REVIEW REPORT FOR

Influence of Moho topography on seismic beamforming: implications for mantle scattering

published in *Seismica*

ROUND 1 REVIEWS

<u>REVIEWER 1</u>

I read the manuscript "Influence of Moho topography on seismic beamforming: implications for mantle scattering" with great interest.

The paper presents an interesting and relevant study on the impact of Moho topography on backazimuth measurements using seismic arrays in Alaska. The topic of the study is very interesting and important for the reliability of beamforming analyses and deep Earth imaging. The paper is generally well structured, and the methodology is clearly described. The combination of synthetic tests and real data (application to the seismic arrays in Alaska) presents a solid approach, and the results are potentially significant for the seismological community.

However, some aspects of the methodology and interpretation would benefit from further clarification. In particular, some key parameters (e.g. frequency band, array aperture, and selection thresholds) are chosen empirically without apparently a systematic assessment of their impact on the results.

In general, I find the study very promising (I particularly appreciated Section 4.2 "Discussion of Uncertainties") and suitable for publication after minor revision.

You can find my line-by-line suggestions for the manuscript in the attached PDF file.

I hope my review is helpful.

See annotated pages following this compilation of text comments.

<u>REVIEWER 2</u>

See review on separate pages following this compilation of text comments.

Answers to these reviews are inserted below, following this compilation of text comments.

ROUND 2 REVIEWS

REVIEWER 1

The authors have addressed all the comments point by point, also adding a new figure and a dedicated section discussing the possible origins of backazimuth offsets.

I appreciate the authors' efforts in improving the manuscript; in my opinion, the manuscript is now in very good shape and can be accepted for publication.

REVIEWER 2

See review on separate pages following this compilation of text comments.

Answers to these reviews are inserted below, following this compilation of text comments.

Influence of Moho topography on seismic beamforming: implications for mantle scattering

by Shubham Agrawal, Dan Frost and Philip Crotwell

1 Introduction

L. 31-32 I prefer to keep one opening sentence instead of 2 separate. If you follow this suggestion, please add "with" between "decade." and "regions"

L. 33 Please add a reference also for US Transportable Array in order to provide the reader with a reference source

L. 36 "... extract key seismic wavefield parameters"
here you can explicitly state which seismic wavefield parameters are meant (e.g., backazimuth, wave amplitude, etc.) rather than leaving this sentence generic.

L. 64-70 "... Demonstrating that a dipping Moho causes seismic waves to deviate from their expected paths..."

This effect is also noted in the paper of Colavitti et al. (2022), when is shown the effect of Moho local dip and local crustal velocities in the shooting of P-to-S converted waves (see Fig. 2). If you think it is necessary, you can add this work to your reference list:

Colavitti, L., Hetényi, G., and AlpArray Working Group (2022). A new approach to construct 3-D crustal shear-wave velocity models: method description and application to the Central Alps. Acta Geod Geophys 57, 529–562 (2022). https://doi.org/10.1007/s40328-022-00394-4.

L. 76-77 "Backazimuth offsets originating from Moho tilt could be up to 10°".
Please could you specify if this estimations are based on real observations or simulations?

L. 79 "P.P-wave" -> "P-to-P wave"

L. 81-82 "This illustrates how even moderate backazimuth anomalies can introduce large spatial errors in mantle imaging"
"Large" is a bit "qualitative", please could you explain in a more quantitative way (not necessarily to the text but as a reply to this comment) how it influences the choice of arrays or the design of tomography experiments?"

L. 85 "... such as IMS arrays" probably here you can say that IMS is the acronym for International Monitoring System (I guess ;))

2 Effect of tilt on seismic stations: tilted Moho case

2.1 Application to single station

L. 161-166 This is a nice summary of the application of Moho tilt to a single station but I was wondering what are the practical implications of these deviations?
In other words, at what level do they become significant for interpretation?

2.2 Application to an array of stations

L. 189-190 "The estimated backazimuth and incidence angle (for the whole array) are then compared to the true input values (for the central station of the array), and the differences are plotted as circles in Figure 2b".

Probably I did not understand why the true input are referred to the central station of the array rather than, for example, to an average over the array or to the global theoretical model?

In addition, are the differences you showed simple subtractions, absolute errors, RMS values, ... Please explain exactly how the differences shown are computed.

3 Effect of tilt on real data: Alaska

3.1 Data and pre-processing

L. 224 "... source depths greater than 60 km to ensure simpler source time functions..."
Why was a threshold of 60 km chosen? Would the results change significantly for a different cutoff (e.g., 50 or 100 km)?

3.2 Beamforming methodology

L. 243-244 "Rather than analyzing individual phases separately, ..."
I understand this is a methodological choice, but please explain why you combine the phases you mentioned. How does combining multiple phases (with different slowness and backazimuth) within a single time window affect the beamforming results?

In my opinion, this part is well described; however, several methodological choices would benefit from further justification. In particular, the slowness and backazimuth search ranges (0–10 s/° and ±50°) appear quite broad. Can the authors justify these choices and assess whether they may introduce contamination from incoherent or off-path energy?

3.3 Coherence and quality control

L. 263-264 "If this Max/Mean ratio exceeds 20 (based on trial and error), ..."
I understand this threshold is stated to be based on trail and error tecnique, but please could you clarify if similat trhresholds have been adopted in the literature? As it stands, the criterion appears a bit arbitrary and may affect the reproducibility of the method.

3.4 Constraining backazimuth offsets

L. 282-284 "Significant peaks were selected from this coherence timeseries using a prominence-based algorithm: only those exceeding 25% of the global maximum in amplitude and 10% in prominence were retained".
In my opinion, here there is a similar problem than before. The choice of 25%

(amplitude) and 10% (prominence) thresholds is not justified.
Are these values empirically tuned for this dataset, or adopted from previous studies? Please explain your choice replying here or in the manuscript.

3.5 Understanding backazimuth offsets in Alaska

OK

3.5.1 Impact of the Alaskan Moho

L. 315-316 “Our sub-arrays have apertures of 300-550 km (Figure 3), effectively averaging the Moho gradient over a large scale”.
Given the aperture of the sub-arrays, the Moho gradient is effectively averaged: could this spatial feature obscure smaller-scale structures that might influence the observed backazimuth offsets?

3.6 Control Case: Flat Moho region in US

OK

4 Correction and uncertainties

4.1 Correcting vespagrams for near-surface structure

L. 351-353 “Assuming that the entire offset arises due to refraction at the Moho, we back-project these rays to the base of the crust and apply Snell’s law to compute their incidence angles beneath the interface”. Probably I missed it, but here could the authors clarify which velocity model is used?

L. 358-365 If I understand correctly, the interpretation that any “remaining” backazimuth deviations can be attributed to mantle heterogeneities assuming that the Moho correction fully accounts for receiver-side effects. Could the authors justify this assumption and briefly discuss whether residual crustal complexities (e.g., intra-crustal structure, small-scale Moho variability) might still contribute?

