Dear Editor,

Thank you for your time and revision. We have considered the Reviewers' suggestions to improve the manuscript.

We have attached two versions of the revised manuscript: a copy with the underlined modifications (manuscript_R2_including_corrections) and a cleaned copy (manuscript_R2). We have provided a response for each of the Reviewer's comments in the following.

There are two comments that are connected to each other on which we would kindly request the Editor's decision:

Reviewer#2 in his first comment suggests performing the inversion using one of the presented numerical examples but with lower data coverage to confirm the conclusion that straight-ray tomography works quite well for data acquired with optimised acquisition layout while curved ray provide an improvement in low coverage due to non-optimal acquisition layout. We have hence taken one of the synthetic data and carried out the inversion after decimating the DC to reduce the coverage and included the results in the response to that comment. However, we did not include this result in the manuscript because we think that simply decimating the data to reduce the coverage of a dataset acquired with an optimised acquisition layout would not be equal to having a lower coverage due to a non-optimal acquisition layout, as in the CNR field example. So, we do not feel that this example is very significant and that it adds value to the manuscript. If the Editor decides that it is necessary to include this example also in the manuscript, we would do so.

On the other hand, the last comment of Reviewer#1 is: "Since the highlighted findings in the conclusions section include the influence of acquisition layout and the benefit of curved-ray SWT in un-optimized acquisition layout, providing some additional noise-free tests using a numerical model (similar to the CNR site) might be helpful for convincing the readers. The test could be submitted as supplementary materials".

We think that this request (not mandatory according to the reviewer if we properly understand) would represent a more meaningful example than the one described above, and it would be certainly possible to carry it out. However, we have not provided such new examples since the CNR site is a controlled site with known features and has been used as the benchmark in several previous studies. More importantly, our conclusion from the CNR example is sufficiently soft not to include more studies.

We would like to leave it to Editor to decide whether the presented examples are sufficient, or it is necessary to provide any new numerical example.

Best regards,

The Authors

Responses to the comments of Reviwer#1:

It is great to see that the authors revise their manuscript carefully according to the remarks made by reviewers and editors, and the quality of the paper is greatly improved. It could be ready for publication after some minor revisions:

(1) Line 160, Page 6. Since the regularization term in the objective function is the l2-norm of the differences between each neighboring model and the optimization algorithm always tends to minimize the objective function toward zero, the sentence "it means that the vs difference between the neighboring cells is constrained to 1000 m/s" could make the readers confusing.

We have removed the sentence "it means that the vs difference between the neighboring cells is constrained to 1000 m/s" to avoid confusion.

(2) Line 200, Page 8. It might not need to describe the advanced features of the utilized simulation tool (such as viscoelastic) since they are not used in the paper.

Following this suggestion, we have removed the following sentence from the manuscript:

"… consider viscoelastic wave propagation effects such as attenuation and dispersion, …".

(3) Line 395, Page 24. "Figure 14 depicts the ray paths of …", please check the figure number, here it should be figure 16, right?

Correct. We have corrected the reference to the figure.

(4) Since the highlighted findings in the conclusions section include the influence of acquisition layout and the benefit of curved-ray SWT in un-optimized acquisition layout, providing some additional noise-free tests using a numerical model (similar to the CNR site) might be helpful for convincing the readers. The test could be submitted as supplementary materials.

We agree that providing more examples could be helpful for the readers to understand the importance of the source positions optimisation. However, it is not the topic of this paper. At the same time, the CNR site is a controlled site with known features and has been used as the benchmark in several previous studies (Boiero, 2009; Teodor et al., 2017; Khosro Anjom et al., 2019; Hu et al., 2021).

Providing a new numerical example is possible, but it would be a very time-consuming process (takes approximately a few weeks) and cannot be performed within the given time frame for minor revision (one week).

More importantly, as we have mentioned in the Conclusions that "We showed that the classical cross-spread acquisition layout (which was used in the CNR example) may not provide high DC coverage. In this case, the improvement of inversion results from curved-ray SWT can be significant." We think that this conclusion is sufficiently soft to not include more studies. This can be an interesting topic of further investigation.

