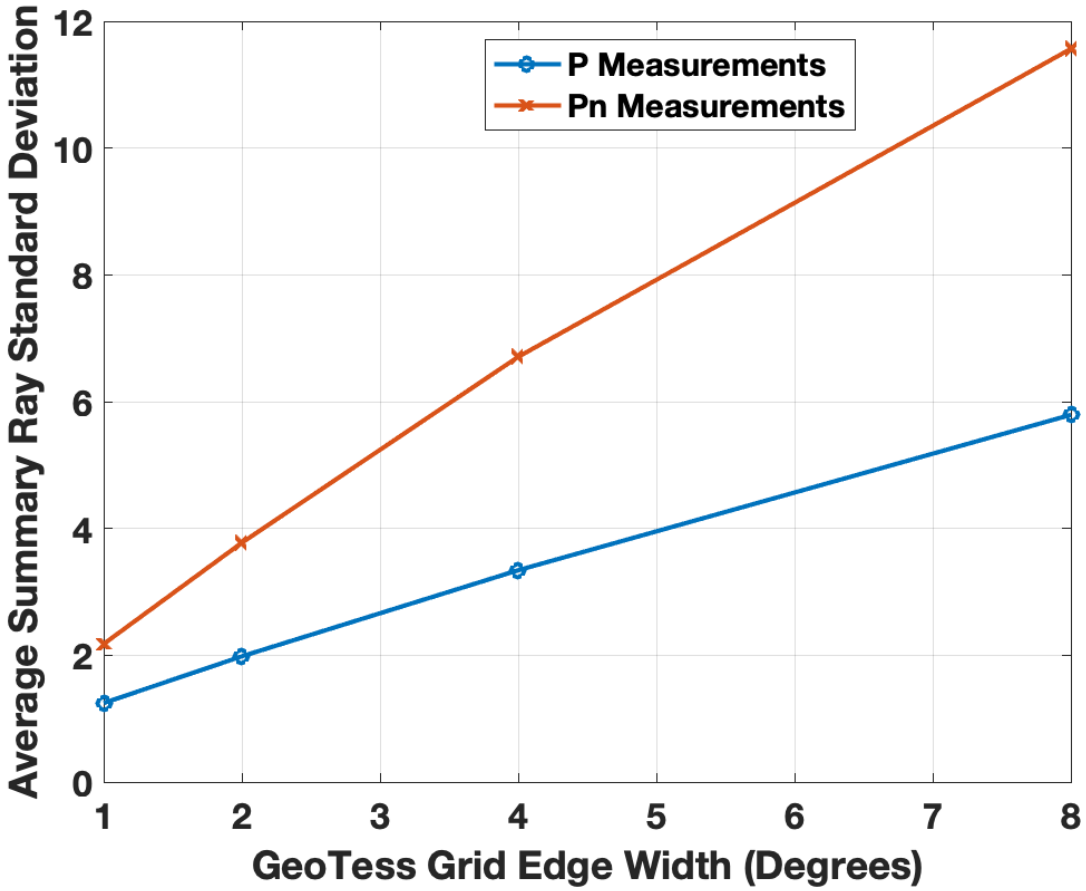

Figure 6: Average standard deviation of contributing measurements in a summary as a function of the GeoTess triangle grid edge width, for P and Pn measurements. Only GT50 measurements are used, and only summary rays with more than one ray in the ray bundle are considered.

The standard deviation plotted on the left-hand y-axis represents the combination of several factors, including random error due to measurement noise and the contribution of variations in individual raypaths that are averaged into the summary ray. The approximately linear increase in the average summary ray standard deviation with increasing GeoTess cell size reflects the fact that with sensitivity to a wider range of earthquake hypocenters, there is an increase in the range of mantle heterogeneity that a ray will experience, resulting in a wider spread in the range of traveltimes that a summary ray is sensitive to.

Overall, the reduced standard deviation with decreasing cell size provides a justification for us to use the narrowest ray bundles possible, corresponding to a GeoTess grid with edge width of one degree for our summary rays; Figure 5 shows that such a calculation still reduces the size of the GT50 dataset by more than a factor of two for the direct P measurements and more than a factor of 3 for the Pn measurements.

3.2. Approaches to Calculating Summary Ray Uncertainty

As they represent a statistical average of a set of observations, summary rays present a natural avenue to quantify measurement uncertainty; the uncertainty of a given summary ray can simply be taken as the standard deviation of all contributing measurements. It is necessary to develop a mechanism to quantify the uncertainty for summary rays as these data uncertainties are necessary in order to form the diagonal of the data covariance matrix in a tomographic inversion (e.g. Ballard et al., 2016). However, alternate approaches to determine these uncertainties, based on, for instance, the analyst-assigned uncertainties in the individual contributing rays, are also possible. In this section, we explore the distribution of measurement uncertainties and speculate on the factors that might impact them, causing the variance we observe in the measurements that contribute to a summary ray.

Two approaches to describe measurement uncertainty are to (1) take the standard deviation of the traveltimes going into a summary and to (2) take the average of the analyst-assigned measurement uncertainties for all the measurements contributing to a summary ray. In Figure 7, we compare the distribution of uncertainties obtained for both approaches for our dataset of P and Pn summary rays.

When summary ray uncertainties are defined using the standard deviation of contributing measurements, we obtain a smooth and approximately single-sided gaussian distribution of uncertainties; the average Pn uncertainty (2.18 s) is larger than the average P uncertainty (1.25 s) and the distribution (Figure 7) shows a much larger kurtosis (i.e., more measurements fall within the tail of the distribution).

On the other hand, if the uncertainties are calculated using the average analyst-defined uncertainties, the values exhibit discrete peaks at certain values due to discrete values of the underlying individual uncertainties; for instance, the majority of the individual P measurements are assigned an individual uncertainty of 1 s, resulting in a corresponding peak for the average analyst-assigned uncertainty. The average P uncertainty in this case is 1.06 s and the average Pn uncertainty is 1.38 s.
For both cases, the uncertainties for Pn measurement summary rays are larger on average than

the uncertainties for summary ray P measurements, but the variance captured in the summary ray standard deviation is subtly higher than that inferred from the analyst-assigned uncertainties.

To understand the origin of this discrepancy, we compare the average analyst-assigned uncertainty for a measurement directly to its corresponding summary ray uncertainty (Figure 8). For P measurements, the correlation coefficient between both metrics is 0.026 and for Pn measurements, the correlation coefficient between both metrics is -0.003, suggesting virtually no agreement between the approaches. Rather, both approaches to uncertainty quantification appear to represent fundamentally different characteristics of the measurements. In the next section, we explore why this is.

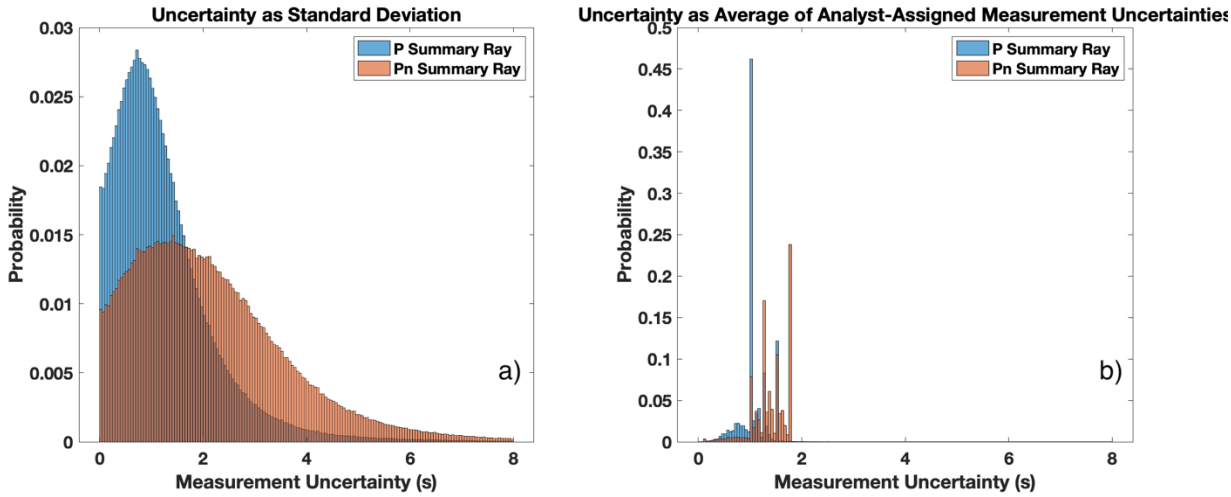

Figure 7: (a): Distribution of uncertainties, calculated as the standard deviation of all measurements in a ray bundle, shown separately for P and Pn measurements. (b): Distribution of analyst-assigned uncertainties for individual P and Pn measurements. Since there are far more P than Pn measurements, the histograms are shown as probability density functions normalized to 1 instead of with the y-axis corresponding to the absolute count in each bin.