4.2 Discussion of uncertainties

L. 382-384 “Therefore, the aperture and frequency values used in this study were selected through trial and error to suit the Alaskan context and are not necessarily generalizable to other station geometries”.

Could the authors briefly clarify to what extent their conclusions depend on these methodological choices?

4.3 Other possible origins of backazimuth-offsets

OK

5 Conclusions

In my opinion the conclusions are clear and concise; however, for completeness, it would be appropriate to include a statement acknowledging the limitations of the study, given the underlying assumptions. For example, that the backazimuth offsets are primarily controlled by Moho structure, but may also be influenced by uncertainties in velocity models and array configuration.

I hope my suggestions will be helpful in improving the manuscript!

Review: Influence of Moho topography on seismic beamforming: implications for mantle scattering, by Shubham Agrawal, Daniel A. Frost and Philip Crotwell

The manuscript performs synthetic tests of the influence of near receiver structure, namely Moho dip, onto measured incidence angle (slowness and backazimuth) of an incident wavefront and compares the results with recorded data from Alaska as an example of complex Moho structure and the central US with simpler Moho structure. There has been a larger body of work related to "mislocation vectors" in the CTBTO community and with the increased usage of ad-hoc arrays facilitated by dense station deployments the results are interesting for a wider audience. The manuscript is well written and documented although a few points remain unclear and should be improved.

1) The authors mainly consider Moho structure as the source of the mis direction with a few other possibilities discussed in section 4.3 (instrument effects, surface topography, sedimentary cover, anisotropy, mantle structure). What seems to be missing from this list are velocity variations in the crust. As far as I can see the authors assume a constant velocity crust in their analysis. I think both the impact of lateral velocity variations and the limitations of the 1D velocity structure assumption need to be discussed (e.g. in the form of velocity gradients in the crust perhaps?) as part of the manuscript.
2) On line 170ff the authors state that most IMS arrays are cross-shaped. This is actually incorrect as most IMS arrays are circular with stations on concentric circles or randomly distributed to reduce sidelobes in the Array Response function with many much smaller than apertures of 5deg (with most apertures below 20 km). Only the 1960s UKAEA arrays as part of the IMS seem to be cross shaped. E.g. Gibbons, Pure Appl. Geophy,. 2012, doi: 10.1007/s00024-012-0613-2. I recognise that the array likely has limited impact on this analysis but might break some of the symmetries the authors see in the deviations.
3) I note that the authors do not reference the work by Flanagan, Myers and Simmons (Model-based corrections to observed back azimuth and slowness observations from a dipping Mohorovičić discontinuity LLNL-CONF-563592, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2012). Although it follows a slightly different approach and estimates Moho structure from the backazimuth deviations (vice versa to this study) it seems worthwhile to mention.
4) In their theoretical study the authors use ray theory. Is there a limitation to this approach based on wavelength related effect and the array aperture that is not captured by this approach (the authors discuss Fresnel zone in another context later, but that seems to be unrelated to wavefield effects that might make their estimates an upper bound)?

5) Line 187: The authors describe an optimization routine to the delay times to determine wavefront direction, but do not provide any detail. Could the authors please add more information here.
6) Line 322: Here the authors state that Fig. 6 "demonstrates that the sign of the observed backazimuth offsets is consistent with the sign of the direction Moho gradient". Unfortunately, I don't see that in this figure and it seems a crucial constraint to their analysis. The means of the backazimuths for the two populations are close to zero and given the scatter in the results I find it hard to see any relationship. The statement that "the relationship is not strictly linear" seems overly positive. Perhaps the authors can design a statistical test to make this point quantitatively? For consistency the authors also might want to show the same analysis for the central US TA data example.
7) Line 360ff: The authors apply a correction to bring the main phases back to the great circle path. Is the take-home message for readers here that if a study looks at relative deviations between phases, that these can be interpreted without concern for the near receiver structure unless the incidence angles of the phases are drastically different? As mentioned earlier this correction is as proposed by Flanagan et al.
8) Line 405ff: The authors mention topographic corrections in this paragraph which seem to have been applied throughout the analysis. As these seem to be applied in the processing I think they need to be discussed earlier in the manuscript.

Minor points:

a) Line 55: To be consistent with the naming of the other arrays in the list it would be better to spell out Gräfenberg (rather than GRF).
b) Line 56: There is a wider catalog of studies into mislocations that the authors might want to reference e.g. Hao and Zheng, BSSA, 2009 and Hao and Zheng, BSSA, 2010 for arrays in China (and there might be more), Koch & Kradolfer, 1999 (GSETT3), Pirli et al, 2007, 10.12681/bgsg.16876 (Tripoli Array), Faber et al., J. Geophysics, 60, 1986 (GRF), Bokelmann, BSSA, 1995, 10.1785/BSSA0850051456 (GERESS)
c) Line 89: The work by Ward et al (GJI, 2021, https://doi.org/10.1093/gji/ggab196) might be worth adding to this list of work using ad-hoc arrays and array processing in an automated way.
d) Very minor: Line 133 and Figure 1 – The Moho dip is stated as 2.9° and 2.86° in text and caption. Should perhaps be consistent.
e) Figure 1 a) shows an array but the test is for a single station which is somewhat confusing. As an array is shown in Figure 2 for the array test again the array here seems not needed?

f) Line 185: The authors write about largest possible source-receiver distance. Does that mean P going into the core shadow? As Pdiff is described later the statement seems confusing.
g) Figure 4: The figure is quite busy. The authors describe vertical error bars in the text, but there seem to be also white contour lines which might describe prominence? Please update the caption. Also e and f seem to be for the event - array example in c) / d)? Please specify.
h) Line 284: How do you define prominence in this context?
i) Line 460: The link to code and scripts seems to be missing.

Response to Reviewers

Influence of Moho topography on seismic beamforming: implications for mantle scattering

Shubham Agrawal, Dan Frost, Philip Crotwell

We thank the editor for handling our manuscript.

We have addressed several minor comments by both the reviewers and added text in the manuscript. The comments were mostly to clarify our choices and how they might affect the overall interpretation. Further, both the reviewers mentioned to consider the effect of lateral velocity heterogeneity (intra- crustal structure) on array based measurements. We had now added a new Figure 9 and a discussion in section 4.3 “Other possible origins of backazimuth-offsets”.

In the following, the comments of the two reviewers are addressed, with our response in *red*.

Reviewer A:

I read the manuscript "Influence of Moho topography on seismic beamforming: implications for mantle scattering" with great interest.

The paper presents an interesting and relevant study on the impact of Moho topography on backazimuth measurements using seismic arrays in Alaska. The topic of the study is very interesting and important for the reliability of beamforming analyses and deep Earth imaging. The paper is generally well structured, and the methodology is clearly described. The combination of synthetic tests and real data (application to the seismic arrays in Alaska) presents a solid approach, and the results are potentially significant for the seismological community.