References

Boiero, D.: Surface wave analysis for building shear wave velocity models: Ph.D. thesis, Politecnico di Torino, 2009.

Hu, S., Zhao, Y., Socco, L.V., and Ge, S.: Retrieving 2-D laterally varying structures from multistation surface wave dispersion curves using multiscale window analysis. Geophys. J. Int., **227**(2), 1418-1438, 2021.

Khosro Anjom F., Teodor, D., Comina, C., Brossier, R., Virieux, J., Socco L.V.: Full waveform matching of vp and vs models from surface waves, Geophys. J. Int., **218**, 1873-1891, 2019

Teodor, D., Comina, C., Socco, L.V., Brossier, R., Trinh, P.T., and Virieux, J.: Initial model design for full-waveform inversion—preliminary elastic modeling from surface waves data analysis: in Extended Abstract in 36th GNGTS national convention, 733–756, 2017.

Responses to the comments of Reviewer#2

Dear Editor,
Dear Authors,

the manuscript has significantly improved and (almost) all of my comments have been addressed appropriately.
Comment that has not been taken into account
"One of your key results is that in a scenario with low data coverage, the curved-ray approach performs significantly better. But your synthetic examples do not prove that since you run no example with low coverage. I would suggest that you take one of your two synthetic tests and test whether this conclusion holds."
I still think it would be important to properly test this hypothesis. Doing so in a synthetic test would give the reader a better feeling what influence the optimized shot positions have and where to draw the limit between poor and good data coverage. In your conclusions you highlight the importance of optimized shot positions, but is this really the problem in case study 4? If I compare figure 11 and figure 17, it seems like there is no data for many source positions. So even with optimized shot positions, I might run into the same problem of too low data coverage?

Figure 11 (in the previous version of the manuscript) shows the acquisition scheme of the CNR example while Figure 17 (in the previous version of the manuscript) depicts the impact of the data weighting on the inversion results for the Pijnacker example. I assume the reviewer meant "If I compare figure 11 and figure 16" instead of "If I compare figure 11 and figure 17".

The acquisition layout for the CNR example has not been optimized and Figure 16 shows the obtained low data coverage from this set up for the wavelength range of 0-1 m. It is true that there is no data for some shots (by comparing Figures 11 and 16) and the main reason is that no receiver pair is aligned with some shots, and therefore no corresponding DC has been estimated from them. So, it shows the importance of shot position optimization which can prevent this before acquiring the data.

So, to answer the last question (So even with optimized shot positions, I might run into the same problem of too low data coverage?), in the optimization of shot positions the data illumination and

alignment of receiver pairs with each shot are considered and the best shots are picked to generate the data. Therefore, the risk of running into the problem of too low data coverage would significantly decrease.

Moreover, the data coverage is not the only problem here. It is the obtained low data coverage because of using a non-optimized acquisition set up. So, reducing the data coverage for one of the numerical examples might not show the importance of shot positions optimisation and it would not provide a very similar example to the CNR site.

