REVIEW REPORT FOR

Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector

published in *Seismica*

ROUND 1 REVIEWS

REVIEWER 1

Dear Authors,
I have completed the review of your manuscript "Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector".

The study is timely and of current interest, as it adds to a series of works aimed at the characterization of Sos Enattos, the Italian candidate site for the Einstein Telescope. In particular, it focuses on the analysis of the seismic noise in the frequency range 10-20 Hz. Although the applied techniques *per se* are not a novelty, the case study is needed to complete the picture of seismic noise at the site of interest and adds to the previous knowledge available in literature. Moreover, an interesting comparison is provided between arrays using different sensor number and configuration (concentric versus radial).

Overall, the paper is robust in its theoretical framework and clear in describing the applied techniques and presenting the results. The title is adequate, and the figures are helpful for the article's understanding. However, I believe that some details regarding the experimental setup are missing and that the structure of the manuscript is not optimal, especially lacking a clear initial description of the knowledge about the site already gathered through previous studies (especially regarding the site geology). This knowledge is essential to justify the choices made regarding the arrays configurations and the subsequent analyses, and would also help highlight the new contributions provided by this article to its context.

More specifically, I would appreciate if you could expand on the "Introduction" section, providing a fuller picture of the context in which the work is carried out, and on the and "Array description and data acquisition" section, adding some details that are needed to support your interpretation of the results (see the list of comments below) and to better motivate how you came to the choice of performing these analyses.

Comments:

- This work is strongly tied to the context for which it was ideated, but I did not find a clear and comprehensive description of such context. I suggest to expand the Introduction section to give a broader overview of the studies already published for the seismic characterization at the Sos Enattos site, some of which you already cite. Instead of a simple list of citations, I believe an overview of what analyses were already conducted and what specific issue of seismic characterization they tackled would provide a much clearer context for your manuscript, and help the reader understand some of your choices (for example, why do you focus on the 10-20 Hz range? Because other studies already covered the 1-10 Hz band, and the overall band of possible impact on the ET is that from 1 Hz and above; why do you focus on short-term characterization? Because previous studies already covered the long-term temporal variation of seismic noise; and so on). I understand this could also imply increasing the numbers of citations to works by the same authors, but in this case this would be due to the work belonging to a series of group efforts rather than to an attempt at self-citation.

- Consequently, the Abstract would be clearer if you included a sentence to inform the reader that this work is carried out in the context of many other works that already covered other aspects of the site characterization.
- Following a similar reasoning, I find that an additional section (however short) dedicated to describing the geological setting of the site would be needed to support your choices regarding for example the array configurations. In the "Methods and results" section you comment on your results in the context of the findings by Villani et al. 2025, but their (and possibly others') findings on the subsoil characterization should be more explicitly summarized also in the beginning. More specifically, the manuscript currently gives the impression that you ‘blindly’ selected an array configuration with a total aperture of 400 m and subsequently found that your results allow you to characterize the 10-20 Hz band and the upper 100-150 m of subsoil, but I believe it was the other way around: based on the previous knowledge on the subsoil properties (especially the geological and velocity profiles) and on the characterizations already available for other frequency bands, you decided to use a 400 m wide array configuration to fill in the gaps of the previous knowledge on site seismic noise. The explanation of the reasons that led you to choose some specific analyses to be carried out with some specific array configurations should be stated forefront and supported by an overview of the a priori knowledge you had when taking these decisions.
- For the same reason, I believe it would be useful to also cite the site selection criteria for the ET (Amann et al. 2020) and especially their recommendations regarding seismic noise characterization, which you likely used as guidelines in designing your experiment. More in general, since in this kind of studies the results are strongly dependent on the used sensors, it would be useful to provide a better description of your instrumentation, for example by showing the instrumental response curves of the sensors or their self noise curves (perhaps as Supplementary material).
- Additional details regarding the seismic arrays should be provided, to fully support the results:
    - Could you indicate the station codes in Figure 1, so that the faulty ones can be clearly identified? Could you also indicate which are the ones shown in the PPSD plots?

    - In the description of the P2-2024 configuration, please consider more clearly stating that you only used surface installations.
    - What is the sampling frequency of the recordings?
    - There is no indication on the time periods in which the recordings took place: in which season, in which part of the week (working days or not; this is especially important for the P2-2024 configuration which only covers 24 hours); were the recordings of P2 and P3 synchronous?
    - You correctly comment in the end that your configuration does not allow to assess for possible seasonal variations; this could be stated already in the "Array description" section.
- In Figure 1, could you improve the map to show also the location of the wind park and of the nearby towns? For example, the town of Bitti is quite near to site P2, and the town of Lodé to P3 (both roughly 6 km away), but this aspect is never commented on.
- About the PPSD analysis, at lines 89-90 you state that “Higher frequencies [above 3 Hz] … approach the Peterson’s NLNM”, but this is not true at even higher frequencies (starting at roughly 10 Hz the curves increase again). In the P2-2024 case, you motivate the overall higher noise levels with the stormy weather; do you think it could also partially be a consequence of the different array configuration (surface sensors only)?
- On the beamforming section:
    - Array synchronization is a key aspect in the beamforming analysis; how did you ensure/check that all array elements had high precision relative timing?
    - You mention the results of beamforming analysis performed on the horizontal components and even propose an interpretation, but you do not show them; these results (or at least a selection) could be provided as supplementary material.
    - I am not fully convinced of the explanation of the difference observed between P2 and P3 at 10-20 Hz. P3 is also quite near to a road that covers almost all the 270°-45° directions; is this a less trafficked road?
    - There is a typo at line 120, where the sector with no incoming signal should be the SE one (not SW).
    - The labels inside the radial plots of Figure 5 are not readable, please increase their font.
- On the HVSR section:
    - Overall, it appears your analysis was compliant to the SESAME guidelines, except maybe for the P2-2024 dataset which if I understood correctly was taken under strong winds condition. You already cite the guidelines, but I suggest you to also explicitly state the compliancy of your analysis to such guidelines, which are broadly used in the community.
    - I would however be careful in stating that, based only on HVSR results, there are “no major resonance effects” (line 174): surely the HVRS curves indicate no sharp contrasts at depth, but they do not guarantee no amplification whatsoever...if the sites are located on soil deposits: if you provide an initial