However, some aspects of the methodology and interpretation would benefit from further clarification. In particular, some key parameters (e.g. frequency band, array aperture, and selection thresholds) are chosen empirically without apparently a systematic assessment of their impact on the results.

In general, I find the study very promising (I particularly appreciated Section 4.2 "Discussion of Uncertainties") and suitable for publication after minor revision.

You can find my line-by-line suggestions for the manuscript in the attached PDF file.

I hope my review is helpful.

We thank the reviewer for their constructive comments and suggestions.

1 Introduction

L. 31-32 I prefer to keep one opening sentence instead of 2 separate. If you follow this suggestion, please add “with” between “decade.” and “regions”.

The opening line now reads as:
“Seismic station coverage across the globe has burgeoned significantly over the past decade, with regions of Earth bereft of seismic stations seeing new deployments.”

L. 33 Please add a reference also for US Transportable Array in order to provide the reader with a reference source

Added.

L. 36 “... extract key seismic wavefield parameters” here you can explicitly state which seismic wavefield parameters are meant (e.g., backazimuth, wave amplitude, etc.) rather than leaving this sentence generic.

The reason to not mention them here was to explain the parameters (slowness and backazimuth) first time they are mentioned, which is in the second paragraph. We understand reviewers' concern about leaving the statement generic, but we think the explanation works better in the second paragraph.

L. 64-70 "… Demonstrating that a dipping Moho causes seismic waves to deviate from their expected paths…"

This effect is also noted in the paper of Colavitti et al. (2022), when is shown the effect of Moho local dip and local crustal velocities in the shooting of P-to-S converted waves (see Fig. 2). If you think it is necessary, you can add this work to your reference list:

Colavitti, L., Hetényi, G., and AlpArray Working Group (2022). A new approach to construct 3-D crustal shear-wave velocity models: method description and application to the Central Alps. Acta Geod Geophys 57, 529–562 (2022). https://doi.org/10.1007/s40328-022-00394-4.

We thank the reviewer for bringing this study to our notice. We have added it.

L. 76-77 : "Backazimuth offsets originating from Moho tilt could be up to 10°".

Please could you specify if this estimations are based on real observations or simulations?

This now reads as:
"It has been observed that backazimuth offsets originating from Moho tilt…."

L. 79 : "P.P-wave" -> "P-to-P wave"

Done.

L. 81-82: "This illustrates how even moderate backazimuth anomalies can introduce large spatial errors in mantle imaging"

"Large" is a bit "qualitative", please could you explain in a more quantitative way (not necessarily to the text but as a reply to this comment) how it influences the choice of arrays or the design of tomography experiments?"

By "large" we mean large relative to the spatial accuracy needed for locating mantle scatterers. As mentioned in the sentence preceding the quoted sentence, a 5° backazimuth error can shift the inferred location of a mid-mantle P-to-P scatterer by up to ~500 km. The process of finding the scatterer location is independent of arrays as it involves fitting two P rays for the observed time, slowness, and backazimuth.

L. 85: "… such as IMS arrays" probably here you can say that IMS is the acronym for International Monitoring System (I guess ;))

We have defined the acronym in the second paragraph, hence the subsequent usage.

2 Effect of tilt on seismic stations: tilted Moho case

2.1 Application to single station

L. 161-166: This is a nice summary of the application of Moho tilt to a single station but I was wondering what are the practical implications of these deviations?

In other words, at what level do they become significant for interpretation?

We have now added this summary to convey the significance of effect of Moho tolt on single station:

"These deviations become significant when they exceed the resolution of the beamforming, or when they are large enough to substantially misplace inferred scatterer locations. As mentioned in Section 1, a backazimuth offset of 5°, which is reasonable for Moho tilts representative of Alaska, can shift a scatterer location by hundreds of kilometers. The phase-dependence of these offsets is also of relevance as differential Moho-induced offsets between P and PP waves could be misinterpreted as evidence of off-path scattering if near-receiver crustal structure is not properly accounted for."

2.2 Application to an array of stations

L. 189-190: "The estimated backazimuth and incidence angle (for the whole array) are then compared to the true input values (for the central station of the array), and the differences are plotted as circles in Figure 2b".

Probably I did not understand why the true input are referred to the central station of the array rather than, for example, to an average over the array or to the global theoretical model? In addition, are the differences you showed simple subtractions, absolute errors, RMS values, ... Please explain exactly how the differences shown are computed.

We have added the following text to explain a similar query of the other reviewer related to optimization routine used:

"We measure time delays of this wave at each station in the array, using Equation 5. We trace rays through the mantle, Moho, and crust to each station and calculate the arrival time at each station. We then estimate the wavefront direction and incidence angle by minimizing the sum of squared residuals between observed arrival times and those predicted by a planar wavefront model, treating backazimuth and incidence angle as free parameters. The optimization is performed using Nelder-Mead algorithm, initialized from a fixed starting point (5°, 5°), with the incidence angle constrained to physically reasonable values (0-90°). The resulting best-fitting parameters represent the apparent wavefront direction and incidence angle for the array, plotted as circles in Figure 2b."

3 Effect of tilt on real data: Alaska

3.1 Data and pre-processing

L. 224 "... source depths greater than 60 km to ensure simpler source time functions..." Why was a threshold of 60 km chosen? Would the results change significantly for a different cutoff (e.g., 50 or 100 km)?

We chose earthquakes deeper than 60 km to ensure simpler source time functions and better separation of the direct P arrival from the depth phases in the vespagrams. Shallower events, especially in subduction settings, commonly produce more complicated source-time functions and smaller P–pP separation, which makes the arrivals harder to identify reliably. We do not expect the main conclusions to change with a stricter cutoff such as 100 km, although the number of usable events would decrease. A shallower cutoff such as 50 km would likely admit more events with less phase separation. It is to be noted that using a depth threshold is a standard choice in studies targeting teleseismic P-wave arrivals and mantle scattering (e.g. Kaneshima and Helffrich, 2010; Bagley et al., 2013; Haugland et al. 2017, Kaneshima, 2018).

We have clarified this choice in text:

".....and source depths greater than 60 km to ensure simpler source time functions and clearer separation of the direct P arrival from its depth phases (e.g. Kaneshima and Helffrich, 2010; Bagley et al., 2013)."

3.2 Beamforming methodology

L. 243-244 "Rather than analyzing individual phases separately, ..."

I understand this is a methodological choice, but please explain why you combine the phases you mentioned. How does combining multiple phases (with different slowness and backazimuth) within a single time window affect the beamforming results?