Nonetheless, we have randomly picked 10 % of the original dispersion curves for the Blocky model and carried out the inversions only using those dispersion curves. We have shown the results in the figure below (Fig. 1):
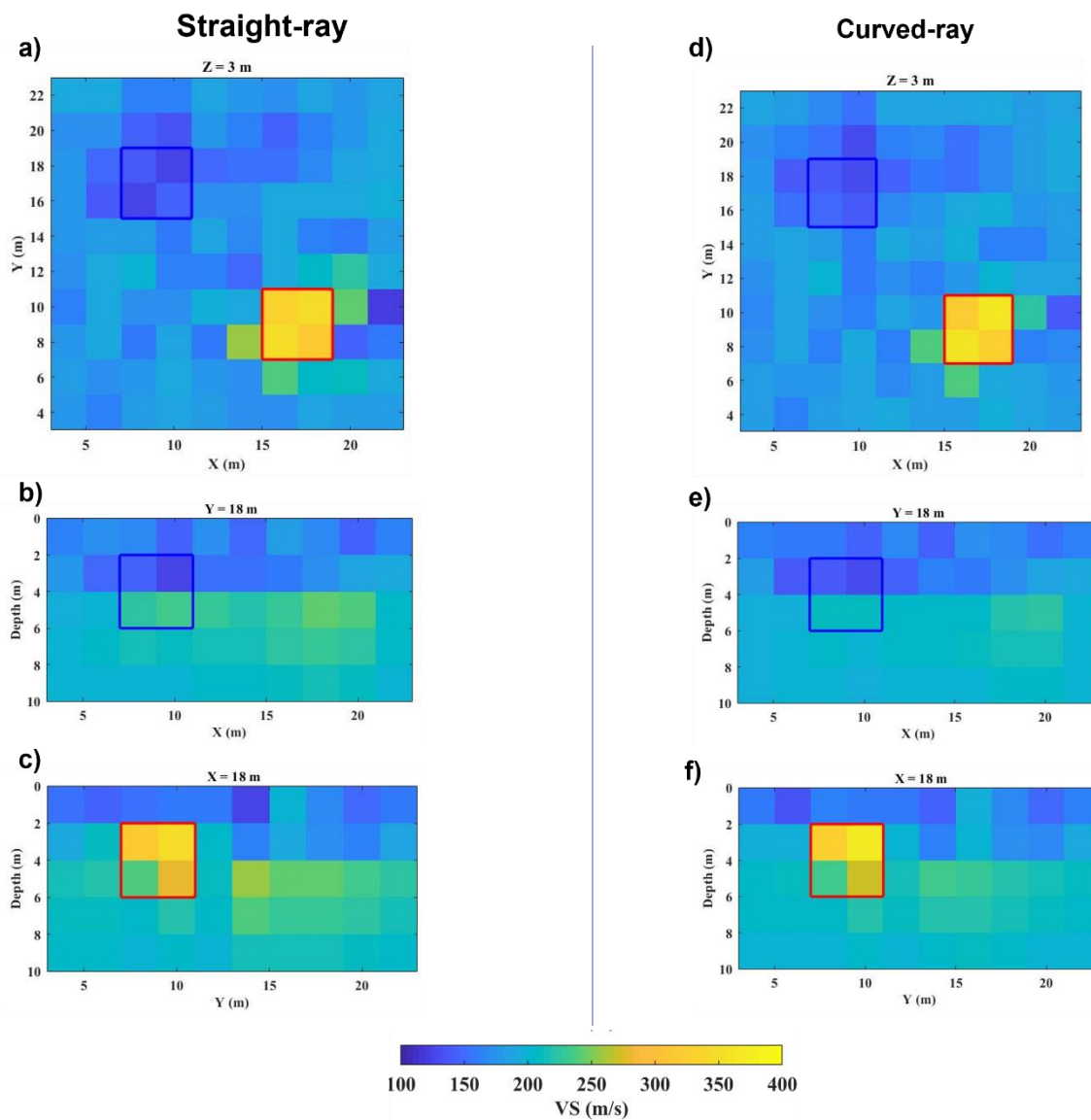

Figure 1. The retrieved VS models for the Blocky model with lowered data coverage from: (a-c) straight-ray SWT inversion, (d-f) curved-ray SWT inversion.

In this case, the curved-ray approach has produced a more accurate VS model than the straight-ray method. The obtained relative model error (compared to the true model) from the straight-ray and curved-ray methods are 12.4 % and 9.7 %, respectively.

Additionally, it would be interesting if the results in case study 4 could be made more similar by applying more smoothing in the straight ray example.

