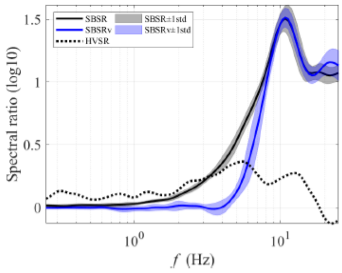
This study presents a compilation of mostly existing site information in search of reference sties in Greece. It collects some important site data which are crucial for may downstream studies, serving as an important dataset for Greece and the international community. The manuscript is relatively well written, and I enjoyed reading it very much. Most of the information here is collected from existing sources, in my view, addition efforts could be devoted to uniformly deriving some extra information, e.g., empirical site response, site kappa, and topographic parameters, to render the dataset even more valuable. The following are my specific comments:
➔ We thank the reviewer for their positive feedback. We appreciate his suggestions and agree with many of them, but will endeavor to explain in what follows why most of the additional studies proposed are at this moment out of the scope of the present work.

Major comments:

1.
GIT or deltS2S site response: since waveforms are available, why don't the authors derive site response from observations using either generalized inversion or residual analysis to the median prediction from a ground motion model (GMM)? These two techniques are equivalent and can give same results when the same constraints are applied. Though reference sites need to be assumed, it can be used as the average site response over all sites in the dataset. Consequently, the resulting site responses are relative to this condition. However, in my view, these observed (relative) site responses carry the most important information in terms of reference site selection. A good reference site shall exhibit smooth change in amplitude with frequency, thus teasing out sites with significant resonance (with clear peaks or throughs).
➔ We really like the techniques mentioned by the reviewer and will be very happy to test them at a later date, after our database has been curated with respect to all the necessary parameters. At this time, it is not straightforward to implement neither a broadband spectral inversion for source, path and site components, nor a full residuals analysis with respect to GMMs. Both these techniques would necessitate further curation and homogenization of non-site parameters, most notably catalogue parameters such as magnitudes and fault mechanisms. Both GIT and ds2s can suffer from trade-offs and need the user to exercise care so that e.g. source-related factors do not map onto site-related (or even path-related) ones and vice versa (e.g. stress-drop vs moment or path attenuation vs site amplification in the one case, τ vs φ trade-offs in the other case, etc.). For a successful result, we believe it is also generally desirable to have a well-balanced source-station distribution with as many events as possible being simultaneously recorded at as many stations. None of these issues or limitations really hinder HVSR; this is why we believe that spectral ratios are indeed the first technique we should be using in this new dataset we compiled, before delving into less straightforward ones that require additional checks, experience and judgement for correct interpretation. Thus, we would like to respectfully save the reviewer's recommendations for when we are ready to move on to a full inversion and/or residuals analysis, which will constitute a new research step. We also point out that adding to the scope of the work performed implies additional resources.
Based on the reviewer's recommendations for these additional analyses, we are concerned that perhaps our work may have struck the reviewer as somewhat limited in scope or simple. In an effort to rectify this impression, we would like to permit ourselves to point out a few considerations:
1. A large part of this work comprises the compilation/curation of the database itself. It might perhaps be easy to pass by, yet it actually took over a year to carefully analyse the waveforms with the in-house software we developed ad hoc for this purpose. It was an iterative (and rather painful, despite having some significant past experiences in strong-motion processing) procedure, during which our processing procedure was improved, the data reanalysed, and several errors in the data or metadata were discovered, investigated and eventually amended. We do not claim to have done a perfect job, but our experience taught us that creating a database from scratch to perform analyses is truly challenging compared to using existing curated ones. We would thus like to respectfully draw some attention to the value we feel this effort carries in its own right.
2. Although our empirical analysis only includes HVSR, we would like to point out that we did try to use the technique in as exhaustive a way as possible, including features that are –in our experience- quite rarely explored, such as:
a. going to great pains to control the usable bandwidth within which it is computed, so as not to attribute significance to results where the data are too noisy;
b. carefully examining the full range of rotational possibilities, so as to sample the site effects not only at the as-recorded directions but in all possible directions (since geomorphological features do not have to oblige by following the orientation of the sensor), so as to find the maximum amplification in

whichever direction it takes place, and even assessing the rotation-related variability of the amplification over selected frequency ranges in a quantitative way, which is something not done before -to our knowledge.
c. considering the known fact that HVSR underestimates the amplification level and likely the higher modes, we implemented the correction method of Ito to provide an approximate estimate of the potential amplification level that could be expected.