We combine these phases in a single time window because the beamforming is evaluated on a grid of time, slowness, and backazimuth. Therefore, arrivals with different slowness and backazimuth can still be visualized and identified within the same vespagram. In other words, our method allows all the phases (including P, pP, sP, and PP) to be analyzed consistently without making any assumptions about their respective slowness/backazimuth. The main practical consequence is increased computation time.

In my opinion, this part is well described; however, several methodological choices would benefit from further justification. In particular, the slowness and backazimuth search ranges (0–10 s/° and ±50°) appear quite broad. Can the authors justify these choices and assess whether they may introduce contamination from incoherent or off-path energy?

The search windows were chosen to include the expected slowness and backazimuth range of the phases of interest over the distance range (65°-120°) used in this study, together with some margin for deviations from the 1-D prediction. As for backazimuth, we agree that the search range for backazimuth is broader than the backazimuth offsets observed (~ 10°). However, the broader search does not by itself imply contamination as the subsequent analysis retains only well-constrained high-coherence arrival with peaks exceeding 25% of the global maximum. These steps remove most incoherent or weak off-path energy.

3.3 Coherence and quality control

L. 263-264 "If this Max/Mean ratio exceeds 20 (based on trial and error), …"

I understand this threshold is stated to be based on trail and error tecnique, but please could you clarify if similat trhresholds have been adopted in the literature? As it stands, the criterion appears a bit arbitrary and may affect the reproducibility of the method.

As far as we know, this specific Max/Mean metric and threshold are not standard in the literature. Many array-based studies manually assess vespagram quality as the target arrivals are often weak. In our case, because we focus on direct phases rather than weak scattered phases, we were able to use an automated vespagram selection criterion. The Max/Mean threshold should therefore be viewed as a quality-control parameter rather than a universal value, akin to signal-to-noise ratio. We have clarified this in the manuscript so that the method is reproducible and its empirical nature is explicit.

"If this Max/Mean ratio exceeds 20 (based on trial and error), meaning that there are clear peaks of coherence associated with well-resolved phases, the vespagram is retained for further analysis. This process removes 28% of the vespagrams. This threshold was introduced as an automated quality-control criterion for this dataset and may require retuning for different noise levels, station density, or array geometry."

3.4 Constraining backazimuth offsets

L. 282-284 "Significant peaks were selected from this coherence timeseries using a prominence-based algorithm: only those exceeding 25% of the global maximum in amplitude and 10% in prominence were retained".

In my opinion, here there is a similar problem than before. The choice of 25% (amplitude) and 10% (prominence) thresholds is not justified. Are these values empirically tuned for this dataset, or adopted from previous studies? Please explain your choice replying here or in the manuscript.

Similar to Max/mean ratio, these thresholds were tuned for this dataset and were not adopted from previous array studies. To our knowledge, this type of automated peak selection is not commonly used in array-based analyses, where phase identification is often done manually. The 25% amplitude threshold and 10% prominence threshold were introduced as practical quality-control criteria to retain only well-defined coherence peaks and suppress weaker incoherent or off-path energy. We have added the text to justify this choice:

"Significant peaks were selected from this coherence timeseries using a prominence-based algorithm: we retained only peaks whose amplitude exceeded 25% of the global maximum and whose prominence, defined as the peak height relative to the surrounding local background, exceeded 10% of the global maximum (plus symbols in Figures 4 e,f). These thresholds were chosen for this dataset so that only well-defined, physically meaningful peaks were used, suppressing noise and weaker incoherent or off-path energy."

3.5 Understanding backazimuth offsets in Alaska

OK

3.5.1 Impact of the Alaskan Moho

L. 315-316 "Our sub-arrays have apertures of 300-550 km (Figure 3), effectively averaging the Moho gradient over a large scale".

Given the aperture of the sub-arrays, the Moho gradient is effectively averaged: could this spatial feature obscure smaller-scale structures that might influence the observed backazimuth offsets?

Yes; the reviewer is right in thinking that large aperture means that each sub-array averages structure over a broad area, so smaller-scale Moho variations may be obscured. We mentioned this in Fresnel-zone argument in Section 4.2. We have now expanded it a bit to state this more clearly:

"However, because each station has its own Fresnel zone, the array effectively samples a region comparable to its surface aperture plus approximately one Fresnel-zone radius on either side, which remains small relative to the ~275 km radius used to compute the local Moho gradient and therefore does not significantly influence our gradient estimates. Despite this, the large aperture of our sub-arrays mean that the Moho structure is averaged over a similarly broad area. As a result, smaller-scale heterogeneity in Moho topography may be smoothed and may contribute less to the observed backazimuth offsets."

3.6 Control Case: Flat Moho region in US

OK

4 Correction and uncertainties

4.1 Correcting vespagrams for near-surface structure

L. 351-353: "Assuming that the entire offset arises due to refraction at the Moho, we back- project these rays to the base of the crust and apply Snell's law to compute their incidence angles beneath the interface". Probably I missed it, but here could the authors clarify which velocity model is used?

We use the same crust and mantle P-wave velocities as in the synthetic optimization in Section 2.2. We have now clarified this in Section 4.1:

"In this back-projection, we use the same P-wave velocities adopted in Section 2.2, with Vcrust, Vmantle as 5.8 km/s and 8 km/s, respectively."

L. 358-365: If I understand correctly, the interpretation that any "remaining" backazimuth deviations can be attributed to mantle heterogeneities assuming that the Moho correction fully accounts for receiver-side effects. Could the authors justify this assumption and briefly discuss whether residual crustal complexities (e.g., intra- crustal structure, small-scale Moho variability) might still contribute?

Reviewer C also had a similar query. We kindly ask the reviewer to look at our response to their Comment #1. We had added a discussion and a new figure to address potential contribution from intra- crustal structure.

4.2 Discussion of uncertainties

L. 382-384 "Therefore, the aperture and frequency values used in this study were selected through trial and error to suit the Alaskan context and are not necessarily generalizable to other station geometries".

Could the authors briefly clarify to what extent their conclusions depend on these methodological choices?

Thank you. We observed that what depends on the methodological choices is the quality and resolution of individual vespagrams, not necessarily the overall interpretation. We now clarify this further:

"Therefore, the aperture and frequency values used in this study were selected through trial and error to balance aliasing, station availability, and vespagram quality in Alaska. These choices affect the resolution and robustness of individual slowness and backazimuth measurements, and may influence the exact magnitude of measured offsets, but they do not alter the main conclusion that systematic backazimuth deviations in Alaska are consistent with Moho structure and are substantially larger than those observed in the flat-Moho control region."

4.3 Other possible origins of backazimuth-offsets

OK

5 Conclusions

In my opinion the conclusions are clear and concise; however, for completeness, it would be appropriate to include a statement acknowledging the limitations of the study, given the underlying assumptions. For example, that the backazimuth offsets are primarily controlled by Moho structure, but may also be influenced by uncertainties in velocity models and array configuration.