Following this suggestion, we have applied stronger smoothing in the straight-ray example. We decreased the values of C_Rp in the straight-ray approach from 10^6 to 100 and noticed that the VS model has become more similar to the curved-ray approach (with C_R values of 10^6). We have presented the results in the figure below:
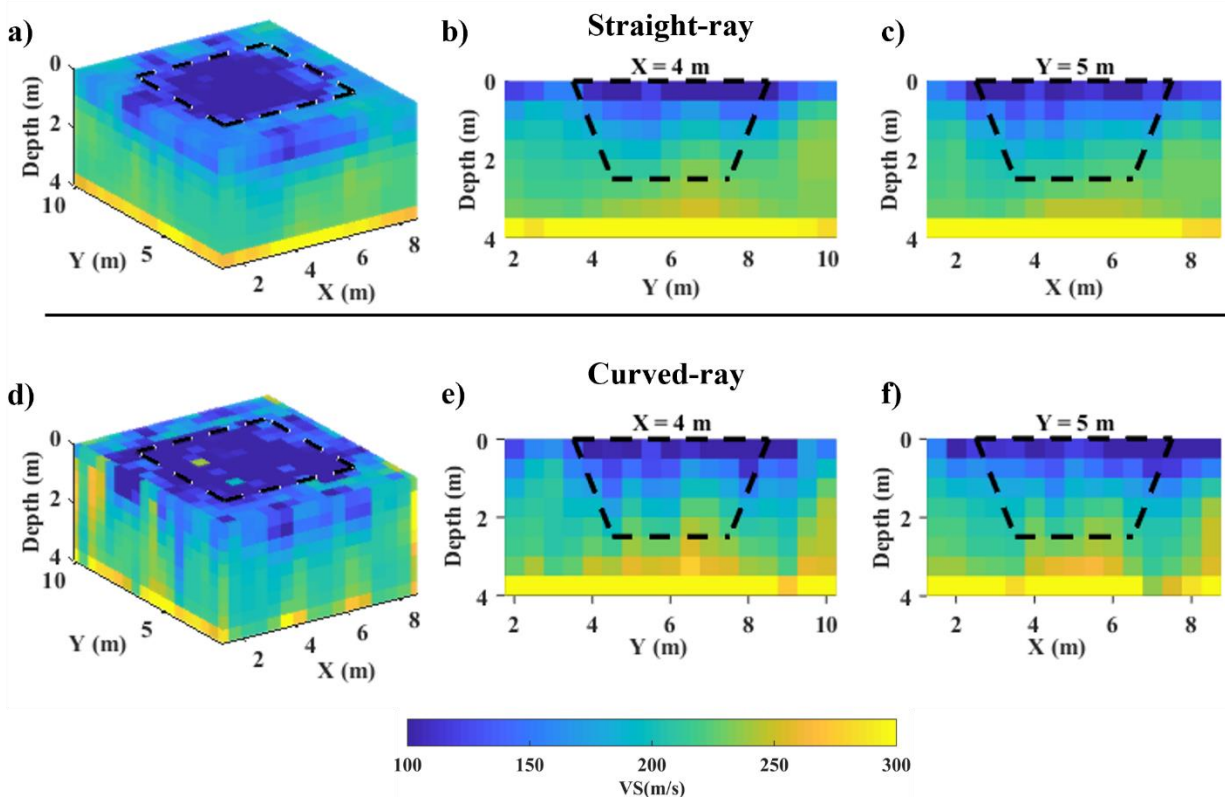

Figure 2. *The obtained VS models from: (a-c) straight-ray approach with C_Rp = 100, (d-f) curved-ray approach with C_Rp = 10^6*

In the matrix in eq. 12, I think the last element on the trace should not have index j=1?

True. We have corrected the last subscripts of the last diagonal element from "N,1" to "N,j".

line 208: I think the parameter "h" appears here for the first time and should be explained.

We have explained the parameter "h" in line 202 of the "manuscript_R2" file.

Table 4: The unit of "Cost" should be time instead of Dollar?

Responses to the comments of Reviewer#3

Dear Editor,

Dear Authors,

I found the manuscript of Karimpour et al. much improved, and I am mostly satisfied by the applied corrections. I still believe that the manuscript needs some polishing though, and I suggest that the authors spend one last effort in improving the style of writing and the quality of the English. (And double check for typos.) Besides this, I identified several points, reported below, that I would be happy to see addressed in the revised version of manuscript. Having done so, I will be happy to accept the manuscript for publication.