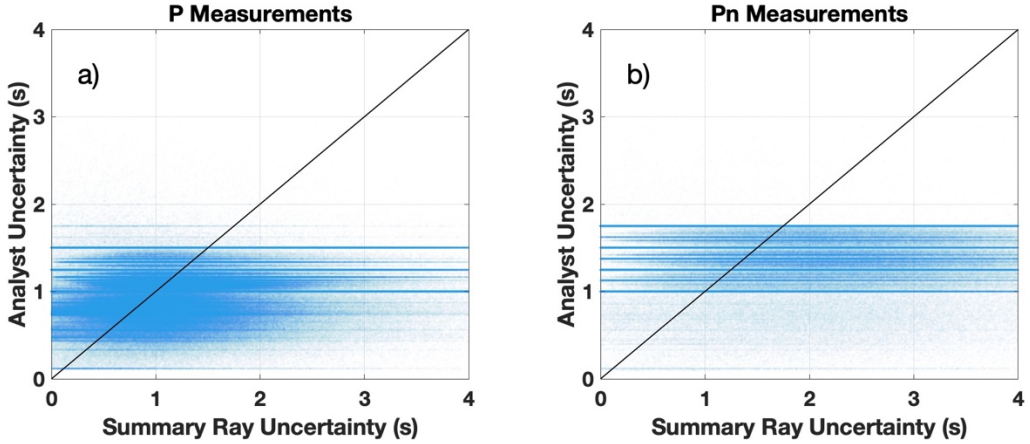

Figure 8: Comparison between summary ray uncertainties calculated as the standard deviation of the contributing traveltimes and uncertainties calculated as the average of the analyst-assigned

uncertainties. Left: Comparison for P measurements. Right: Comparison for Pn measurements. A 1-1 line is shown for comparison and is not a fit.

3.3. The Relative Strengths of Different Sources of Noise in Summary Rays

To understand which approach to uncertainty quantification to use for summary rays and to accurately quantify how much random noise is in the summary rays, we pose a question. Since the different measurements in a summary ray can travel along subtly different raypaths and each sample slightly different mantle heterogeneity, how much scatter is this contribution from path-dependent earth structure expected to introduce in a realistic dataset of summary rays? To answer this question, we perform a synthetic test.

By examining our dataset of measurements, we first choose the 26 stations with the largest numbers of P-wave measurements. For every measurement at these stations, we generate a synthetic traveltime prediction using the PCalc application (Ballard et al., 2009) and by ray-tracing paths through the published SALSA3D Earth model (Ballard et al., 2016). We thus have a dataset of synthetic traveltimes that matches the distribution of measurements at these well-sampled stations.
Using our new dataset of synthetic, noise-free traveltimes, we then calculate summary rays on these traveltimes and calculate the standard deviation for each of these summary rays in the same manner that we calculate it on the real data. The scatter in these synthetic summary rays can then be attributed to solely the contributions from mantle heterogeneity; the test therefore allows a realistic estimate of the scatter that we expect from mantle heterogeneity to the measurements in a summary ray. Overall, this test involves 2,104,859 individual GT50 measurements and 237,478 summary rays.

Figure 9 presents a comparison of the uncertainties for the corresponding summary rays generated from synthetic and real data. A one-to-one plot of every synthetic measurement against its real counterpart shows that the uncertainty in the real and synthetic data sets are well correlated, with a correlation coefficient of 0.77. This shows that the scatter of the measurements in a summary ray is well-described by the subtle variations in raypaths in a ray bundle that result from raypaths predicted using the SALSA3D model.

Furthermore, by examining the distributions of uncertainties for the real and synthetic datasets (Figure 9) we find that the average of the uncertainties in both sets of summary rays are similar, with the real dataset only having a slightly higher average standard deviation (median = 1.17 s) than the synthetic dataset of summary rays (median = 1.02 s). This shows that most of the 'scatter' in our measured summary rays are due to variations in realistic mantle heterogeneity that are captured by the summary ray, and that the contribution from data noise is also relatively small in comparison. The dominant role of mantle heterogeneity as a source of scatter in the summary rays is also shown in Figure 6; if scatter was only due to random noise, we would see no increase in error with increasing grid size. Instead, we observe a clear increase in standard deviation with increasing grid size. Overall, the ability of the synthetic noise to closely match the noise measurements on the real summary rays likely suggests that using the analyst-assessed uncertainties may not be appropriate for summary rays, and thus we proceed using the standard

deviation of the contributing measurements as our chosen metric for measurement uncertainty.

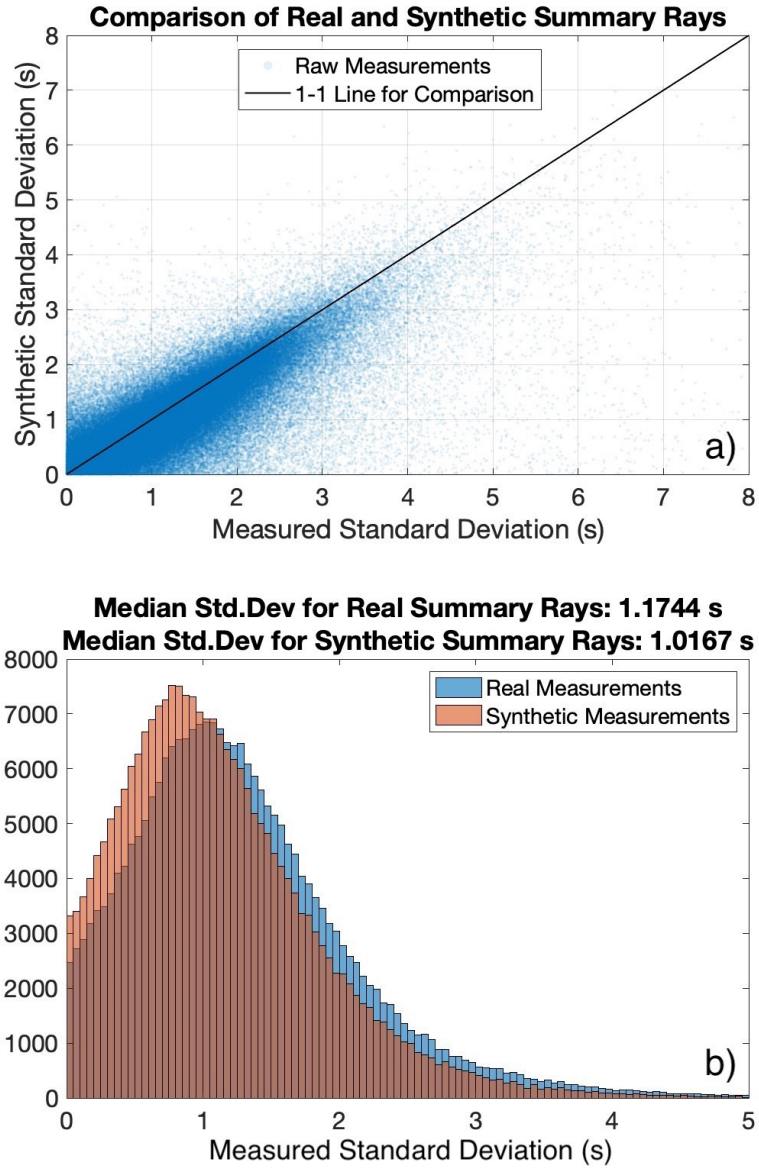

Figure 9: (a): 1-1 comparison between the uncertainty for a measured summary ray using real data, and the corresponding uncertainty calculated for the same ray but using synthetic data. (b): Histogram of the uncertainties for real and synthetic data.

3.4. Validating the Accuracy of Summary Rays using SALSA3D Traveltime Predictions

Before using the summary rays in a tomographic inversion, we wish to confirm that the measurements are sufficiently accurate, particularly in comparison to the individual measurements in the original dataset. To do this, we perform a synthetic test similar to that described in the previous section, this time with the goal of assessing the accuracy of our summary ray measurements. For this test, we limit our attention to P-wave measurements, which represent the bulk of the dataset.