Though HVSR is used here, it is well known, and is clearly pointed out by the authors, that HVSR only approximates (horizontal) site response at frequency range where there is no significant vertical site response. However, we know as little in which frequency range that vertical site response is negligible at a specific site, as the horizontal site response.
Thus, it is valid that: a good reference -> flat HVSR over broad frequency range, however, invalid: flat HVSR → a good reference. I just provide one real example (KiK-net site) below. The HVSR seems to be relatively flat, which is because the horizontal peak is canceled by the vertical.



Therefore, I suggest the authors to consider using GIT or residual analysis to derive empirical (relative) site response to compliment HVSR, as did Lanzano et al. (2020).
For this very reason, consider to use “Earthquake HVSR” directly, rather than ‘Transfer Function’ as section title. The former is clearer to readers.
➔ We thank the reviewer for this comment. We agree that a flat HVSR does not necessarily a perfect reference site make, though the reverse should be true. However, we’d allow ourselves to comment that the plot given as an example of HVSR limitations seems a bit extreme to us: it shows a site whose amplification according to SSR seems to reach 10^1.5=32 at 10 Hz. We have never come across such a site, and although we do not doubt the correctness of this plot, we do believe it describes a somewhat special case. (We also wonder whether the horizontal downhole FAS may not have a sharp trough at 10 Hz that might perhaps cause an exaggerated peak of surface SSR -- in other words, is the surface being divided with a ‘perfect’ reference site or could this instance be affected by the limitations of the SSR method itself? But this is outside the scope of the discussion, only a passing thought.)
Though we have no means of deriving SSR for Greece, we are happy to include a stronger ‘disclaimer’ in our manuscript as warning that even ‘Ito-corrected’ HVSR amplifications may still underestimate resonance, but we honestly do not expect such a shocking case among our stations.
We’d also like to mention that, although we neglected to include it in our list of references, we are acquainted with the reviewer’s recent work of Zhu et al. (2020) entitled ‘Detecting Site Resonant Frequency Using HVSR: Fourier versus Response Spectrum and the First versus the Highest Peak Frequency’ and we agree -both in theory and in practice, i.e. in this paper- with many of its conclusions. Their conclusions included propositions such as: that FAS-based HVSR yields an estimate of where the maximum site amplification takes place, even if it provides a lower bound; that network operators should make public their fp frequency estimates for their stations as a matter of priority, and preferably their entire HVSR curves for users to scrutinise and reach their own assessment of f0/fp. This is what we have tried to do. We agree it is intriguing to further investigate vertical amplification, but at this stage in our work, we feel we should leave that for the future.
We can certainly use the term ‘Earthquake HVSR’ rather than ‘Transfer Function’ in the paper, if the latter encourages a misapprehension that it is free from the HVSR limitations.

2.
Site kappa: likewise, given available waveforms, I think site kappa can be another source of information for site selection. Please consider deriving site kappa at these sites.
➔ Thank you for this comment, which was also raised by Reviewer 2. We would certainly like to study kappa for the stations, but will need to construct a different database for this purpose. Plots S1 and S2 show the magnitude-distance distribution of our sets per station, and because of our criterion for M4 and above (and the fact that HH stations can clip in the near field), most stations do not have nearby

recordings. For Greece, ‘nearby’ can mean within 20-25 km, as in Ktenidou et al. (2015 - https://doi.org/10.1093/gji/ggv315) we showed that there can be quite an increase of κ_r with distance after 30 km. Only a handful of our stations have more than a couple of data points within 30 km. For the vast majority of them an extrapolation of κ_r to R=0 would yield a significant uncertainty in κ_0, as the latter would be rendered strongly dependent on the path attenuation being removed. So again we stumble upon the issue of site-path trade-offs, and to resolve this to a satisfactory extent we feel we must address it in a future step with an improved dataset.

3.
Data availability: I strongly suggest providing the information in various tables in the paper in a single flatfile (e.g., csv) as electronic supplemental materials, which will greatly facilitate the use of these results by end users. Ideally, only symbols in the table header that are readable to machines can be used such that the table can be read directly by computers.
➔ We can certainly try to do that to facilitate the uptake of our results by readers (and machines!).

4.
On HVSR computation: it is mentioned that the two horizontal components are combined as the square root of the sum of squares, rather than the root-mean-squares. Thus, would HVSR be unity, or sqrt(2), even at perfect rock sites? This is important for the following statements in Line 150.
➔ The reviewer has a point. We can amend the manuscript to clarify the effect of how components are combined.

5.
Table 4: some sties lack topographic data, i.e., slope. There topographic data can be readily derived from openly available DEM. I suggest the authors devote some efforts to deriving them.
➔ We can look into that.

Minor comments:
1.
Line 225: S1 should be A1?
➔ Yes indeed, thank you for catching that.

I have happy to waiver my anonymity.
Best regards,
Chuanbin