description of the geological condition at your sites, instead of only giving it at the end (lines 259-264), it would also help you support your HVSR interpretation here.
    - Where are the two stations with a peak in the HVSR curve (case P3) located? As above, you could add the stations codes to Figure 1. Are they near each other?
- On the f-k section:
    - you state you used ”12 hours of continuous recording in the 6:00-12:00 time interval” (line 179); which is it?
    - As for the difficulty in recovering the phase velocities for the P3 dataset, you state it is likely due to the low background noise levels at the site; however, the PPSD curves shown in Figure 3 do not seem to differ so much for the cases P3 and P2 (this is also evident in Figure 4). Have you considered the possible influence of a difference in the nature of the wavefield at the two sites? According to the beamforming results, the noise wavefield seems more homogeneous in P3, whereas in P2 is relatively more azimuthally constrained, which would favor the f-k methodology (e.g., https://doi.org/10.2312/GFZ.NMSOP-2_ch9 and references therein). Have you considered applying a different kind of analysis such as MSPAC, for which a random source distribution is most favorable, to the P3 data?
    - The legend of Figure 7 is incorrect: the description of panel b) is reported as panel c), and the description of actual panel c) is missing. Similarly, in text panel c) is referred as panel b) (line 197).
    - According to the sensitivity plot in Figure 7.c, it seems that for the range 10-20 Hz the data inversion would be informative of the fist 50-100 m maximum, more than 100-150 as stated.
- I strongly advise the Authors to already provide a repository, for example a Zenodo project, instead of stating that “seismic recordings will be made available”. You could embargo the dataset until the paper is published, so that in the meantime at least the metadata for the dataset is FAIRly accessible. Also, because of this, the analyses are not currently reproducible.
- References:
    - I would appreciate if you could add the DOIs in your References.
    - I believe you could find more direct sources for the claims in section 3.3. As for the one regarding wind gusts (lines 166-167), Mucciarelli and Gallipoli 2001 provide a good overwiev of the whole HVSR technique in microtremors, but the original sources for the case at hand should be the cited Mucciarelli paper (Mucciarelli 1998) and even better the SESAME guidelines, which report precisely the warning of not recording data under strong wind conditions (> 5 m/s). As for the claim regarding the Rayleigh-Love partitioning depending on noise source variability (line 167-168), I do not find any correspondence with the paper you cited (Xiao and Wang 2022), which analyzes data from Mars and contains no similar claims, at least to my

understanding. Please explain your choice of using this reference or provide a different one.
- CMDS and BayesBay could be cited in the "Data code and availability" section, instead of in-text.

Once these issues have been dealt with, I believe your manuscript has the potential to be of interest and usefulness in the context of the site characterization related to the Einstein Telescope project, and also to provide some broader insights to the field of local seismic noise characterization.

I hope I have not missed some detail or misinterpreted your views and I wish you good work for the revision.

REVIEWER 2

Dear authors, dear editor,
Please find my comments in the attached review report.

Kind regards

(See annotated 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

Dear Authors,
I have completed the review of the latest version of your manuscript “Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector”.

I appreciate your work in reviewing the manuscript according to both mine and the other Reviewer’s comments. I particularly appreciate your care in providing the context for your work, both in terms of previous experiments and of the site geology (although maybe you could consider slightly simplifying that section). I believe you work is now improved in terms of clarity and strength.

I only have a few more minor comments that should be easy to address:

- General comments:
    - L136-137: you state that the configuration of the spiral array “allows to push the limit of non-aliased frequency up to 100 Hz”. This consideration is only true strictly speaking in terms of array geometry, but it should be remarked

(although it may seem trivial) that the sampling frequency also affects the total non-aliased frequency band of the recorded signal. In your case that would lower the maximum frequency to 50 Hz.

  - Tables 2, 3 and 4: could you please check all of the values reported in these tables? It seems to me like you forgot to update some of them (e.g., the reported wavelength/aperture ratio is the same even if D is different; in table 3 it may be due to rounding (and I would suggest adding a second decimal), but in tables 2 and 4 the values just seem wrong for the provided v, f and D values)
  - L292-293: please provide (either here or in the supplementary material) some reference for the methodology used to calculate coherency
  - L295: here you talk about "extremely low level of ambient noise at all frequencies". However, as you also state in your reply, you interpret it as a low level of *seismic background* noise, not of *ambient* noise overall (which would include e.g. the wind); you actually state many times that the discrepancy in noise comes from weather conditions. This line should be rephrased, or you should specifically refer to seismic background noise instead of ambient noise.
  - Data and Code Availability section: I appreciate that the datasets are now publicly available at EIDA. Could you add a more direct link to the webservice where they are available, in addition to citing the paper describing the node?

- Typos:
  - L106: "18-28/09-2021" should be "18-28/09**/**2021"
  - L112: please check, I believe there should be a comma where there is now a full stop
  - L130: "of each of the array**s**"
  - L295: "at all frequency" should be "at all frequenc**ies**"

- Figures:
  - Fig.8: please increase the font size of all subplots (axes, ticks and legend labels) which are now difficult to read
  - Fig.8: the legend still reports the sentence "Dashed lines indicate the limit of the resolved region." which is now not applicable

I wish you good work for the revision.

<u>Answers</u> to these reviews are inserted below, following this compilation of text comments.

Title:	Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector
Authors:	*Giovanni Diaferia, Marco Olivieri, Irene Molinari, Annalisa Allocca, Enrico Calloni, Giovanni Luca Cardello, Andrea Contu, Domenico D'Urso, Rosario De Rosa, Matteo Di Giovanni, Luciano Errico, Luca Naticchioni, Davide Rozza, Lucia Trozzo, Carlo Giunchi*
Journal:	SEISMICA
Manuscript ID:	*Not provided*

REVIEW REPORT

The manuscript presents the results of three seismic array campaigns conducted at the Sos Enattos site, the proposed Italian candidate location for the Einstein Telescope (ET). The stated objective is to characterize the local seismic noise field and to derive shallow subsurface information through beamforming, f–k analysis, and HVSR measurements. The dataset is valuable, and the topic is definitely relevant for the ET community, since the assessment of both seismic and geological site conditions is a cornerstone of the site-selection process. The data acquisition is clearly described, the graphical material is well prepared, and several results, particularly the overall consistency between PPSD, HVSR, and dispersion curves, are promising. Nevertheless, a number of methodological, interpretative, and contextual issues limit the robustness of the study in its current form, and merit careful revision.