For every summary ray (only using summary rays with multiple GT50 measurements) calculated at the 26 most prolific stations in the SALSA3D dataset, we predict a corresponding traveltime by ray-tracing through the SALSA3D model using PCalc and assuming the same source location as the one calculated for the summary ray. In addition, we randomly sample one of the contributing GT50 measurements to that summary ray and generate a corresponding traveltime prediction for the same GT50 measurement, using its source location. By randomly sampling a single GT50 measurement for every summary ray, we ensure that the comparison between the individual and summary ray measurements has approximately the same source/station geometry, distribution of azimuth and epicentral distances, and samples the same heterogeneity inside the Earth, controlling for the confounding effects that such factors might otherwise introduce in a comparison between two sets of measurements. For reference, we also generate traveltime predictions for every GT < 50 measurement at the stations used in this test. We note here that comparing the GT50 observations here with predictions of the SALSA3D model from Ballard et al., 2016 is not a circular comparison since the Ballard et al., 2016 model did not use any measurements with GT = 50.

We then calculate the summary ray errors as the absolute difference between the SALSA3D predicted traveltime for the summary rays and the measured traveltime for the summary rays. Similarly, we calculate the individual ray errors as the absolute difference between the individual GT50 measured traveltimes and their corresponding SALSA3D predicted traveltimes. In Figure 10, we compare these quantities, and find that the summary rays (which have a median difference of 0.47 s relative to the predictions) are more consistent with the predictions than the individual GT50 measurements (which have a median difference of 0.572 s relative to the predictions). We also calculate the corresponding errors for the GT < 50 measurements and compare these to the summary ray measurement errors in Figure 11, finding that once again, the summary rays are on average more consistent with the traveltime predictions than the GT < 50 measurements (which have a median error of 0.570 s, slightly better than the individual GT50 measurements). We note here, however, the caveat that a comparison between the GT<50 measurements and the summary rays is not as much of an apples-to-apples comparison as with the GT50 measurements because we can no longer guarantee that the measurements traverse along approximately identical paths. There are also fewer GT<50 measurements, and the distributions shown in Figure 11 are consequently normalized.

Overall, this test suggests that, under the valid assumption of the accuracy of the SALSA3D model, the fact that the GT50 summary rays are an aggregate of individual measurements allows

them to average out random noise and, on average, be noticeably more accurate than the individual measurements used to construct them.

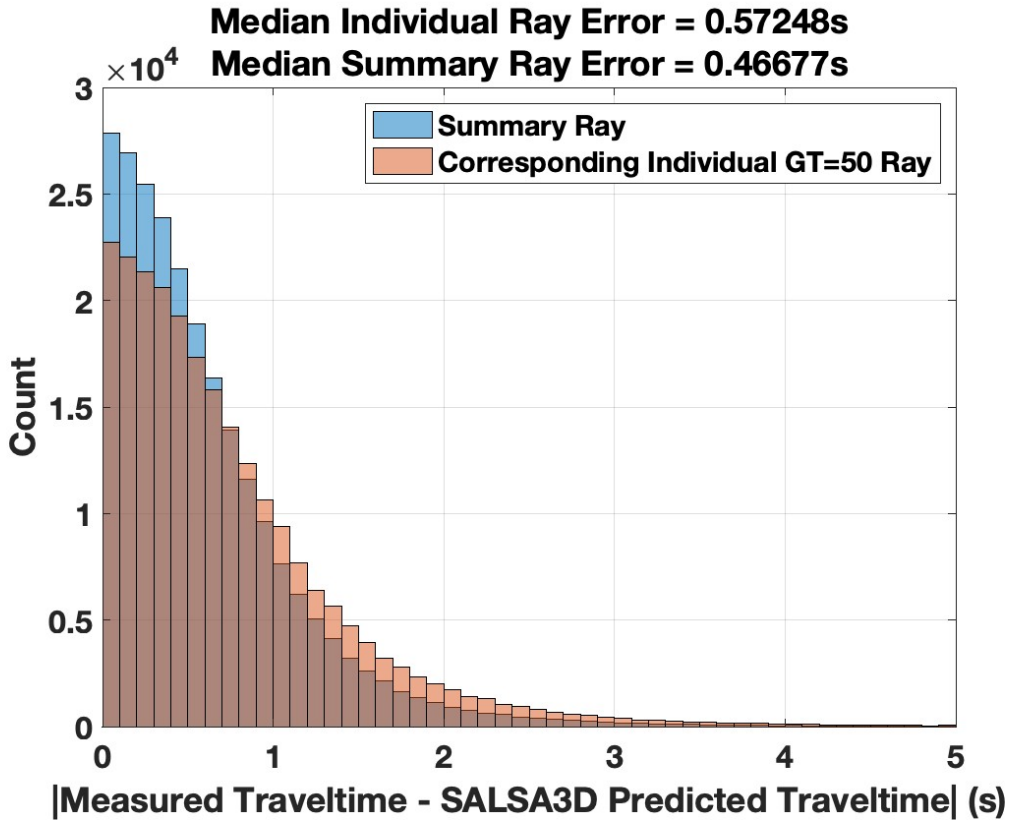

Figure 10: Comparison of the distribution of (blue) the difference between measured traveltimes for a summary ray and predicted traveltimes for the same summary ray, and (orange) the difference between measured traveltimes for a single GT50 measurement in the corresponding summary ray and predicted traveltimes for the same GT50 measurement.

We also explore other controls on the accuracy of a summary ray relative to the predictions. We find evidence to suggest that the number of individual measurements in a summary ray can have a significant impact. Figure 12 shows that, when the difference between a summary ray and its prediction is plotted as a function of the number of measurements, this error decreases from approximately 0.7 s on average to 0.5 s as the number of measurements increases, converging after the summary ray includes approximately 6 measurements. This conclusion suggests an alternate approach for data weighting that promotes more accurate data, where weighting is based on the number of measurements in a summary ray and not their standard deviation. Such a scheme is explored later in this report.

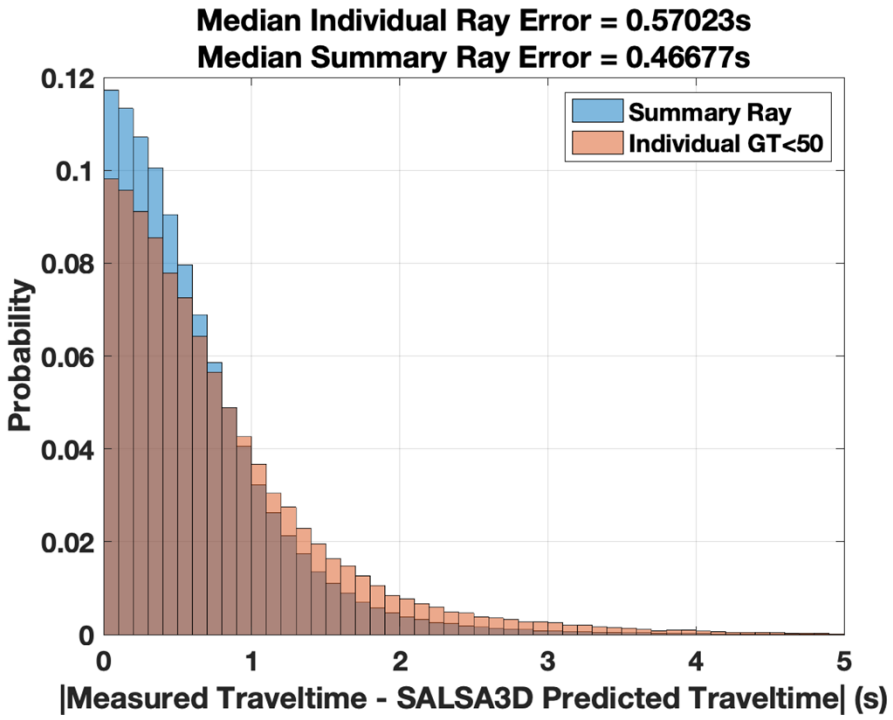

Figure 11: Comparison of the distribution of (1) the difference between measured traveltimes for a summary ray and predicted traveltimes for the same summary ray, and (2) the difference between measured traveltimes for a single GT<50 measurement and a prediction for the same measurement.