Sincerely,
Fabrizio Magrini

Notation in mathematical formulas:
I found the use of "DC" (eqs. (7), (8)) and "fw" (eqs. (8), (9), (11)) quite unconventional, i.e. closer to what one would normally write in a computer program rather than in a mathematical formula. I suggest that the authors replace these "symbols" with more conventional ones in their equations. For example, "DC" could be replaced with "c" and "fw(m)" with "g(m)". (Note that the symbol "g" already appears in equation (3), albeit never employed throughout the rest of the manuscript. For this reason, and because equation (3) can be easily and fully explained using plain English, this equation could simply be removed.)

Following this suggestion, we have replaced the symbol "DC" (in eqs. (7), (8) of the previous version) with "c", and "fw(m)" (eqs. (8), (9), (11) of the previous version) with "g(m)". We have also deleted Eq. (3) and explained it in the text.

Dispersion curve vs phase-velocity profile:
The authors employ the term "dispersion curve" (or "DC") to refer to two quite different "objects": (i) the measured -- experimentally or through ray theory -- dispersion curve (which can be seen as the average inter-station phase velocity), and (ii) the 1-D phase-velocity profiles, computed forwardly and associated with each location identified in the xy plane of a 3-D model. To avoid confusion (e.g., paragraph 130), and consistent with many other works (e.g., Magrini et al. 2022), I suggest that the authors use two different terms for the two circumstances, e.g., (i) "inter-station dispersion curves", or simply "dispersion curves", and (ii) "1-D phase-velocity profiles" or "phase-velocity profiles".

To avoid confusion, we have used the term "local DC" for the computed forward model at the position of the model points. We think that using the term "local DC" (instead of "phase velocity profiles") would clarify the concept for the readers and keeps the manuscript consistent with the majority of the surface waves related works in the literature.

We have applied this modification in lines 115-116 of "manuscript_R2" file as:

"For each model point, the **local DC** is computed using a Haskell (1953) and Thomson (1950) forward model modified by Dunkin (1965)."

Paragraph 130:
- Is Vp treated as an unknown? Please clarify

Vp (or equivalently Poisson's ratio) is assumed to be known. We had mentioned it in Section 2.3 of the previous version as:

"We carry out our experiments in a Cartesian coordinate system. The subsurface is discretized into a set of 3D grid blocks where it is assumed that the only unknown parameter of each grid block is the VS value while Poisson ratio (v) and density (ρ) are assumed to be known as a priori information.".

- Equation (3) is not very informative: it is simply stating that the kth phase-velocity profile is computed forwardly from the k_th Earth's model parameter. Moreover, the symbol "g" is not employed throughout the rest of the manuscript. I suggest that this equation is removed.

Following this suggestion, we have removed this equation in the new version of the manuscript.

Paragraph 140:
- In response to a comment I made in my first revision, the authors produced a figure of a traced ray (presented in their rebuttal) and wrote in the manuscript "To evaluate the accuracy of the ray-tracing algorithm, we have applied it in a to a homogeneous media and noticed that the error (i.e., deviation from straight-line) in this condition is almost zero (not shown here)." However, I did not find this test satisfactory. The fact that the traced arrivals of a wavefront propagating from the source towards azimuths of ~90°/180°/270°/360° is zero or close to zero is a well known property of ray-tracing algorithms such as fast-marching and fast-sweeping (see, e.g., Fig. 3 in White et al. 2020). What I was suggesting last time is that the authors produce a map of the relative errors between the theoretically predicted arrival times (which are easily computed from a homogeneous medium) and those computed through their ray-tracing algorithm. Something similar to Fig. 3 of White et al. 2020.
- Depending on the result of this test, it might be worth showing the 2-D map of relative errors and spending a few words about it