The article opens with a description of the study site and the motivations associated with the ET candidature, but it lacks a fundamental component: a geological and geophysical (e.g., seismotectonic) overview of the area. The interpretation of dispersion curves, flat HVSR spectra, high velocities, and the apparent absence of impedance contrasts requires explicit reference to the local geological structure, including lithology, degree of fracturing, weathering profiles, presence or absence of superficial cover, and any previous surveys (including those related to the mine). Without such context, many of the interpretations remain unsupported and difficult to evaluate. Studies of this kind benefit greatly from a concise geological framework that allows the reader to assess whether the observed seismic response is consistent with existing knowledge of the subsurface or whether alternative explanations should be considered.

The section describing the array configurations is clear regarding geometry, aperture, and instrumentation. However, the choice of adopting three markedly different configurations—two concentric arrays with 20–24 stations operating for 10 days each, and one spiral array with 15 stations operating for only 24 hours—is insufficiently justified. In the absence of a discussion explaining why these different geometries were selected and how they affect the resolving capability of the arrays (in terms of aliasing, sidelobes, kmax, and azimuthal coverage), direct comparisons between them remain ambiguous. A quantitative (clearly analytical) evaluation of resolution limits for each configuration would be especially valuable, given the differences observed in the dispersion and beamforming results.

Regarding the spectral analysis, PPSD results are stated in the text to be performed for only three "randomly selected" stations per array. Although the method is correctly applied, such a limited sample does not allow for a complete assessment of data quality nor for the systematic identification of anomalous sensors. Since PPSD computation is automated and computationally light, it would be more informative to present either the full set of PPSDs or a statistical summary covering all stations. This would strengthen the evaluation of noise levels and their spatial consistency.

In the beamforming section, the analysis is performed on very broad frequency bands (1–10, 10–20, and 20–40 Hz). While this may lead to smoother graphical results, it is not methodologically appropriate. Beamforming is inherently a narrowband technique, and applying it over such wide intervals inevitably produces smearing in both slowness and azimuth, mixing contributions from different wave types and severely reducing the ability to track dispersive behavior. This is particularly problematic when discussing possible Love waves or higher-mode Rayleigh waves (as it done later). The lack of coherent directional patterns in the 20–40 Hz band cannot be taken as an indicator of "incoherent noise" when, at such frequencies, arrays with apertures on the order of 400 m lie entirely outside their resolvable domain. Such an interpretation requires comparison with theoretical resolution bounds, which the manuscript does not provide. Explicit plots of resolvable wavenumber versus frequency (kmax–f or λ–f curves) would be essential to distinguish between true absence of coherence and theoretical impossibility of resolving the signal.

The interpretation of Love waves, suggested on the basis of slightly higher apparent velocities on the horizontal components, is also not convincingly supported. At a site where HVSR curves are flat and devoid of significant peaks—thus lacking strong impedance contrasts—the efficient generation of Love waves is unlikely. Moreover, the manuscript does not include any dedicated evidence (horizontal-component beamforming, polarization analysis, or directional energy partitioning) to substantiate this claim. A similar issue arises with the mention of (Rayleigh?) higher modes. The figures do not show clear modal branches, and the observed patterns are compatible with noise or sidelobe artefacts rather than true higher-mode dispersion. These interpretations should therefore be reconsidered more cautiously.

A further point that would benefit from clarification concerns the treatment of the horizontal components in the beamforming (and in the f–k analyses, if this was also the case). The manuscript does not specify whether the authors applied the appropriate directional projection of each station's horizontal recordings onto the trial slowness vectors, which is required when estimating coherent horizontal wavefields such as Love waves. Without this projection, the horizontal-component beamforming may fail to enhance the true directionally coherent energy and could instead mix azimuthally inconsistent contributions. Clarifying how the horizontal components were combined would therefore strengthen the interpretation of the results and help assess the robustness of the inferred wave types. Please provide some clarification.

The HVSR analysis is consistent with a high-velocity, non-resonant shallow structure, and the acquisition is well executed. However, the choice to plot HVSR amplitudes using a linear vertical axis is statistically very inappropriate. HVSR amplitudes form a strictly positive quantity that typically follows a log-normal distribution, i.e., a skewed distribution in which large values have disproportionate weight and the arithmetic mean does not represent the central tendency. Plotting HVSR in linear scale suppresses subtle but meaningful variations, which are instead clearly visible in logarithmic scale. This choice is particularly limiting in low-amplification rocky sites, where small deviations from unity may reveal velocity gradients or thin weathered layers. The manuscript would benefit from a revised representation using logarithmic amplitude scaling, which more faithfully reflects the statistical nature of the HVSR estimator.

The f–k analysis yields dispersion curves that are consistent with the beamforming results (actually, Geopsy's f-k is made through beamforming): all configurations recover the fundamental mode between 10 and 20 Hz, whereas low-frequency information is much weaker or absent, especially for the P3 array. This discrepancy is acknowledged but not adequately explained. Simply attributing it to "low noise amplitude" is plausible but not sufficiently substantiated. No measure of interstation coherence is provided, which would be extremely helpful to discriminate between array limitations and characteristics of the noise field. In addition, the strong slowness peaks observed from ENE at P2-2024 likely reflect wind conditions and local vegetation, but also in this case some quantification would strengthen the argument.

The inversion of the dispersion curves employs a modern trans-dimensional MCMC approach, which is appropriate in principle. However, the presentation of the results is not statistically sound. The ensemble of models produced by the sampler includes layers with variable depth; averaging these models point-wise does not yield a physically meaningful velocity profile, but rather a smoothed artefact unrelated to any realizable structure. A more rigorous Bayesian summary should present median velocities at fixed depths, credible intervals, and posterior distributions of interface depths. Although the comparison with borehole Vp profiles is a valuable addition, it should be more carefully limited to the depth range actually resolved by surface waves.

Overall, the manuscript contains valuable material and the dataset itself is of high interest. However, several key aspects require significant revision. The impression is that the work can be substantially improved, but requires a careful and structured revision to ensure that the conclusions rest on solid ground. I therefore suggest major revision at this stage.