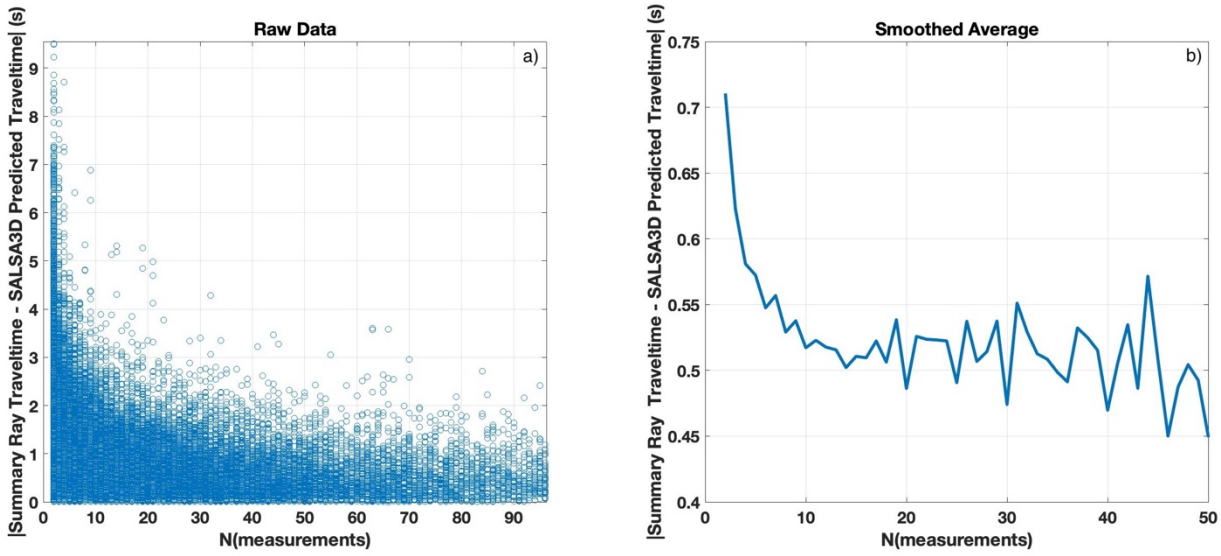

Figure 12: Impact of the number of rays in a ray bundle on summary ray accuracy. (a): plot of difference between summary ray measurement and prediction against number of rays in the ray bundle. (b): Average of left panel, displaying the mean of all error measurements for every value of N. The x-axis in b) is truncated at a lower value to avoid oscillations at larger values of N that are due to lower numbers of measurements in the corresponding bins.

This page left blank

4. IMPACT OF INCLUDING THE GT50 SUMMARY RAY DATASET IN TOMOGRAPHIC INVERSIONS

Having established the quality and accuracy of the P-wave summary dataset, we now explore inversions using the summary ray dataset as a preliminary test of their feasibility for inclusion in the dataset used for SALSA3D.

We first establish the questions that we seek to answer concerning the impact of using GT50 data and summary rays in a tomographic inversion:

1. Is the quality and distribution of the GT50 data sufficient to generate a tomographic model that presents a reasonable depiction of Earth structure, whether using summary rays or individual GT50 measurements?
2. Are any differences manifest when a model is inverted for using individual GT50 data as compared to summary ray measurements?
3. To what extent, if at all, does the information and coverage afforded by the GT50 summary rays complement the information in the GT < 50 data?
4. How does the use of GT50 data or GT50 summary rays impact posterior model covariance, relative to the GT <50 data?

To answer the above questions, we generate three versions of the P and Pn datasets described in the previous section. Firstly, to isolate the impact of using data other than the GT50 data, we generate one dataset which includes solely the P and Pn measurements with GT < 50, amounting to 3,783,266 measurements. Secondly, we generate a dataset which includes only GT50 measurements and incorporates all the summary rays calculated for that dataset, amounting to 12,820,416 measurements. Thirdly, to properly compare the effect of using summary rays to the effect of just using individual GT50 measurements, we follow the same approach as was used in the synthetic tests and generate a dataset of GT50 measurements that has exactly the same number of measurements as in the summary ray dataset, but for each summary ray, we randomly pick one of the individual GT50 measurements that contributes to the ray bundle and use that instead. Therefore, the spatial distribution of sources and receivers in the dataset of individual GT50 measurements and the resulting number of measurements in this dataset matches that of the summary ray dataset.

For the tomographic inversion, we follow the approach and framework established in Ballard et al., 2016. The inversion for a model of slowness in the Earth's mantle uses a Bayesian inference approach, incorporating an *a priori* model covariance matrix in lieu of *ad hoc* damping and Tikhonov regularization. We iteratively solve the least-squares problem for model slowness as a function of depth. The full workflow in Ballard et al., (2016) comprises multiple sets of model iterations as grid refinement and event relocation is performed. In the results presented here, we focus primarily on comparing results from a single set of iterations; this allows us to isolate the variations we see in the obtained solution to the impact of using a different dataset.

Firstly, we compare the performance of the two GT50 inversions in terms of the fit of the data. We find that for the dataset that includes summary rays, the inversion consistently achieves a lower (better fitting) RMS at all iterations (Figure 13) than for the dataset without summary rays. Since all other aspects of the inversion are identical between these inversions, we attribute the difference in RMS to a difference in data quality; as they represent an average of traveltimes from similar raypaths, the summary rays likely eliminate some of the inherent random error in the individual measurements and thus represent more accurate, higher-quality data, consequently achieving a lower traveltime RMS error. This is consistent with the conclusions inferred in Figure 10, where the published SALSA3D model is shown to be more consistent with summary ray traveltime measurements than GT50 measurements randomly picked from the same ray bundles.

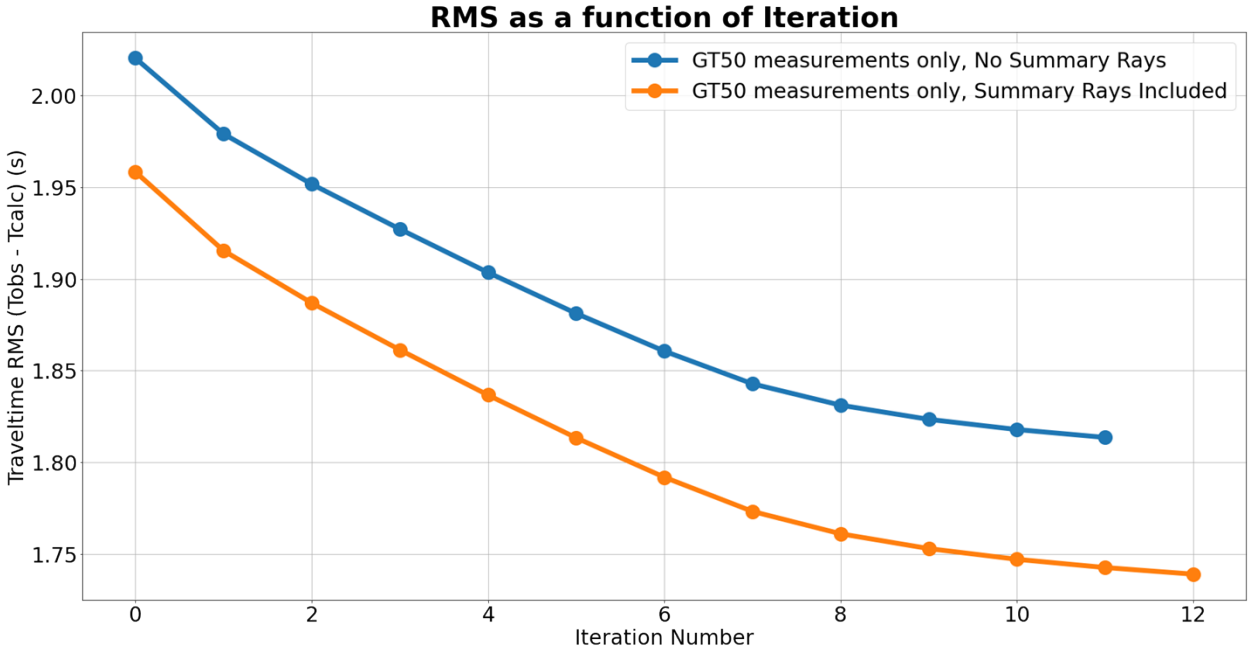

Figure 13: Inversion RMS as a function of iteration number, shown for two inversions that use only GT50 data. The inversions also use the same number of measurements and an effectively identical configuration of sources, receivers, and damping.

We also perform a comparison of the two models that are solved for using GT50 measurements. A visual inspection of the two models at several depths (Figure 14) shows no major differences, and in fact, a quantitative comparison of the models by taking the correlation coefficient between model slices at constant depths finds that they are virtually identical at all depths, achieving correlation coefficients above 0.99 at all depths analyzed (Figure 14). This comparison shows that using summary rays enables the inversion to still resolve similar heterogeneity to that resolved by the individual measurements. It may be the case, however, that as later adaption iterations use more refined grids and probe smaller length-scales in heterogeneity, the impact of the lower accuracy of the individual GT50 measurements may result in more noticeable differences in models.