As suggested, we have provided the map of the relative errors and have shown the results in the figure below:
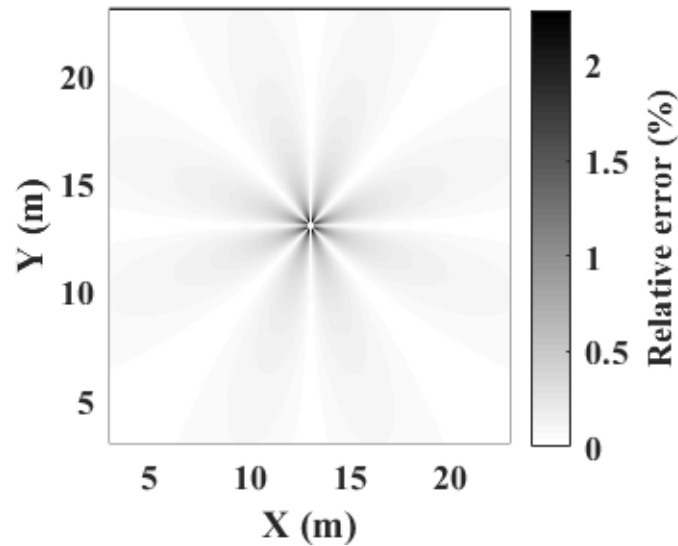
Figure 3. The computed errors of traveltimes for a source at (X, Y) = (13 m, 13 m)

We have not included this figure in the manuscript because this figure would not be very informative for the readers, and it could be a deviation from the main topic of the manuscript. More importantly, this figure would not add anything new since we have used the work of Noble et al. (2014) in which they have already discussed the traveltimes errors in detail (e.g., Fig.6 of Noble et al., 2014).

Anyhow, we have added the following explanations regarding the possible ray-tracing error in line 121-122 of the "manuscript_R2" file as:

"The accuracy of the employed ray-tracing algorithm has been already discussed by Noble et al. (2014)."

Equation (5):
Is the distance travelled by the wavefront from a source to a receiver calculated through an integral, or, as I suppose, in discrete form through a sum over the lengths of the several segments constituting each path? If you calculate it in discrete form, please rewrite the equation accordingly and/or be explicit about this in the text.

The inter-station paths are discretised. We have rewritten the equation and replaced the integral with the summation. Moreover, we have clarified in the line before the equation that the path is discretised.

"The path-average phase slowness along the **discretised** path for each frequency ($p_{R_1R_2}\left(f\right)$) is …".

Equation (6):
This is probably simply trivial, I suggest that this equation is removed

We have deleted this equation in the newest version of the manuscript.

Paragraph 160: I found this paragraph unclear.
- "The vector of the forward response of the model fw(m)", seems to be a matrix, based on eq. (8)

The forward response is a vector not a matrix. As shown in eq. (8) of the previous version of the manuscript, fw(m) consists of several vectors (DC_i) that are concatenated vertically forming a vector.

- The authors write: "It should be noted that each estimated DC may have a frequency band different from the others and therefore, the lengths of DCs are not necessarily the same", but I am not sure what this implies. Does this mean that the rows of fw(m), each corresponding to a different frequency, have different dimensions depending on the considered inter-station pair? If so, do the authors carry out some sort of interpolation to homogenize the dimensionality of d_obs and fw(m)? Please clarify in the text and eventually rephrase.

Based on eq. (7) of the previous version, the vectors of inter-station phase velocities (DC_i) are computed and then they are put together in a vector to form the forward response of the model (fw(m) in eq. (8) of the previous version). The dimension of fw(m) and d_obs are essentially the same and no interpolation is needed to be done. We have noticed that the explanations about the different frequency band of each DC might have been confusing to the readers. We have modified the text accordingly to make it more straightforward and understandable as:

"It should be noted that each estimated DC may have a frequency band different from the others. The vector of the forward response of the model (g(m)), is then obtained as …".

Reference

Noble, M., Gesret, A., and Belayouni, N.: Accurate 3-D finite difference computation of traveltimes in strongly heterogeneous media, Geophys. J. Int., 199, 422 1572-1585, https://doi.org/10.1093/gji/ggu358, 2014.

REFERENCES
- Magrini et al. 2022. Surface-Wave Tomography of the Central-Western Mediterranean: New Insights Into the Liguro-Provençal and Tyrrhenian Basins. Journal of Geophysical Research: Solid Earth 127.3 (2022): e2021JB023267.
- White et al. 2020. PyKonal: a Python package for solving the eikonal equation in spherical and Cartesian coordinates using the fast marching method. Seismological Research Letters 91.4 (2020): 2378-2389.