Kind regards.

Dear Editor,

We are pleased to submit the revised version of our manuscript. We have carefully addressed all points raised by the reviewers and we are convinced that the article has been substantially improved in terms of clarity, quantitative substantiation of the results, and their overall impact.

Here is a summary of the main revisions made to the manuscript:

- A new section has been added after the Introduction to describe the geological and geophysical setting of the study area. This section provides the necessary context for the geophysical results presented later in the manuscript
- The theoretical capabilities and limitations of the arrays, previously discussed only qualitatively, are now quantitatively evaluated and summarized in three tables, one for each reference frequency (3, 10, and 20 Hz). Parameters such as the ratio between dominant wavelength and array aperture, the maximum resolvable frequency, and the minimum and maximum wavenumbers are reported. These quantities provide a quantitative assessment of the resolving power of each array and complement the interpretation of the array transfer functions.
- As suggested by the rev#1r, the Supplementary Material now includes the results of a narrow-band (1 Hz) beamforming analysis, over the entire 1–19 Hz frequency range. This allows a more detailed evaluation of possible variations in slowness and backazimuth with frequency, or the absence thereof.
- Since we decide to solely rely on analysis of the vertical component (with the exception of the HVSR analysis), all references to Love waves have been removed. Given the near-homogeneous geological setting, the occurrence of Love waves is expected to be unlikely. Moreover, a rigorous assessment of a potential Love wave contribution is beyond the scope of this study and would require a dedicated investigation into seismic energy partitioning including body waves and surface waves.

We attach the revised manuscript, a version with tracked changes, and the Supplementary Material. It follows the rebuttal letter to the reviewers' comments.

We thank you for your time and consideration.

All the best,
Giovanni Diaferia (corresponding author)

REBUTTAL LETTER

REV#1

The manuscript presents the results of three seismic array campaigns conducted at the Sos Enattos site, the proposed Italian candidate location for the Einstein Telescope (ET). The stated objective is to characterize the local seismic noise field and to derive shallow subsurface information through beamforming, f–k analysis, and HVSR measurements. The dataset is valuable, and the topic is definitely relevant for the ET community, since the assessment of both seismic and geological site conditions is a cornerstone of the site-selection process. The data acquisition is clearly described, the graphical material is well prepared, and several results, particularly the overall consistency between PPSD, HVSR, and dispersion curves, are promising. Nevertheless, a number of methodological, interpretative, and contextual issues limit the robustness of the study in its current form, and merit careful revision.

1 - The article opens with a description of the study site and the motivations associated with the ET candidature, but it lacks a fundamental component: a geological and geophysical (e.g., seismotectonic) overview of the area. The interpretation of dispersion curves, flat HVSR spectra, high velocities, and the apparent absence of impedance contrasts requires explicit reference to the local geological structure, including lithology, degree of fracturing, weathering profiles, presence or absence of superficial cover, and any previous surveys (including those related to the mine). Without such context, many of the interpretations remain unsupported and difficult to evaluate. Studies of this kind benefit greatly from a concise geological framework that allows the reader to assess whether the observed seismic response is consistent with existing knowledge of the subsurface or whether alternative explanations should be considered.

We agree with the suggestion from the reviewer. After the Introduction, we added a brief but comprehensive section dedicated to the geological and geophysical context of the study area. We expect this would also help the reader to better understand the geophysical results provided in the manuscript.

2 - The section describing the array configurations is clear regarding geometry, aperture, and instrumentation. However, the choice of adopting three markedly different configurations—two concentric arrays with 20–24 stations operating for 10 days each, and one spiral array with 15 stations operating for only 24 hours—is insufficiently justified. In the absence of a discussion explaining why these different geometries were selected and how they affect the resolving capability of the arrays (in terms of aliasing, sidelobes, kmax, and azimuthal coverage), direct comparisons between them remain ambiguous. A quantitative (clearly analytical) evaluation of resolution limits for each configuration would be especially valuable, given the differences observed in the dispersion and beamforming results.

We thank the reviewer for pointing this out. At the time of the three acquisition campaigns the noise characteristics of the area were poorly known and there were tight logistic/permits limitations on the accessibility of the area. For these reasons, we opted for those specific array layouts, with rather limited aperture. We now better clarify this in the manuscript. However, the lack of a quantitative justification in terms of resolution of the arrays and overall performance did not allow a clear understanding of the limitation of the data acquisition and in the obtained results. To tackle this point, we now provide three different tables (one for each frequency, 3-10-20 Hz, coupled with a different propagation velocity of the plane wave) where the ratio wavelength/aperture, maximum resolvable resolution, min/max resolvable wavenumber are indicated. By referring to these tables, the section has been expanded, and the three arrays are now better presented in terms of their theoretical capabilities and limitations. We now point out that the similarity of the arrays in their resolution capabilities in the 10-20 Hz range allows to disentangle the role of the array characteristics from the site condition and noise source distribution in the similarity and differences of each arrays into the different analysis (e.g. beamforming, HVSR, estimation of the surface wave dispersion curve)

3 - Regarding the spectral analysis, PPSD results are stated in the text to be performed for only three “randomly selected” stations per array. Although the method is correctly applied, such a limited sample does not allow for a complete assessment of data quality nor for the systematic identification of anomalous sensors. Since PPSD computation is automated and computationally light, it would be more informative to present either the full set of PPSDs or a statistical summary covering all stations. This would strengthen the evaluation of noise levels and their spatial consistency.
The reviewer suggests a summary which is given in Fig. 4, where we provide the median of the PPSD for all the stations in the three arrays. These allowed us to assess the quality and consistency of acquired data (which is excellent for the P3 and P2 (2024) arrays), in addition to the possible presence of faulty sensors (which is the case for two sensors at P2).