We next compare the model solved for using summary rays to the model that is solved for using GT<50 data. A visual examination between these two models shows pronounced differences,

both in the strength of heterogeneity and in the locations of heterogeneity. The model solved for using only GT < 50 data shows stronger amplitudes than the summary ray model. This is visually observable when looking at the distribution of RMS dVp as a function of depth; this quantity is consistently higher for the GT<50 model, reflecting larger deviations from the mean. The correlation coefficient between the two models is consistently between 0.6 and 0.8, decreasing subtly from shallower to deeper depths; this illustrates significant differences in the pattern of heterogeneity in addition to the amplitude of heterogeneity. Examples of the locations where these differences manifest can be seen along trenches and subduction zones in southeast Asia at a depth of 200 km, where these features are much more distinct for the model using summary rays. At 200 km, there are also significantly higher amplitudes in Europe and the surrounding area in the GT <50 model. Mid-ocean ridges also appear to be more distinct in the model with summary rays. The differences in amplitude likely reflect the impact of the damping used, which was the same value but affects both inversions datasets differently due to the different number of measurements and the different source/receiver geometry involved. Due to these differences, it is difficult to draw conclusions about the resolving power of the summary rays in comparison to the GT < 50 dataset based on a comparison of the models alone.

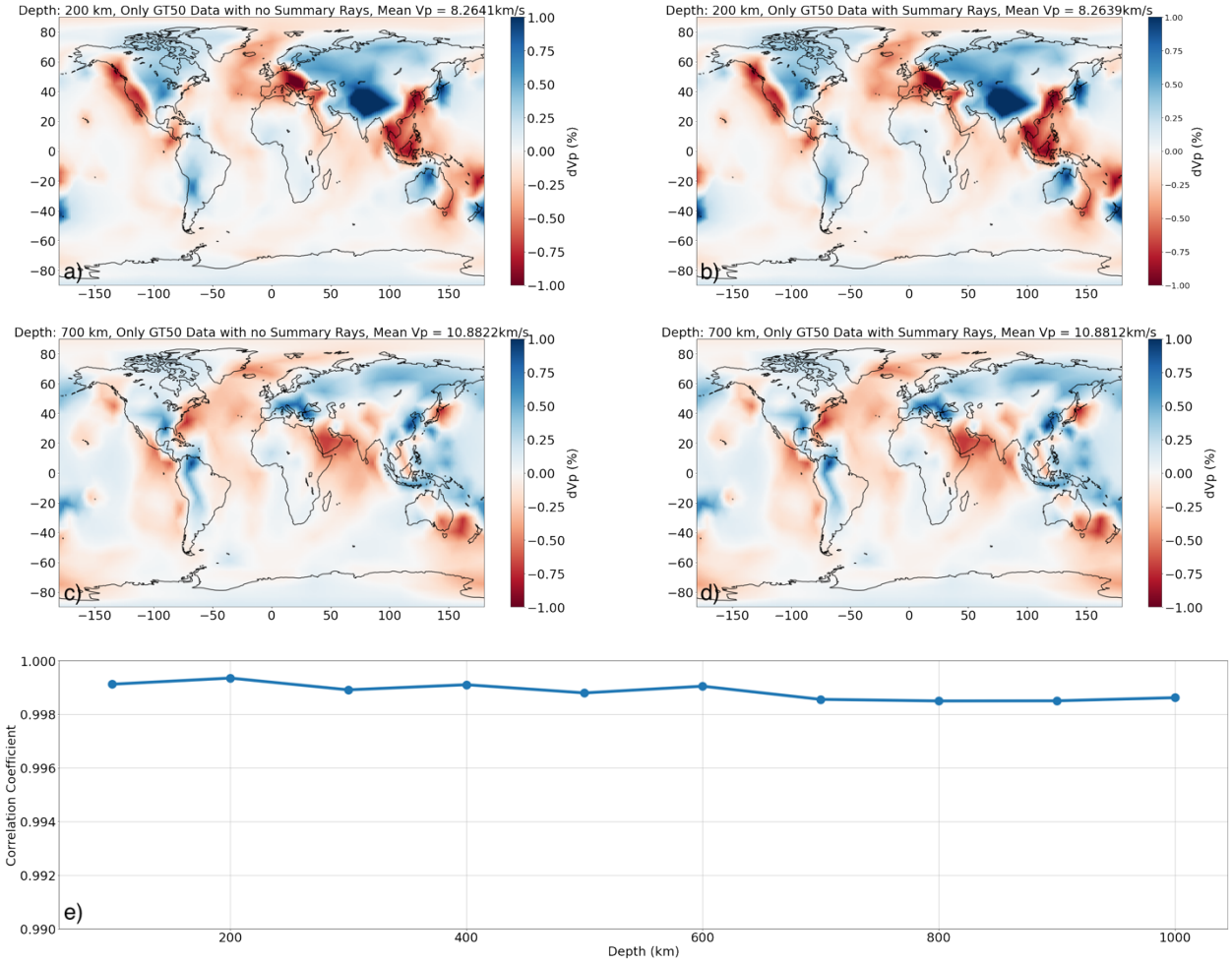

Figure 14: Comparison of dVp heterogeneity (relative to the mean Vp at that depth) solved for in inversions with and without summary rays. (a,b): Map-view plots of models at a depth of 200 km for the cases without (a) and with (b) summary rays. (c,d): Same as top, but for a depth of 700 km.

(e): Correlation coefficient as a function of depth between both models. Note the narrow range on the y-axis on the bottom panel.

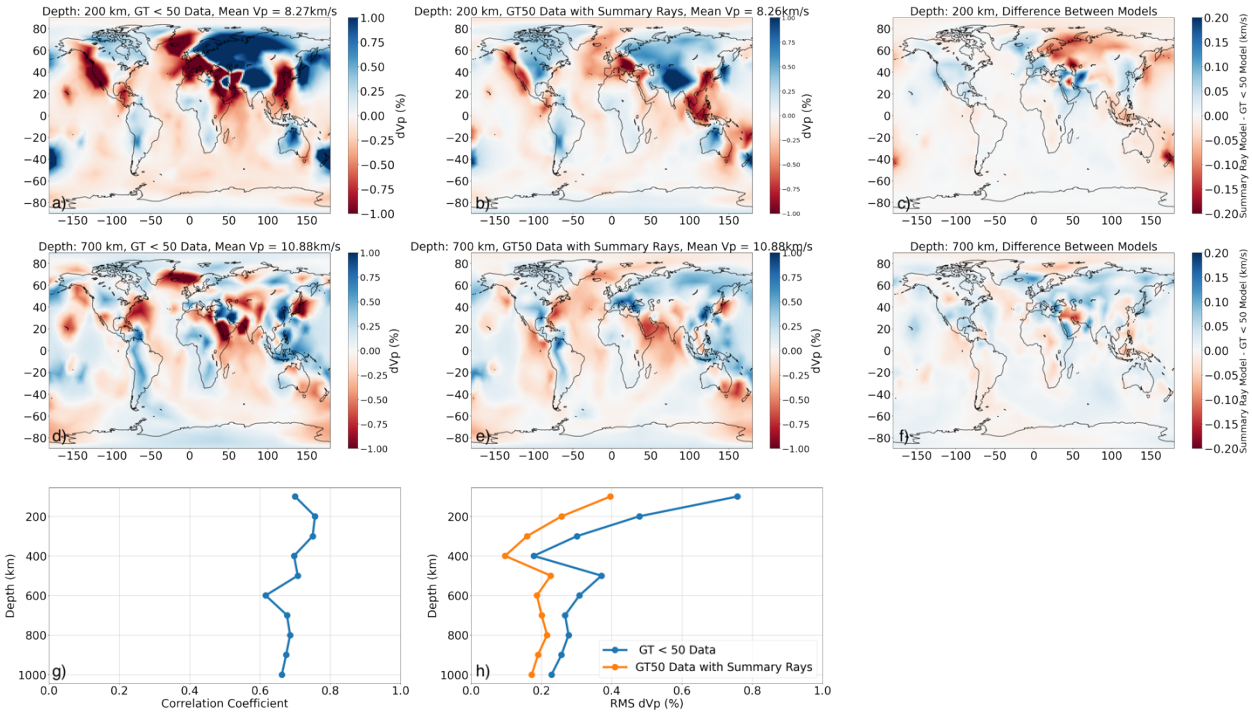

Figure 15: Comparison of dVp heterogeneity solved for in inversions using GT50 summary rays and using only GT < 50 data. (a,b): Map-view plots of models at a depth of 200 km for the cases using only GT < 50 data (a) and with GT50 summary rays (b). (c): Difference between panels (a) and (b) (d,e,f): As in the top row, but at a depth of 700 km. (g): Correlation coefficient as a function of depth between both models. (h): RMS dVp of both models as a function of depth.