We have added the following line in the conclusion as per reviewer's suggestion:

"Although Moho topography explains the majority of the backazimuth offsets, other factors, including lateral crustal velocity structure, array geometry, crustal anisotropy, sediments, and slab-related heterogeneity, may also contribute."

I hope my suggestions will be helpful in improving the manuscript!

We are grateful to the reviewer for their positive outlook towards our manuscript. Their feedback was indeed constructive and helped improved manuscript's clarity!

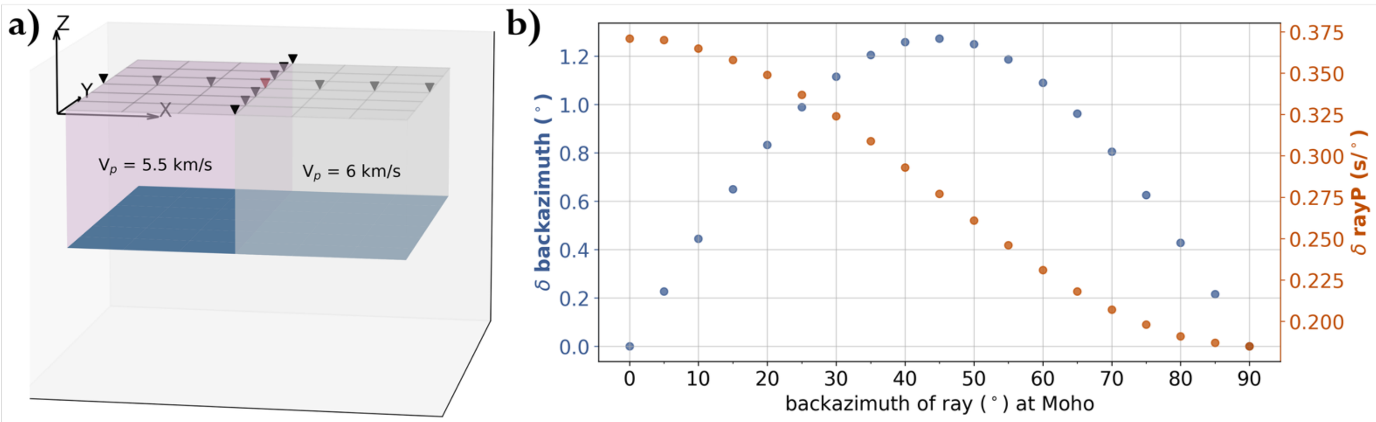
Reviewer C:

The manuscript performs synthetic tests of the influence of near receiver structure, namely Moho dip, onto measured incidence angle (slowness and backazimuth) of an incident wavefront and compares the results with recorded data from Alaska as an example of complex Moho structure and the central US with simpler Moho structure. There has been a larger body of work related to "mislocation vectors" in the CTBTO community and with the increased usage of ad-hoc arrays facilitated by dense station deployments the results are interesting for a wider audience. The manuscript is well written and documented although a few points remain unclear and should be improved.

We thank the reviewer for their positive outlook and useful suggestions.

1) The authors mainly consider Moho structure as the source of the misdirection with a few other possibilities discussed in section 4.3 (instrument, surface topography, sedimentary cover, anisotropy, mantle structure). What seems to be missing from this list are velocity variations in the crust. As far as I can see the authors assume a constant velocity crust in their analysis. I think both the impact of lateral velocity variations and the limitations of the 1D velocity structure assumption need to be discussed (e.g. in the form of velocity gradients in the crust perhaps?) as part of the manuscript.

We thank the reviewer for raising this point. A similar concern was also raised by Reviewer A. While we have focused on Moho topography as the dominant source of wavefront deflection, we acknowledge that intra-crustal velocity variations, both vertical and lateral, can also influence backazimuth and slowness measurements. The wavefront deviations are most sensitive to the largest velocity contrast along the raypath, which is the crust-mantle boundary; therefore, intra-crustal velocity variations are expected to produce smaller deflections. Nonetheless, to assess the potential contribution of lateral velocity variations in the crust, we designed a simple experiment. We introduced a 10% lateral velocity contrast in the crust (Vp = 5.5 km/s vs. 6 km/s, as shown in the figure below) while keeping the Moho flat, and applied the same optimization procedure used in Section 2.2 to estimate the resulting backazimuth and slowness deviations. We find that this lateral velocity variation produces backazimuth deviations of up to 1.2° and slowness deviations of up to 0.375 s/°. Relative to the tilted Moho case, the backazimuth deviations are smaller by up to 80%, suggesting Moho topography remains the dominant control on backazimuth. However, the slowness deviations reach up to 70% of the values produced by Moho tilt, indicating that slowness is considerably more sensitive to lateral crustal velocity variations than backazimuth is.



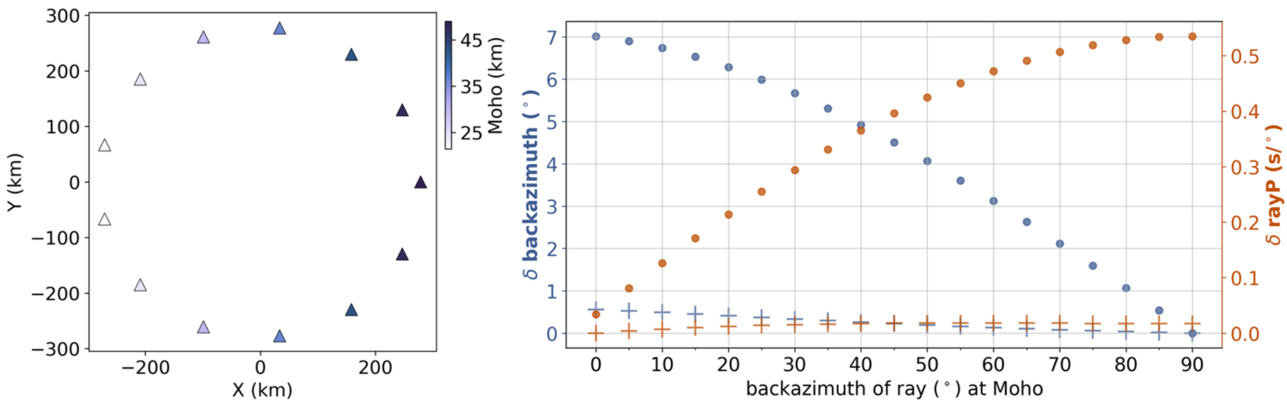
We have now added the above figure and the following text in section 4.3 to address this:

"Lateral velocity variations within the crust represent an additional potential source of backazimuth and slowness anomalies. In the synthetic tests, we assume a laterally uniform crustal velocity, but real crustal structure across Alaska includes significant velocity heterogeneity. To assess the potential influence of such variations, we performed a synthetic test in which we introduced a 10% lateral velocity contrast across the crust (Vp = 5.5 km/s vs. 6.0 km/s) while keeping the Moho flat, and applied the same wavefront optimization procedure as in Section 2.2 (Figure 9a). This configuration produces backazimuth deviations of up to 1.2° and slowness deviations of up to 0.375 s/° (Figure 9b). Compared to the tilted Moho case (Figure 2b), the resulting backazimuth deviations are up to 80%, demonstrating that Moho topography is the dominant contributor to the backazimuth offsets we observe. However, the slowness deviations approach up to 70% of the values produced by Moho tilt, suggesting that slowness is more sensitive to lateral crustal velocity variations than backazimuth. Given that our analysis focuses primarily on backazimuth offsets, and that our observed offsets are well reproduced by Moho tilt alone, we do not apply explicit corrections for lateral crustal velocity variations. Nevertheless, in regions with strong crustal heterogeneity, such effects may contribute to residual slowness anomalies that are not fully accounted for by the Moho correction alone. We also note that our test represents an end-member scenario; a uniform 10% velocity contrast throughout the entire crust is unlikely, and more realistic configurations involving localized or thin-layered contrasts would likely produce smaller effects."

2) On line 170 the authors state that most IMS arrays are cross-shaped. This is actually incorrect as most IMS arrays are circular with stations on concentric circles or randomly distributed to reduce sidelobes in the Array Response function with many much smaller than apertures of 5deg (with most apertures below 20 km). Only the 1960s UKAEA arrays as part of the IMS seem to be cross shaped. E.g. Gibbons, Pure Appl. Geophy,. 2012, doi: 10.1007/s00024-012-0613- 2. I recognise that the array likely has limited impact on this analysis but might break some of the symmetries the authors see in the deviations.

We recognize how that statement can be misconstrued. We have now removed the mention of IMS arrays being cross-shaped.

As for breaking the symmetries, using a cross-shaped array results in the most asymmetry. For instance, for an incoming backazimuth of 90°, all stations on Y axis will experience the wavefront at the same time. However, if it were any other shape, it would not be the case. Nonetheless, we performed the analysis with a circular array of same dimension and number of stations and recreated Figure 2 (below). As is evident, the values of backazimuth and slowness offsets remain the same.



We have added the following text:
"We also performed the same experiment using a circular array of radius 2.5° and found that the results were equivalent to the cross case."

3) I note that the authors do not reference the work by Flanagan, Myers and Simmons (Model-based corrections to observed back azimuth and slowness observations from a dipping Mohorovičić discontinuity. LLNL-CONF-563592, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2012). Although it follows a slightly different approach and estimates Moho structure from the backazimuth deviations (vice versa to this study) it seems worthwhile to mention.

We have added this study.

4) In their theoretical study the authors use ray theory. Is there a limitation to this approach based on wavelength related effect and the array aperture that is not captured by this approach (the authors discuss Fresnel zone in another context later, but that seems to be unrelated to wavefield that might make their estimates an upper bound)?

In Section 2 use ray theory as an idealized reference case. This approach does not include finite-frequency wavefield effects, so it does not capture the spatial averaging associated with wavelength and Fresnel-zone width. As discussed later in the manuscript, the sampled structure is averaged over a finite area, which reduces sensitivity to small-scale heterogeneity. We have added the following the end of Section 4.2 to make this explicit:

"Further, it is important to note that our ray-theory based synthetic tests in Section 2 do not include finite-frequency wavefield effects. As a result, they should be regarded as the upper limit estimate of the effect of Moho topography. Wavelength-dependent spatial averaging may reduce sensitivity to small-scale Moho heterogeneity relative to the ray-theoretical prediction and could result in smaller values than predicted by ray-theory."

5) Line187: The authors describe an optimization routine to the delay times to determine wavefront direction, but do not provide any detail. Could the authors please add more information here.

We have added the following as clarification for the optimization process:

"We measure time delays of this wave at each station in the array, using Equation 5. We trace rays through the mantle, Moho, and crust to each station and calculate the arrival time at each station. We then estimate the wavefront direction and incidence angle by minimizing the sum of squared residuals between observed arrival times and those predicted by a planar wavefront model, treating backazimuth and incidence angle as free parameters. The optimization is performed using Nelder-Mead algorithm, initialized from a fixed starting point (5°, 5°), with the incidence angle constrained to physically reasonable values (0-90°). The resulting best-fitting parameters represent the apparent wavefront direction and incidence angle for the array, plotted as circles in Figure 2b."

6) Line322: Here the authors state that Fig.6 "demonstrates that the sign of the observed backazimuth offsets is consistent with the sign of the direction Moho gradient". Unfortunately, I don't see that in this figure and it seems a crucial constraint to their analysis. The means of the backazimuths for the two populations are close to zero and given the scatter in the results I find it hard to see any relationship. The statement that "the relationship is not strictly linear" seems overly positive. Perhaps the authors can design a statistical test to make this point quantitatively? For consistency the authors also might want to show the same analysis for the central US TA data example.

Thank you for raising this. We agree that the wording might be considered a bit strong given the scatter in Figure 6 and the relationship is not clearly identifiable from the figure alone. To test if the signs of backazimuth and Moho gradient are related, we performed a sign-agreement test. We found that 62% of observations have consistent polarity between the directional Moho gradient and the backazimuth offset. Further, we also computed the Spearman ratio to test the monotonic relationship between the directional Moho gradient and the backazimuth offset. Spearman rank correlation yields $\rho = 0.25$, indicating a monotonic relationship, albeit a bit weak. We have now edited the text to convey this information:

"Figure 6 shows the observed backazimuth offsets with respect to the directional Moho gradient. We find that 62% of observations have consistent polarity between the directional Moho gradient and the backazimuth offset, indicating a modest but statistically significant relationship. A Spearman rank test yields $\rho = 0.25$, indicating a weak monotonic relationship between the two parameters. However, there is considerable scatter in the data, which suggests that Moho topography is a primary but not exclusive control on the observed offsets, and we explore the potential causes of this in Section 4.3."

7) Line360: The authors apply a correction to bring the main phases back to the great circle path. Is the take-home message for readers here that if a study looks at relative deviations between phases, that these can be interpreted without concern for the near receiver structure unless the incidence angles of the phases are drastically different?

Thank you. We agree with the reviewer's intuition, but our results indicate that relative deviations between phases are not, in general, independent of near-receiver structure. As shown in Figure 2b, the effect of a tilted Moho on backazimuth and slowness measured at an array is non-linear with respect to the incoming wave direction. For example, two phase pairs separated by the same angular difference (e.g., 0–10° versus 80–90°) can produce different relative backazimuth anomalies. In our synthetic example, this corresponds to a change of ~1° compared to ~0.2°, demonstrating that relative measurements do not necessarily cancel the receiver-side effect.

8) Line405: The authors mention topographic corrections in this paragraph which seem to have been applied throughout the analysis. As these seem to be applied in the processing I think they need to be discussed earlier in the manuscript.