4 - In the beamforming section, the analysis is performed on very broad frequency bands (1–10, 10–20, and 20–40 Hz). While this may lead to smoother graphical results, it is not methodologically appropriate. Beamforming is inherently a narrowband technique, and applying it over such wide intervals inevitably produces smearing in both slowness and azimuth, mixing contributions from different wave types and severely reducing the ability to track dispersive behavior. This is particularly problematic when discussing possible Love waves or higher-mode Rayleigh waves (as it done later).
We thank the reviewer for this suggestion, as it allows us to better explain the rationale behind the presentation of the beamforming results. We are aware that such a technique is narrow-band but we noted that, due to the poor dispersion behavior of surface waves in the area (as a consequence of the almost homogeneous velocity profile of the site), the polar plots do not show appreciable changes as a function of frequency in the 10-20 Hz range, with only minimal changes of slowness and direction of arrival across different frequencies. Thus, in order to avoid an excessive number of plots in the manuscript for each narrow-band and for each of the three arrays, we opted for a more compact plotting format without sacrificing the results and their interpretations. However, in order to follow the reviewer's suggestions and also support our claim, we now provide (in the Supplementary Material) the beamforming results in the entire 1-20 Hz band, in steps of 1 Hz for P2, P3 and P2 (2024).

5 -The lack of coherent directional patterns in the 20–40 Hz band cannot be taken as an indicator of “incoherent noise” when, at such frequencies, arrays with apertures on the order of 400 m lie entirely outside their resolvable domain. Such an interpretation requires comparison with theoretical resolution bounds, which the manuscript does not provide. Explicit plots of resolvable wavenumber versus frequency (kmax–f or λ–f curves) would be

essential to distinguish between true absence of coherence and theoretical impossibility of resolving the signal.

The reviewer correctly points out that our claim about the incoherent noise above 20 Hz is not fully supported by the observations and requires some clarification. In order to avoid aliasing, the stations in an array should be close enough to correctly sample, spatially, a given wavelength (at least two points per wavelength). In Table 2-3-4 we now provide the maximum non-aliased frequency for each of the array, proving that the array P2 and P3, due their excessive interstation distance, could not correctly reconstruct the seismic wavefield above 20-30 Hz. However, as the reviewer points out, it remains uncertain whether the beamforming results in the 20-40 Hz are due the inadequacy of the array configuration or due to the presence of incoherent noise. To solve such ambiguity, we can consider the array P2 (2024) and its "expanding spiral" configuration, that has interstation distances as short as 9 m and implies (theoretically, see Table 4) a maximum non-aliased frequency well above 40 Hz. The fact that the beamforming results from this array are similar to those from P2 and P3 can solve the aliasing vs. incoherent noise ambiguity and prove that the seismic noise field is mostly incoherent, implying that no slowness, nor direction of arrival, can be thus estimated regardless the resolving capabilities of the considered array.

We now provide this explanation in text.

6 - The interpretation of Love waves, suggested on the basis of slightly higher apparent velocities on the horizontal components, is also not convincingly supported. At a site where HVSR curves are flat and devoid of significant peaks—thus lacking strong impedance contrasts—the efficient generation of Love waves is unlikely. Moreover, the manuscript does not include any dedicated evidence (horizontal-component beamforming, polarization analysis, or directional energy partitioning) to substantiate this claim. A similar issue arises with the mention of (Rayleigh?) higher modes. The figures do not show clear modal branches, and the observed patterns are compatible with noise or sidelobe artefacts rather than true higher-mode dispersion. These interpretations should therefore be reconsidered more cautiously.

8 - A further point that would benefit from clarification concerns the treatment of the horizontal components in the beamforming (and in the f–k analyses, if this was also the case). The manuscript does not specify whether the authors applied the appropriate directional projection of each station's horizontal recordings onto the trial slowness vectors, which is required when estimating coherent horizontal wavefields such as Love waves. Without this projection, the horizontal-component beamforming may fail to enhance the true directionally coherent energy and could instead mix azimuthally inconsistent contributions. Clarifying how the horizontal components were combined would therefore strengthen the interpretation of the results and help assess the robustness of the inferred wave types. Please provide some clarification.

We thank the reviewer for raising this point.

As the reviewer correctly suggests, Love wave generation is probably unlikely in such a peculiar geological context (almost homogenous half-space, no shallow resonant layer). As for the application of beamforming on Love waves, it would introduce a high degree of complexity and its success would strongly rely on a very accurate orientation of the stations' horizontal components, something we cannot be completely sure of at this stage.

For the above reasons, we decide now to solely rely on the stations' vertical component and remove any reference and/or interpretation regarding Love waves. We believe that both results and interpretation presented in the paper remain valid even in the case only Rayleigh waves are considered. A better, more in depth study involving the characterization of Love waves can be a matter for a future study dedicated on wave energy partitioning and overall contribution of body waves vs. surface waves in the noise budget.

9 - The HVSR analysis is consistent with a high-velocity, non-resonant shallow structure, and the acquisition is well executed. However, the choice to plot HVSR amplitudes using a linear vertical axis is statistically very inappropriate. HVSR amplitudes form a strictly positive quantity that typically follows a log-normal distribution, i.e., a skewed distribution in which large values have disproportionate weight and the arithmetic mean does not represent the central tendency. Plotting HVSR in linear scale suppresses subtle but meaningful variations, which are instead clearly visible in logarithmic scale. This choice is particularly limiting in low-amplification rocky sites, where small deviations from unity may reveal velocity gradients or thin weathered layers. The manuscript would benefit from a revised representation using logarithmic amplitude scaling, which more faithfully reflects the statistical nature of the HVSR estimator.
We accept the reviewer's suggestion and now the HVSR results are all presented using logarithmic scale on the y-axis.

10 - The f–k analysis yields dispersion curves that are consistent with the beamforming results (actually, Geopsy's f-k is made through beamforming): all configurations recover the fundamental mode between 10 and 20 Hz, whereas low-frequency information is much weaker or absent, especially for the P3 array. This discrepancy is acknowledged but not adequately explained. Simply attributing it to "low noise amplitude" is plausible but not sufficiently substantiated. No measure of interstation coherence is provided, which would be extremely helpful to discriminate between array limitations and characteristics of the noise field. In addition, the strong slowness peaks observed from ENE at P2-2024 likely reflect wind conditions and local vegetation, but also in this case some quantification would strengthen the argument.
We understand that the difficulty in the extraction of the phase-velocity dispersion curve at P3 is not well substantiated. Following the reviewer's suggestion, we now provide a dedicated plot (in the Supplementary Material) that shows the signal coherence across all possible station pairs at the three arrays. It shows that coherence drops very steeply for the P3 and P2 (2024) arrays, implying a strong contamination by high-frequency, uncorrelated noise at these sites. Such characteristics, in addition to the very low level of the seismic background noise at P3 (as seen in the PPSD), might explain the difficulty in extracting the phase velocity dispersion curve through fk-analysis. On the contrary, the higher signal coherence measured at P2 (likely due to a combination of higher interstation distance and better weather condition) justify the easier and less ambiguous retrieval of the dispersion curve, also in a wider frequency range (below 10 Hz).
Regarding the strong noise source imaged at P2 (2024), one can refer to the work from Diafera at al. (2024) regarding the characterization of the seismic noise produced by windpark in the vicinity of the ET candidate site. In this work, the spectrograms show that, despite the high amplitude of the seismic noise emission, seismic signal in the 10-20 Hz range quickly vanishes within 3-4 km. Therefore, we can expect that the noise source imaged at ENE at P2 (2024) should be fairly close (<<3 km) and, due to the lack of anthropic activities in the area, the only viable explanation could be found in the local vegetation coverage as source of such high-frequency noise emission.