To better understand the differences between both sets of models and explore how the summary ray dataset differs from the GT < 50 dataset in terms of its resolving capabilities, we explore the information contained in the diagonal of the model resolution matrix, which captures the effect of the differences in data coverage between the GT50 summary ray and the GT<50 dataset. Depth slices of the resolution matrix at several depths are shown in Figure 16. These images can be interpreted as more high value regions (red) are more well resolved.

Inspecting these maps of model resolution highlights recurring differences between the two models. The model constructed from data including summary rays consistently shows higher resolution along subduction zones in the southern hemisphere, particularly in South America, along the Andes, and also along the Sunda subduction zone in Southeast Asia, in a consistent pattern that continues southeast along New Zealand. On the other hand, the model constructed from only GT < 50 data shows higher resolution in much of Europe and Russia when compared to the model that uses only GT50 data.

These differences in resolution are likely due to the differences in the distribution of sources in the GT50 and GT < 50 datasets (Figure 3). The GT50 dataset includes far more earthquakes in the southern hemisphere and particularly along the subduction zones; a notable example is South America, where there are many regions of the Andean subduction zone which have no GT<50

events. On the other hand, the GT <50 dataset has many more stations in the Northern hemisphere.

The observed zones of differing resolution also correspond to some of the zones where the model heterogeneity is visually noticeably different. For instance, the summary ray model resolves low-velocity heterogeneity at 200 km that parallels the subduction zones in Southeast Asia, but the GT<50 model resolves a larger low-velocity anomaly east of China instead of recovering heterogeneity along subduction zones. Therefore, the differences in structural heterogeneity that we see in the inverted models likely are due to the underlying differences in the resolution of the data.

These patterns also appear in plots of the diagonal of the model covariance matrix (Figure 17). Regions with higher resolution generally show up as regions of lower model covariance, corresponding to lower slowness uncertainty.

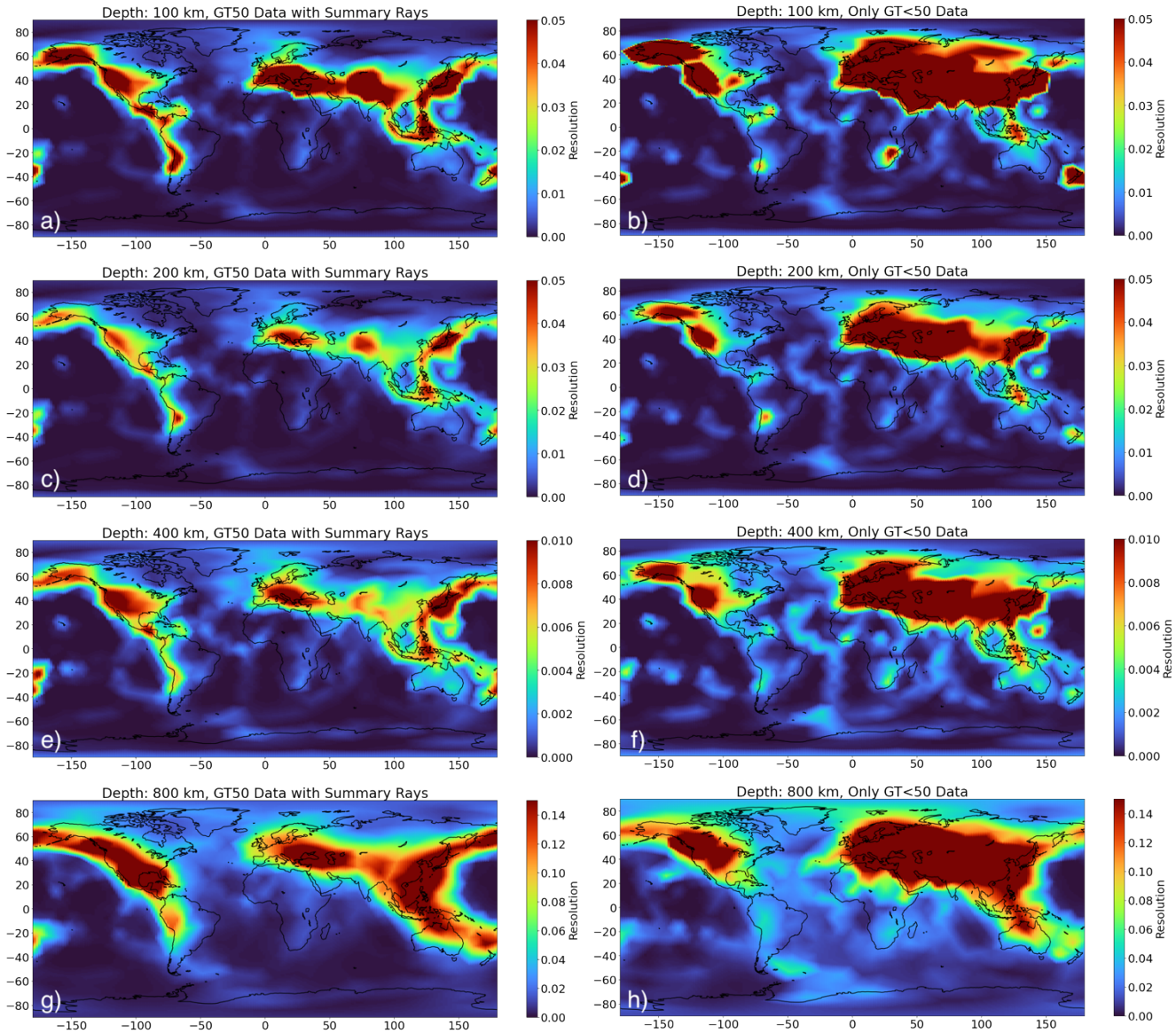

Figure 16: Map view plots of model resolution for (left) model constructed using only GT50 data including summary rays and (right) only GT < 50 data. Each row corresponds to a different depth quoted in the title of the subplot.