The topographic corrections were applied during beamforming and are mentioned earlier in the manuscript at the end of Section 3.2:

"Further, prior to beamforming, we account for the individual topography beneath each seismic station by converting station elevation to a travel-time correction using a P-wave crustal velocity of 3 km/s and incorporating this correction into the moveout calculation."

Minor points:

a) Line55: To be consistent with the naming of the other arrays in the list it would be better to spell out Gräfenberg (rather than GRF).

Done.

b) Line56: There is a wider catalog of studies in to mislocations that the authors might want to reference e.g. Hao and Zheng, BSSA, 2009 and Hao and Zheng, BSSA, 2010 for arrays in China (and there might be more), Koch & Kradolfer, 1999 (GSETT3), Pirli et al, 2007, 10.12681/bgsg.16876 (Tripoli Array), Faber et al., J. Geophysics, 60, 1986 (GRF), Bokelmann, BSSA, 1995, 10.1785/BSSA0850051456 (GERESS)

We are grateful to the reviewer for bringing these studies to our attention. We have now added them to the list.

c) Line89: The work by Ward etal (GJI,2021,https://doi.org/10.1093/gji/ggab196) might be worth adding to this list of work using ad-hoc arrays and array processing in an automated way.

Thanks for the suggestion. We have added this.

d) Very minor: Line 133 and Figure1– The Moho dip is stated as 2.9° and 2.86° in text and caption. Should perhaps be consistent.

Fixed.

e) Figure1a) shows an array but the test is for a single station which is somewhat confusing. As an array is shown in Figure 2 for the array test again the array here seems not needed?

We understand reviewer's concern about Figure 2. However, the stations are coloured differently for helping distinguish the two synthetic experiments. The figure caption illustrates this:

"Backazimuth and ray parameter (horizontal slowness) are measured at the red station for the single-station case, while the black and red triangles indicate the array used in Section 2.2."

f) Line 185: The authors write about largest possible source-receiver distance. Does that mean P going into the core shadow? As Pdiff is described later the statement seems confusing.

We selected earthquake at epicentral distances between 65° and 120° from the stations in Alaska. Therefore, for the array-analysis on real data, Pdiff is mentioned as P enters the shadow zone for distances > ~97°. For the synthetic Snell's law-based experiment, we therefore chose an incident angle corresponding to the largest possible source-receiver distance for direct P wave. We have clarified the wording as below:

"…corresponding to the steepest angle of incidence for a direct P-wave at the largest possible source–receiver distance i.e., before the P shadow zone."

g) Figure 4: The figure is quite busy. The authors describe vertical error bars in the text, but there seem to be also white contour lines which might describe prominence? Please update the caption. Also e and f seem to be for the event - array example in c) / d)? Please specify.

We thank the reviewer for pointing out these omissions on our part. The updated caption is copied below:

"Vespagrams and measurement of backazimuth offsets. (a) An example of a well-constrained vespagram in slowness–time space with contours drawn as solid white lines. The black dashed line separates low- and high-slowness regions. Predicted arrival times and slownesses are shown as pink circles. (b) Same as (a), but in backazimuth–time space, measured relative to the great-circle path (red dashed line). The red diamond marks the point of global maximum coherence. (c,d) Same as (a,b), but for a different earthquake–sub-array pair exhibiting backazimuth offsets. (e,f) Peaks corresponding to high-coherence regions in slowness–time (e) and backazimuth–time (f) space for vespagrams in c and d. Circles are colored by the slowness or backazimuth value of the coherence maximum within a 5-second moving time window. Crosses indicate prominent peaks identified by the peak-picking algorithm. The mean backazimuth offset for prominent peaks in the low-slowness region (≤6 s/°) is 4°"

h) Line284: How do you define prominence in this context?

Thank you. We will clarify that prominence refers to how much a peak stands above the surrounding coherence background, rather than its absolute amplitude alone. It now reads as:

"Significant peaks were selected from this coherence timeseries using a prominence-based algorithm: we retained only peaks whose amplitude exceeded 25% of the global maximum and whose prominence, defined as the peak height relative to the surrounding local background, exceeded 10% of the global maximum (plus symbols in Figures 4 e,f). These thresholds were chosen empirically for this dataset so that only well-defined, physically meaningful peaks were used, suppressing noise and weaker incoherent or off-path energy."

i) Line 460: The link to code and scripts seems to be missing.

It is fixed now.

"The beamforming code and scripts to generate the figure can be found on Zenodo (https://doi.org/10.5281/zenodo.19053135; Agrawal et al., 2026)."

Review – Revision: Aggarwal, Frost and Crotwell – Seismica 2026

I thank the authors for their careful reply to my comments regarding the manuscript. In general I think my comments have been addressed. There are 2 points where potentially clarification is needed:

1) In my original review I commented on the perceived correlation between the observed backazimuth offsets and the directional Moho gradients. In the revised manuscript the authors state that they performed a sign-agreement test with 62% of the observations having consistent polarity between directional Moho gradient and backazimuth offset and a Spearman rank test with $\rho=0.25$. The manuscript now states that the rank test is statistically significant and note a weak, monotonic relationships for the Spearman test.

   62% seems only slightly above random (50%), it might be that the sample size is big enough to state that 62% is significant but I have my doubts.
   Similarly, I am not sure if the Spearman rank test is the best test to show the correlation and the mean of the positive and negative datasets (diamonds in Fig. 6) seem to be clearly within the standard deviations of the dataset. My estimate is that a dataset size of a minimum of 60-70 samples (which the authors clearly fulfil) is required for any kind of statistical significance. But I find judging how robust this estimate is given the sample size very difficult. But I note the careful language the authors use here.

   I am not sure if a permutation test (likely bootstrapping given the sample size) is a better approach?

2) The authors reference Niazi (1966) as laying the foundation of this work. Niazi (1966) develops an analytical expression linking Moho dip and backazimuth deviations. It would be interesting to see how the earlier expression compares to this much older work and its application to the data examples.

As a side comment: Reviewer A asked about the determination of significant peaks (L282-284) with the reply from the authors that this type of automated peak selection is not commonly used in array-based analyses. There are examples in the literature that use permutation tests and unsupervised machine learning (clustering algorithms) very efficiently to this end. E.g. Ward et al., GJI, 2021 (doi: 10.1093/gji/ggab196) and Ward et al., JGR, 2023 (doi: 10.1029/2022JB026260). These approaches might support the authors choice of parameters.

Response to Reviewers: Round 2

Influence of Moho topography on seismic beamforming: implications for mantle scattering

Shubham Agrawal, Dan Frost, Philip Crotwell

We thank the editor Dr. György Hetényi for the prompt handling our manuscript.

We have addressed the three comments by reviewer B and added text in the manuscript. Reviewer A were satisfied with our changes made during the first round of the revision.