11 - The inversion of the dispersion curves employs a modern trans-dimensional MCMC approach, which is appropriate in principle. However, the presentation of the results is not statistically sound. The ensemble of models produced by the sampler includes layers with variable depth; averaging these models point-wise does not yield a physically meaningful velocity profile, but rather a smoothed artefact unrelated to any realizable structure. A more rigorous Bayesian summary should present median velocities at fixed depths, credible intervals, and posterior distributions of interface depths. Although the comparison with borehole Vp profiles is a valuable addition, it should be more carefully limited to the depth range actually resolved by surface waves.

We agree that some details on the presentation of the MCMC inversion results were overlooked. To follow the reviewer suggestion, we have now divided Fig 7 in two different figures. Fig. 7 now has only i) the picked dispersion curves from the fk-analysis at the three arrays and ii) the sensitivity kernels of the Rayleigh wave at different frequencies. Fig. 8 now has, for each array, the ensemble solution of the MCMC inversion, the histogram representing the posterior probabilities of the layer interfaces, and the data fit in terms of 10-90th percentile.

We have also modified the plot with borehole Vp profile at P2 and P3, limiting the depth at the same extent of the ensemble solution from the MCMC inversion.

Overall, the manuscript contains valuable material and the dataset itself is of high interest. However, several key aspects require significant revision. The impression is that the work can be substantially improved, but requires a careful and structured revision to ensure that the conclusions rest on solid ground. I therefore suggest major revision at this stage.

#REV2

Dear Authors,

I have completed the review of your manuscript “Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector”.

The study is timely and of current interest, as it adds to a series of works aimed at the characterization of Sos Enattos, the Italian candidate site for the Einstein Telescope. In particular, it focuses on the analysis of the seismic noise in the frequency range 10-20 Hz. Although the applied techniques per se are not a novelty, the case study is needed to complete the picture of seismic noise at the site of interest and adds to the previous knowledge available in literature. Moreover, an interesting comparison is provided between arrays using different sensor number and configuration (concentric versus radial).

Overall, the paper is robust in its theoretical framework and clear in describing the applied techniques and presenting the results. The title is adequate, and the figures are helpful for the article’s understanding. However, I believe that some details regarding the experimental setup are missing and that the structure of the manuscript is not optimal, especially **lacking a clear initial description of the knowledge about the site already gathered through previous studies (especially regarding the site geology)**. This knowledge is essential to justify the choices made regarding the arrays configurations and the subsequent analyses, and would also help highlight the new contributions provided by this article to its context.

More specifically, I would appreciate if you could expand on the "Introduction" section, providing a fuller picture of the context in which the work is carried out, and on the and **"Array description and data acquisition" section, adding some details that are needed to support your interpretation of the results (see the list of comments below) and to better motivate how you came to the choice of performing these analyses.**
Following the suggestion from rev#1 and rev#2, we have inserted a section that provides an overview on the geological context of the study area. Moreover, in the "Array description and data acquisition" section we now provide quantitative details (see Tables 2-3-4) on the array characteristics and performance. We also better state the rationale behind the choice of such configurations.

Comments:

This work is strongly tied to the context for which it was ideated, but I did not find a clear and comprehensive description of such context. **I suggest to expand the Introduction section to give a broader overview of the studies already published** for the seismic characterization at the Sos Enattos site, some of which you already cite. Instead of a simple list of citations, I believe an overview of what analyses were already conducted and what specific issue of seismic characterization they tackled would provid**e a much clearer context for your manuscript, and help the reader understand some of your choices (for example, why do you focus on the 10-20 Hz range? Because other studies already covered the 1-10 Hz band, and the overall band of possible impact on the ET is that from 1 Hz and above; why do you focus on short-term characterization?** Because previous studies already covered the long-term temporal variation of seismic noise; and so on). I understand this could also imply increasing the numbers of citations to works by the same authors, but in this case this would be due to the work belonging to a series of group efforts rather than to an attempt at self-citation.
Consequently, the Abstract would be clearer if you included a sentence to inform the reader that this work is carried out in the context of many other works that already covered other aspects of the site characterization.
In the introduction we now briefly cite the content of previous works on the characterization of the candidate site. We also better explain the motivation behind this work (e.g. a work on the estimation of noise source distribution was lacking). The Abstract now contains the sentence suggested by the reviewer.

Following a similar reasoning, I find that an additional section (however short) dedicated to describing the geological setting of the site would be needed to support your choices regarding for example the array configurations. In the "Methods and results" section you comment on your results in the context of the findings by Villani et al. 2025, but their (and possibly others') findings on the subsoil characterization should be more explicitly summarized also in the beginning. M**ore specifically, the manuscript currently gives the impression that you 'blindly' selected an array configuration with a total aperture of 400 m and subsequently found that your results allow you to characterize the 10-20 Hz band and the upper 100-150 m of subsoil, but I believe it was the other way around: based on the previous knowledge on the subsoil properties (especially the geological**