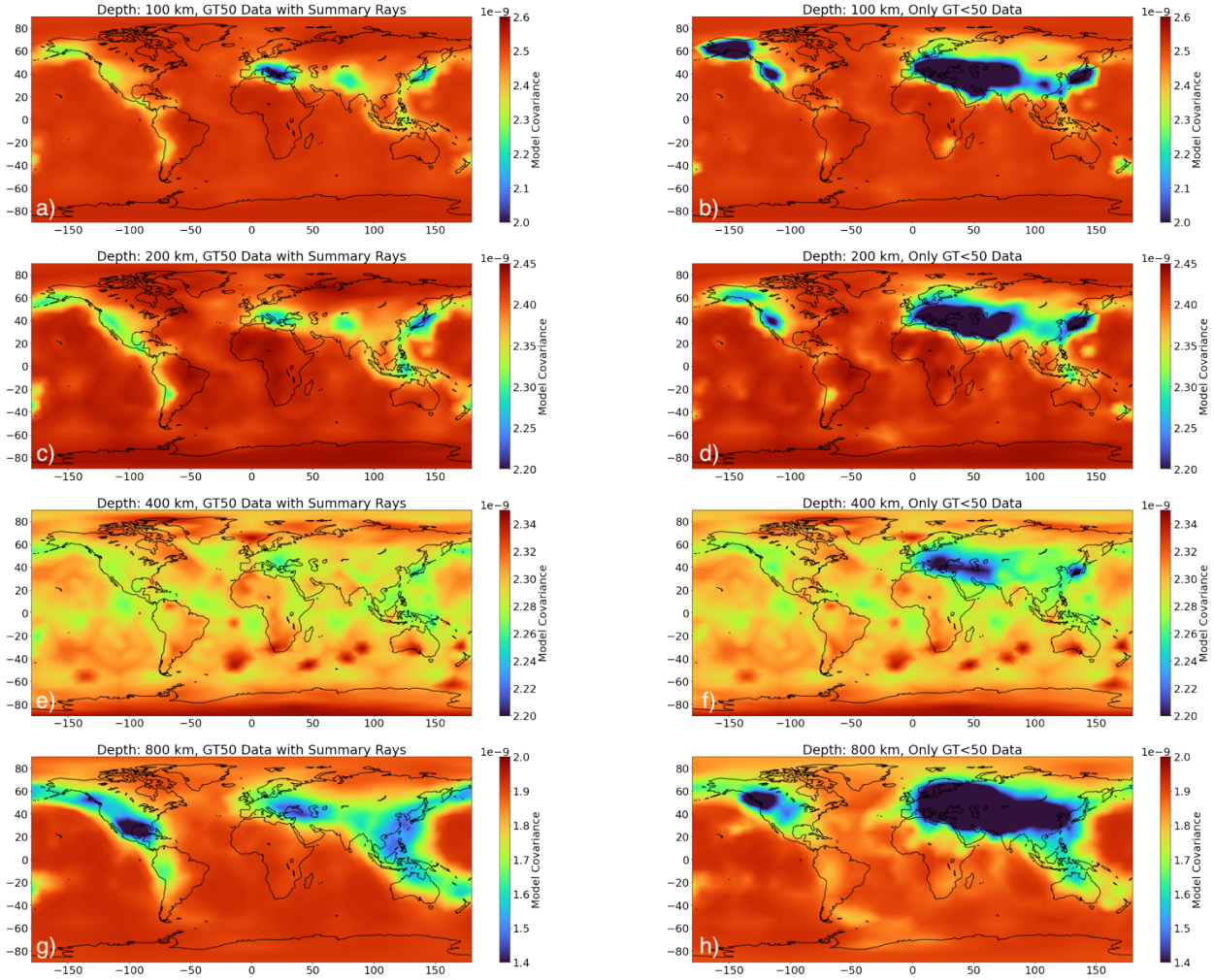

Figure 17: Map view plots of the diagonal of the model covariance matrix for (left) model constructed using only GT50 data including summary rays and (right) only GT < 50 data. Each row corresponds to a different depth quoted in the title of the subplot.

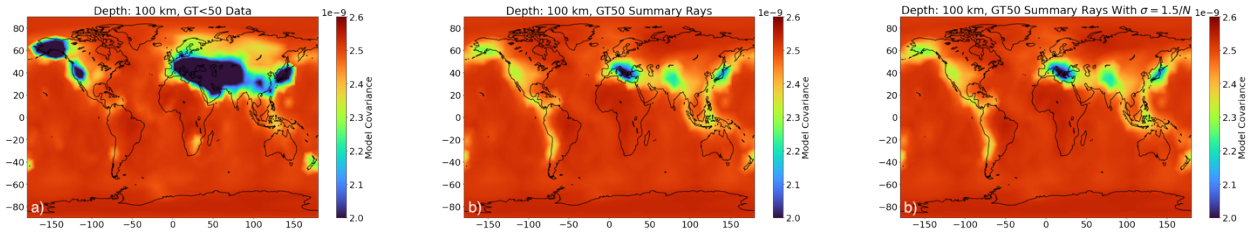

Figure 18: Impact of weighting uncertainty as 1.5/N. Model covariance at a depth of 100 km is plotted for models including only GT < 50 data (a), GT50 Summary Ray data (b), and GT50 Summary Ray data where uncertainty is defined as 1.5/N, where N is the number of measurements (c).

Motivated by our earlier observation of an increase in summary ray accuracy with an increasing number of measurements, we also explore the impact of an alternate data weighting on model covariance, setting summary ray measurement uncertainty to 1.5/N, where N is the number of measurements. Figure 18 shows the model covariance at a depth of 100 km for this inversion configuration. The differences between this model covariance and that obtained for the summary

rays are minimal and the two models are virtually indistinguishable. This result may reflect the control that the non-summary ray measurements in the GT50 dataset exert on the overall result, since the uncertainty of these measurements are unchanged. Alternatively, it may reflect that if we wish to incorporate weighting based on the number of measurements, it may be necessary to use a more dramatic scheme, such as 1.5/N^2, potentially increasing the exponent even further.

Overall, the tomographic inversions we have presented here show that (1) the GT50 summary ray dataset agrees well with the model solved for using the GT50 non-summary ray data, (2) the GT50 summary rays can achieve a better RMS data fit than the non-summary ray GT50 data, and (3) the GT<50 and GT50 datasets have uniquely different coverage and the GT50 summary rays can aid with model resolution, particularly along subduction zones in the southern hemisphere.

5. CONCLUSION

Using seismic data to locate events of interest requires an accurate tomographic model of wavespeed in the Earth's interior. In recent years, the development of such models has had to contend with exponentially growing datasets of measured traveltimes, which serve as fundamental datasets for these models and are particularly crucial for the construction of the SALSA3D model, which aims to accurately predict traveltimes with an objective estimate of the uncertainty. These large datasets inhibit computationally tractable and rapid inversions for these models. To mitigate this issue, we explore summary rays as an approach to reduce the redundancy and size inherent to these datasets and calculate them for a subset of events in the SALSA3D dataset that represent a particularly large and uncertain data source. We catalog the properties and characteristics of this truncated dataset.

We find that summary rays effectively reduce the size of such datasets and improve their accuracy, resulting in improved data fit in tomographic inversions. Furthermore, when summary rays are applied to tomographic inversions, we find certain regions of the Earth's surface are relatively prone to benefiting from the use of summary rays in tomographic inversions via both improved resolution and lower model uncertainty. The results presented here lay the foundation for the incorporation of summary rays in future inversions.

This page left blank

APPENDIX A. PRELIMINARY APPLICATION OF SUMMARY RAYS TO S-WAVE MEASUREMENTS

While the results of summary ray calculations on S waves are not the main focus of this report, we nonetheless briefly examine the effect of using summary rays on our datasets of S-wave measurements.

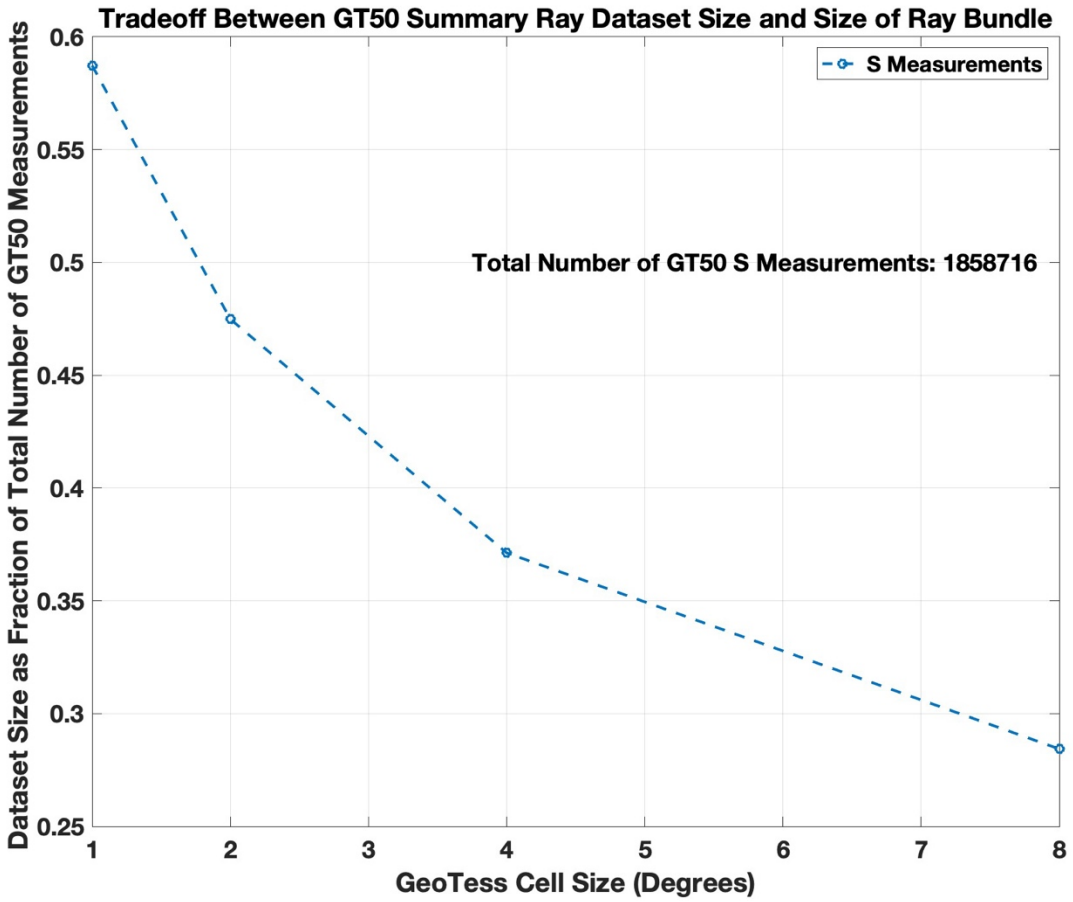

Figure 19: Impact of the use of Summary Rays on the size of the underlying GT50 dataset (as a fraction of the number of rays in original dataset), for S-wave measurements. The absolute number of rays in the original dataset is presented for reference in the top-right corner.