In the following, the comments of the reviewer B are addressed, with our response in ***red***.

Reviewer B:

I thank the authors for their careful reply to my comments regarding the manuscript. In general I think my comments have been addressed. There are 2 points where potentially clarification is needed:

1) In my original review I commented on the perceived correlation between the observed backazimuth offsets and the directional Moho gradients. In the revised manuscript the authors state that they performed a sign-agreement test with 62% of the observations having consistent polarity between directional Moho gradient and backazimuth offset and a Spearman rank test with r=0.25. The manuscript now states that the rank test is statistically significant and note a weak, monotonic relationships for the Spearman test.

62% seems only slightly above random (50%), it might be that the sample size is big enough to state that 62% is significant but I have my doubts.

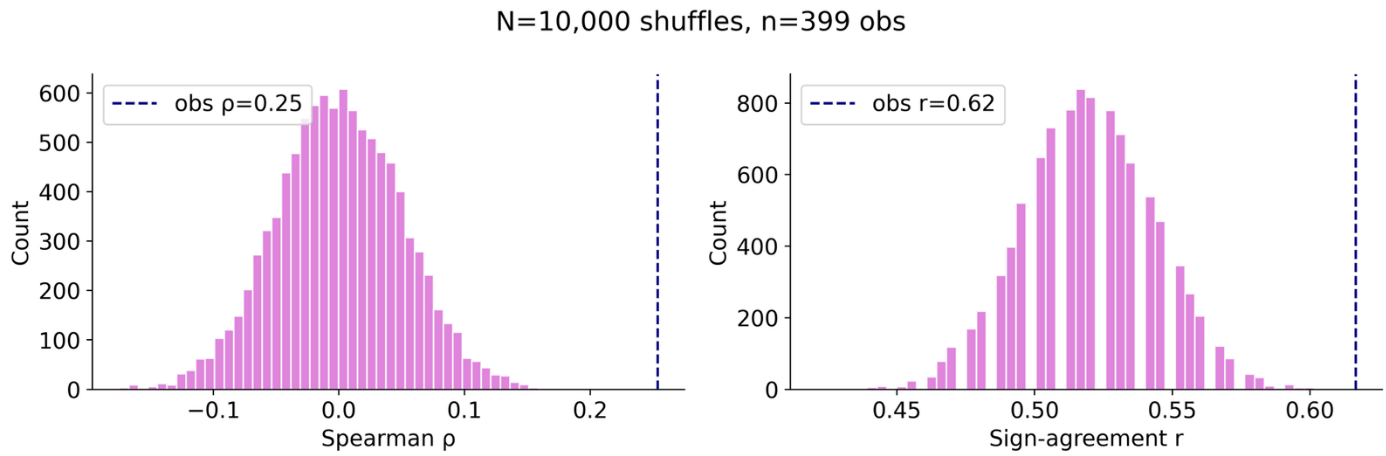
Similarly, I am not sure if the Spearman rank test is the best test to show the correlation and the mean of the positive and negative datasets (diamonds in Fig. 6) seem to be clearly within the standard deviations of the dataset. My estimate is that a dataset size of a minimum of 60-70 samples (which the authors clearly fulfil) is required for any kind of statistical significance. But I find judging how robust this estimate is given the sample size very difficult. But I note the careful language the authors use here.

I am not sure if a permutation test (likely bootstrapping given the sample size) is a better approach?

We agree with the reviewer to use randomization to test the robustness of the spearman (rho) and sign-agreement test (r). We shuffled the backazimuth offsets 10,000 times (breaking any association between grad values and offsets), and calculated rho and r for each iteration. As is observed (figure below), the observed $\rho = 0.25$ and r= 62% both fell in the extreme tail of their respective null distributions, suggesting the observation is clearly non-random.

We have added the following in the text:

“A Spearman rank test yields $\rho = 0.25$, indicating a weak monotonic relationship between the two parameters. A randomization test (10,000 shuffles of the backazimuth offsets) confirms this is unlikely to arise by chance, with both the observed ρ (0.25) and sign-agreement value (62%) exceeding all 10,000 permuted values.”



2) The authors reference Niazi (1966) as laying the foundation of this work. Niazi (1966) develops an analytical expression linking Moho dip and backazimuth deviations. It would be interesting to see how the earlier expression compares to this much older work and its application to the data examples.

We thank the reviewer for this suggestion. Niazi (1966) derived analytical expressions for backazimuth deviation and dT/dΔ corrections at a single station (small-aperture array) above a dipping Moho, tabulating results for dip angles of 2°, 5°, and 8° and velocity contrasts of 0.7, 0.8, and 0.9. They provide corrections for slowness and azimuth for these combinations of Moho dip and velocity contrast. To address the reviewers suggestion, we compare our values in Figure 1 to Niazi's table 2A and 2B (Vl/V2 = 0.7 and dip angle = 5.0°). We have added the following text to the manuscript:

"These observations and trends are consistent with the values reported by Niazi (1966). For a Moho dip angle of 5° and a velocity contrast of 0.7, they estimated backazimuthal deviations of 7° to 4° for incident angles between 20° to 40° (Table 2B in Niazi, 1966). In Figure 1c, for a comparable Moho dip (~5.7°) and velocity contrast (0.725), we observe deviations ranging from 6.5° to 4°, suggesting agreement between our measurements and their predictions."

However, our study extends beyond Niazi's (1966) framework as we apply these corrections to large-aperture arrays (~500 km; Section 2.2), which are substantially larger than the ~10 km aperture TFSO array considered by Niazi. Therefore, our analysis evaluates the effects of realistic Moho structure on modern regional-scale array measurements.

As a side comment: Reviewer A asked about the determination of significant peaks (L282-284) with the reply from the authors that this type of automated peak selection is not commonly used in array-based analyses. There are examples in the literature that use permutation tests and unsupervised machine learning (clustering algorithms) very efficiently to this end. E.g. Ward et al., GJI, 2021 (doi: 10.1093/gji/ggab196) and Ward et al., JGR, 2023 (doi: 10.1029/2022JB026260). These approaches might support the authors choice of parameters.

We thank the reviewer for pointing us to Ward et al. (2021) who present an automated method using bootstrap sampling combined with the DBSCAN clustering algorithm to identify seismic arrivals and estimate their slowness vector uncertainties without visual inspection. Ward et al.'s method and ours are theoretically complementary but there is one key difference: they identify which arrivals (on a phase by phase basis) are present while we characterize the bulk offset of the wavefield. We have added the following line in the main text:

"Ward et al. (2021) developed an automated method to identify arrivals in slowness vespagrams with formal uncertainty estimates. While their approach is optimized for characterizing specific phases within a focused time window, our vespagrams encompass multiple phases simultaneously (up to ~15 energy lobes in some cases), making a prominence-based peak-finding more appropriate."