and velocity profiles) and on the characterizations already available for other frequency bands, you decided to use a 400 m wide array configuration to fill in the gaps of the previous knowledge on site seismic noise. The explanation of the reasons that led you to choose some specific analyses to be carried out with some specific array configurations should be stated forefront and supported by an overview of the a priori knowledge you had when taking these decisions.
For the same reason, I believe it would be useful to also cite the site selection criteria for the ET (Amann et al. 2020) and especially their recommendations regarding seismic noise characterization, which you likely used as guidelines in designing your experiment. More in general, since in this kind of studies the results are strongly dependent on the used sensors, it would be useful to provide a better description of your instrumentation, for example **by showing the instrumental response curves of the sensors or their self noise curves (perhaps as Supplementary material).**
As also suggested by rev#1, we have now inserted a brief but comprehensive section regarding the geological context of the study area. Here we also refer to and summarize the results from previous geophysical campaigns in the area.
The reviewer correctly points out that the reasoning behind the choice of the 400 m aperture arrays was unclear. Now, we specify in the text that this choice was mainly dictated by logistical constraints and permits that, at the time of the investigation, impeded the installation of larger arrays. At the time of writing, we are planning the installation of a spiral, 2000 m aperture array in the vicinity of the P3 site, targeting the 1-10 Hz frequency range. Regarding the instrumentation, we now cite that all the employed instruments, according to the manufacturer's specification, have a flat response in the frequency range analyzed in this study and self-noise is near or below the Peterson's NLNM. Plots of self-noise and instrument response can be easily found at the sensor's manufacturer websites, which are now linked in the main text.

Additional details regarding the seismic arrays should be provided, to fully support the results:
Could you indicate the station codes in Figure 1, so that the faulty ones can be clearly identified? Could you also indicate which are the ones shown in the PPSD plots?
In the description of the P2-2024 configuration, please consider more clearly stating that you only used surface installations.
We have inserted all station codes in Fig 1 and in the PPSD plots. We also state that the P2 (2024) sensors are underneath 20-40 cm of soil to avoid wind disturbance and assure thermal stability.

What is the sampling frequency of the recordings?
There is no indication on the time periods in which the recordings took place: in which season, in which part of the week (working days or not; this is especially important for the P2-2024 configuration which only covers 24 hours); were the recordings of P2 and P3 synchronous?
We now state that all acquisitions are made with a 100 Hz sampling rate. The exact start- and end-time for each array are now given in the text. Acquisition at P2 and P3 were not synchronous.

You correctly comment in the end that your configuration does not allow to assess for possible seasonal variations; this could be stated already in the "Array description" section.
We believe this comment is better suited for the discussion.

In Figure 1, could you improve the map to show also the location of the wind park and of the nearby towns? For example, the town of Bitti is quite near to site P2, and the town of Lodé to P3 (both roughly 6 km away), but this aspect is never commented on.
Fig. 1 has been improved according to the reviewer's suggestion, showing i) the wind park in the NW part of the ET candidate site, ii) the main municipalities in the area.

About the PPSD analysis, at lines 89-90 you state that "Higher frequencies [above 3 Hz] … approach the Peterson's NLNM", but this is not true at even higher frequencies (starting at roughly 10 Hz the curves increase again). In the P2-2024 case, you motivate the overall higher noise levels with the stormy weather; do you think it could also partially be a consequence of the different array configuration (surface sensors only)?
We have now rephrased the sentence, stating that only frequencies in the 3-10 Hz range approach the NLNM. Stations at P2 (2024) were not installed in a different way compared to P2 and P3. In all three cases, sensors were buried in soil.

On the beamforming section:
Array synchronization is a key aspect in the beamforming analysis; how did you ensure/check that all array elements had high precision relative timing?
You mention the results of beamforming analysis performed on the horizontal components and even propose an interpretation, but you do not show them; these results (or at least a selection) could be provided as supplementary material.
All arrays had, beside sensors and digitizer, a GPS receiver that assures the accurate time-stamping of the recorded seismic signal. Following also the comment of rev#1, we decided to discard any interpretation based on the analysis of the horizontal component (and possible Love waves contribution). In fact, this would need a more in depth evaluation of seismic energy partitioning and, most importantly, very accurate orientation of the horizontal components, that we might not be completely sure of at this stage. We decide to solely rely on the analysis of the vertical component.

I am not fully convinced of the explanation of the difference observed between P2 and P3 at 10-20 Hz. P3 is also quite near to a road that covers almost all the 270°-45° directions; is this a less trafficked road?
There is a typo at line 120, where the sector with no incoming signal should be the SE one (not SW).
The labels inside the radial plots of Figure 5 are not readable, please increase their font.
Among the two locations, the P3 site is the most remote and the nearby road has very little traffic (in the order 1-2 vehicles per hour).
The typo has been corrected. Labels size in Fig. 5 have been enlarged.

On the HVSR section:
Overall, it appears your analysis was compliant to the SESAME guidelines, except maybe for the P2-2024 dataset which if I understood correctly was taken under strong winds condition. You already cite the guidelines, but I suggest you to also explicitly state the compliancy of your analysis to such guidelines, which are broadly used in the community.
I would however be careful in stating that, based only on HVSR results, there are "no major resonance effects" (line 174): surely the HVRS curves indicate no sharp contrasts at depth, but they do not guarantee no amplification whatsoever...if the sites are located on soil deposits: if you provide an initial description of the geological condition at your sites, instead of only giving it at the end (lines 259-264), it would also help you support your HVSR interpretation here.
Where are the two stations with a peak in the HVSR curve (case P3) located? As above, you could add the stations codes to Figure 1. Are they near each other?
We now state that our procedure for the extraction of the HVSR is compliant with the best practices (which are properly cited).
Since the geological context is now properly given at the beginning of the paper, we now refer to this to support the interpretation regarding the lack of a major impedance contrast, which is suggested by the retrieved HSRV spectra at the three arrays.
The two stations which show a pick in the HVSR spectrum are stations P308 and P317, which are not close to each other.

On the f-k section:
you state you used "12 hours of continuous recording in the 6:00-12:00 time interval" (line 179); which is it?
Corrected. The window length is 6 hours.