We find that applying summary rays to the S-wave dataset can yield a substantial reduction in the size of the resulting dataset (Figure 13). Though the impact is slightly less pronounced than for the corresponding P-wave dataset, we still find that the use of summary rays can reduce the size of the GT50 dataset by almost a factor of 2.

We also explore the stability and uncertainty of the summary ray calculations by calculating the standard deviation of all the measurements in each ray bundle for S-waves. When this is compared to the corresponding P-wave standard deviations, we find that the standard deviations

for S-waves are vastly larger (Figure 14).

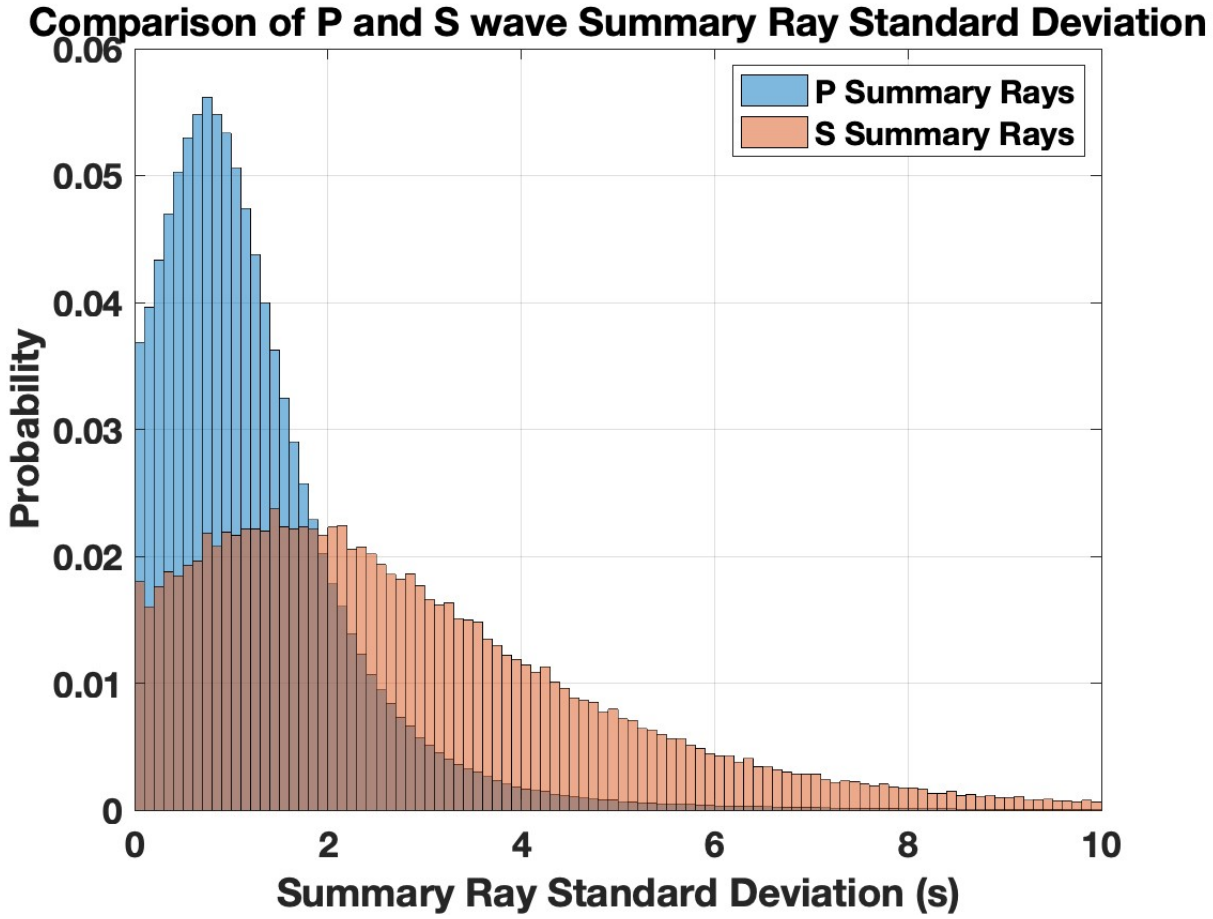

Figure 20: Normalized histograms of summary ray standard deviations for P and S measurements.

Assuming a constant scaling ratio for heterogeneity, the contribution to the standard deviation from structural heterogeneity would be expected to be less for S-waves than for P-waves, suggesting a lower median standard deviation if scatter was as governed by structural heterogeneity for S-waves than for P-waves. However, we find that the S-wave standard deviation is substantially larger than P-waves in general, suggesting a much larger contribution from random error to the underlying scatter contributing to S-wave summary rays. The increased random error inherent in the individual S-wave measurements also presents an opportunity in that the S-wave summary rays might improve substantially on measurement accuracy.

REFERENCES

[1] Ballard, S., Hipp, J. R., & Young, C. J. (2009). Efficient and accurate calculation of ray theory seismic traveltime through variable resolution 3D earth models. *Seismological Research Letters*, *80*(6), 989-999.

[2] Ballard, S., Hipp, J. R., Begnaud, M. L., Young, C. J., Encarnacao, A. V., Chael, E. P., & Phillips, W. S. (2016). SALSA3D: A tomographic model of compressional wave slowness in the Earth's mantle for improved travel-time prediction and travel-time prediction uncertainty. *Bulletin of the Seismological Society of America*, *106*(6), 2900-2916.

[3] Ballard, S., Hipp, J., Kraus, B., Encarnacao, A., & Young, C. (2016). GeoTess: A generalized Earth model software utility. *Seismological Research Letters*, *87*(3), 719-725.

[4] Bondár, István, et al. "Epicentre accuracy based on seismic network criteria." *Geophysical Journal International* 156.3 (2004): 483-496.

[5] Conley, A. (2021). *SALSA3D* (No. SAND2021-8194PE). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

[6] Begnaud, M. L., Ballard, S., Hipp, J. R., Encarnacao, A. V., Young, C. J., Phillips, W. S., & Rowe, C. A. (2021). *Validation and testing of SALSA3D: A global model of compressional wave speed for the crust and mantle* (No. LA-UR-20-27867). Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

[7] Bijwaard, H., Spakman, W., & Engdahl, E. R. (1998). Closing the gap between regional and global traveltime tomography. *Journal of Geophysical Research: Solid Earth*, *103*(B12), 30055-30078.

[8] Davies, J. H., Gudmundsson, O., & Clayton, R. W. (1992). Spectra of mantle shear wave velocity structure. *Geophysical journal international*, *108*(3), 865-882.

[9] Downey, N. J., Ballard, S., Hipp, J. R., & Begnaud, M. A. (2019). *LocOO3D User's Manual* (No. SAND-2019-9542). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

[10] Ekström, G., Nettles, M., & Dziewoński, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200*, 1-9.

[11] Gudmundsson, O., Davies, J. H., & Clayton, R. W. (1990). Stochastic analysis of global traveltime data: mantle heterogeneity and random errors in the ISC data. *Geophysical Journal International*, *102*(1), 25-43.

[12] Li, X. D., & Romanowicz, B. (1996). Global mantle shear velocity model developed using nonlinear asymptotic coupling theory. *Journal of Geophysical Research: Solid Earth*, *101*(B10), 22245-22272.

[13] Morelli, A., & Dziewonski, A. M. (1987). Topography of the core–mantle boundary and lateral homogeneity of the liquid core. *Nature*, *325*(6106), 678-683.

[14] National Research Council. (2012). The comprehensive nuclear test ban treaty: Technical issues for the United States.

[15] Ringler, Adam T., et al. "Achievements and prospects of global broadband seismographic networks after 30 years of continuous geophysical observations." *Reviews of Geophysics (1985)* 60.3 (2022).

[16] Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdağ, E., & Tromp, J. (2019). Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example. *Geophysical Journal International*, *219*(2), 1225-1236.

[17] Spakman, W., & Nolet, G. (1988). Imaging algorithms, accuracy and resolution in delay time tomography. In *Mathematical geophysics: A survey of recent developments in seismology and geodynamics* (pp. 155-187). Dordrecht: Springer Netherlands.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Andrea C. Conley	06752	acconle@sandia.gov
Robert W. Porritt	06756	rwporri@sandia.gov
Technical Library	1911	sanddocs@sandia.gov

This page left blank

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.