As for the difficulty in recovering the phase velocities for the P3 dataset, you state it is likely due to the low background noise levels at the site; however, the PPSD curves shown in Figure 3 do not seem to differ so much for the cases P3 and P2 (this is also evident in Figure 4). Have you considered the possible influence of a difference in the nature of the wavefield at the two sites? According to the beamforming results, the noise wavefield seems more homogeneous in P3, whereas in P2 is relatively more azimuthally constrained, which would favor the f-k methodology (e.g., https://doi.org/10.2312/GFZ.NMSOP-2_ch9 and references therein). Have you considered applying a different kind of analysis such as MSPAC, for which a random source distribution is most favorable, to the P3 data?
As the reviewer correctly points out, P2 and P3 arrays seem to have similar noise conditions, yet the fk-method is able to recover a phase velocity dispersion curve only for P2. The nature of the wavefield can indeed play a role and we believe that it boils down to the coherency of the wavefield rather than to the azimuthal distribution. In fact, as shown in the coherency plots which are now shown in the Supplementary Material (SM4), despite the similarity of the array configuration of P2 and P3, the latter shows a sharp drop in coherency in the 10-20 Hz range (for 15 Hz, coherency is an order of magnitude lower at P3 compared to P2). Therefore, the likely prevalence of incoherent noise overcoming the coherent, low amplitude noise wavefield, might explain the difficulty to recover a dispersion curve through

fk-analysis at this site. The MSPAC methodology has been initially considered for all three arrays but provided unsatisfactory results, especially for the spiral array at P2 (2024).

The legend of Figure 7 is incorrect: the description of panel b) is reported as panel c), and the description of actual panel c) is missing. Similarly, in text panel c) is referred as panel b) (line 197).
To follow rev#1 suggestion, Fig 7 has been completely changed. The caption is now correct.

According to the sensitivity plot in Figure 7.c, it seems that for the range 10-20 Hz the data inversion would be informative of the fist 50-100 m maximum, more than 100-150 as stated.
Corrected.

I strongly advise the Authors to already provide a repository, for example a Zenodo project, instead of stating that "seismic recordings will be made available". You could embargo the dataset until the paper is published, so that in the meantime at least the metadata for the dataset is FAIRly accessible. Also, because of this, the analyses are not currently reproducible.
The recorded seismic data at the three arrays are now available at the EIDA repository. This is now specified in the "Data and code availability".

References:
I would appreciate if you could add the DOIs in your References.
I believe you could find more direct sources for the claims in section 3.3. As for the one regarding wind gusts (lines 166-167), Mucciarelli and Gallipoli 2001 provide a good overwiev of the whole HVSR technique in microtremors, but the original sources for the case at hand should be the cited Mucciarelli paper (Mucciarelli 1998) and even better the SESAME guidelines, which report precisely the warning of not recording data under strong wind conditions (> 5 m/s).
Following the reviewer suggestion, we have also added the Mucciarelli (1998) citation. When the P2 (2024) array acquisition was carried out, a meteorological station was installed at the site, providing wind speed and direction with a 10-min sampling rate for the entire duration of the seismic acquisition. Wind speed never exceeded 5 m/s and this is now also reported in the manuscript, thus assuring the compliance with the HVSR acquisition best practices.

As for the claim regarding the Rayleigh-Love partitioning depending on noise source variability (line 167-168), I do not find any correspondence with the paper you cited (Xiao and Wang 2022), which analyzes data from Mars and contains no similar claims, at least to my understanding. Please explain your choice of using this reference or provide a different one.
We have now inserted the correct citation to support our claim.

CMDS and BayesBay could be cited in the "Data code and availability" section, instead of in-text.
These are now inserted in the "Data code and availability" section.

Once these issues have been dealt with, I believe your manuscript has the potential to be of interest and usefulness in the context of the site characterization related to the Einstein Telescope project, and also to provide some broader insights to the field of local seismic noise characterization.

I hope I have not missed some detail or misinterpreted your views and I wish you good work for the revision.

Reviewer A:

Dear Authors,

I have completed the review of the latest version of your manuscript “Seismic array measurements in the Italian candidate site for the Einstein Telescope, the third-generation gravitational wave detector”.

I appreciate your work in reviewing the manuscript according to both mine and the other Reviewer’s comments. I particularly appreciate your care in providing the context for your work, both in terms of previous experiments and of the site geology (although maybe you could consider slightly simplifying that section). I believe you work is now improved in terms of clarity and strength.

The section regarding the site geology has been slightly shortened.

I only have a few more minor comments that should be easy to address:

- General comments:
    - L136-137: you state that the configuration of the spiral array “allows to push the limit of non-aliased frequency up to 100 Hz”. This consideration is only true strictly speaking in terms of array geometry, but it should be remarked (although it may seem trivial) that the sampling frequency also affects the total non-aliased frequency band of the recorded signal. In your case that would lower the maximum frequency to 50 Hz.

      We state now that temporal aliasing occurs for frequencies above 50 Hz given the 100 Hz sampling rate

    - Tables 2, 3 and 4: could you please check all of the values reported in these tables? It seems to me like you forgot to update some of them (e.g., the reported wavelength/aperture ratio is the same even if D is different; in table 3 it may be due to rounding (and I would suggest adding a second decimal), but in tables 2 and 4 the values just seem wrong for the provided v, f and D values)

      Values have been checked and corrected.

    - L292-293: please provide (either here or in the supplementary material) some reference for the methodology used to calculate coherency

      Reference has been added in the main text and in the SM.

    - L295: here you talk about “extremely low level of ambient noise at all frequencies”. However, as you also state in your reply, you interpret it as a low level of *seismic background* noise, not of *ambient* noise overall (which would include e.g. the wind); you actually state many times that the discrepancy in noise comes from weather conditions. This line should be rephrased, or you should specifically refer to seismic background noise instead of ambient noise.

As the reviewer suggests, we now refer to seismic background noise.

- Data and Code Availability section: I appreciate that the datasets are now publicly available at EIDA. Could you add a more direct link to the webservice where they are available, in addition to citing the paper describing the node?

  The exact urls to the EIDA webpages have been added.

- Typos:
  - L106: “18-28/09-2021” should be “18-28/09/2021” CORRECTED
  - L112: please check, I believe there should be a comma where there is now a full stop CORRECTED
  - L130: “of each of the arrays” CORRECTED
  - L295: “at all frequency” should be “at all frequencies” CORRECTED
  -
- Figures:
  - Fig.8: please increase the font size of all subplots (axes, ticks and legend labels) which are now difficult to read. These have been increased by 20%
  - Fig.8: the legend still reports the sentence “Dashed lines indicate the limit of the resolved region.” which is now not applicable CORRECTED
  -

I wish you good work for the revision.

Recommendation: Revisions Required