We would like to thank the editor for the constructive comments on our manuscript. In the following, we will answer to the editor's comments and suggestions.

First of all, some sentences on the data quality are missing. Far distance events with 5<=M<=5.5 might contain very low energy, the authors have to state if they have made some selection among the recorded data and how; moreover, showing one example of M 5.2 is not enough, considering that its epicentral distance from the array is not mentioned. At least the authors have to show that events with M<=5.5 occurred at the largest possible epicentral distances still retain the energy necessary for this kind of analysis. The authors have to show that their data selection is the best as possible, a way of proving it is to show how the image would look like if only events with M>=6 are used.

We agree that the explanation of the data selection could be improved. We have followed the suggestions and have added more details in the form of supplementary material, as follows,

A total of 81 MW ≥ 5 earthquakes were selected for further processing. The selection of the minimum magnitude to be considered was taken as a balance between the signal quality of the earthquakes and the number of available sources for each deployment.

For each deployment, all earthquakes with MW > 5 and epicentral distance > 120º were selected. All the events within the data were checked, and the power-spectral density was computed to confirm the existence of useful energy within the selected frequency band. This process gave the final selected 81 sources.

For the first deployment a total of 44 earthquakes were analysed and 17 proved to have useful energy. In the second deployment, 59 earthquakes were evaluated, from which 38 were selected for further processing. For the final deployment, 60 earthquakes were analysed and 26 were selected to produce the final image.

For each of the three deployments a different number of earthquakes is available, 17 for the 1st, 38 for the 2nd, and 26 for the 3rd. Among those, there are 3 in the 1st deployment, 16 in the 2nd deployment and 14 for the last deployment that are with MW ≥ 6.This amount of events per deployment is insufficient for application of GloPSI. To suppress retrieval of artefacts, for example due to the PKP triplication, it is important to sum phases from a wide range of ray parameters (Ruigraok and Wapennar, 2012; Nishitsuji et al., 2016).

We have included as well a plot with the total percentages of earthquakes used in the ranges, 5 ≤ MW < 5.5, 5.5 ≤ MW < 6 and MW ≥ 6.

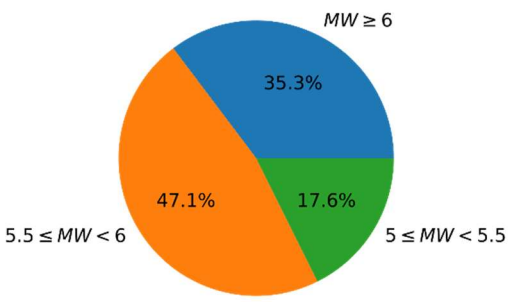
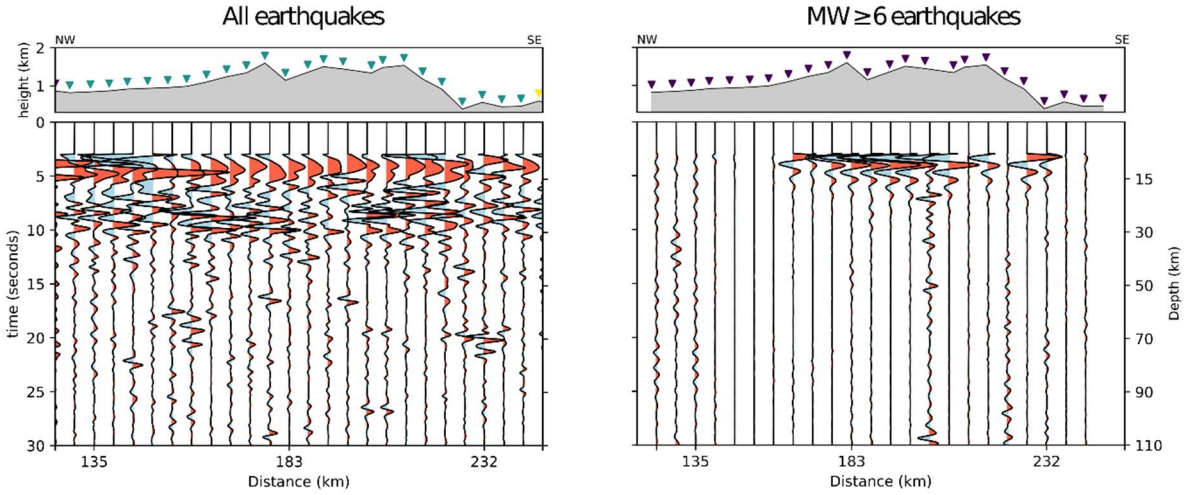


Figure S1. Percentages of events per magnitude for all used earthquakes.

Regarding the 5.2 MW event shown in Figure 3 in the manuscript, the epicentral distance is 147.1º, which, makes it the farthest 5.2 magnitude event. It is worth noting that the events, with magnitudes between 5.1 and 5.4, correspond all to the first deployment. It has been stated in the supplementary material as well.

The use of low magnitude events, between MW 5.1-5.4, is restricted to the first acquisition where the deployment time was the shorter and therefore, we were forced to include lower magnitude events in the processing scheme. These events represent the 35 % of the total events used to produce the central part of the image.

For all these reasons, we believe our data selection is good to produce a reliable lithospheric image. However, we have followed the suggestion of the editor and created an image, just of the central segment where there are only 3 earthquakes with MW ≥ 6, and compare it with our final image (we have not include it in the supplementary material).



In the image with only MW ≥ 6 events, the amplitude and frequency of the autocorrelations it is clearly different, and there is small reflectivity retrieved. This could be due to a lack of information in the stacking process, or due to not retrieving the stationary phase during the stack. This would mean, that these 3 events could interfere destructively in the summation process or that the events do not cover a wide range of ray parameters, needed to produce a quality image.

Finally, we have moved the table of earthquakes presented at the end of the manuscript to the supplementary material.

Then, the several steps of how the GIoPSI image has been created must be shown, in particular, the image before muting the delta pulse at t=0 and (as suggested by R2) the results after the deconvolution of the virtual source-time function and subtraction of the average time function (such image might go in the supplementary material). I would also ask how the authors are dealing with the multiple suppression, and if there are any multiples recognized in the produced image.

An image comprising all the post-stacking processing steps has been included in the supplementary material. The post-stacking process is quite straightforward as only a band-pass filter and muting of the delta pulse are applied.

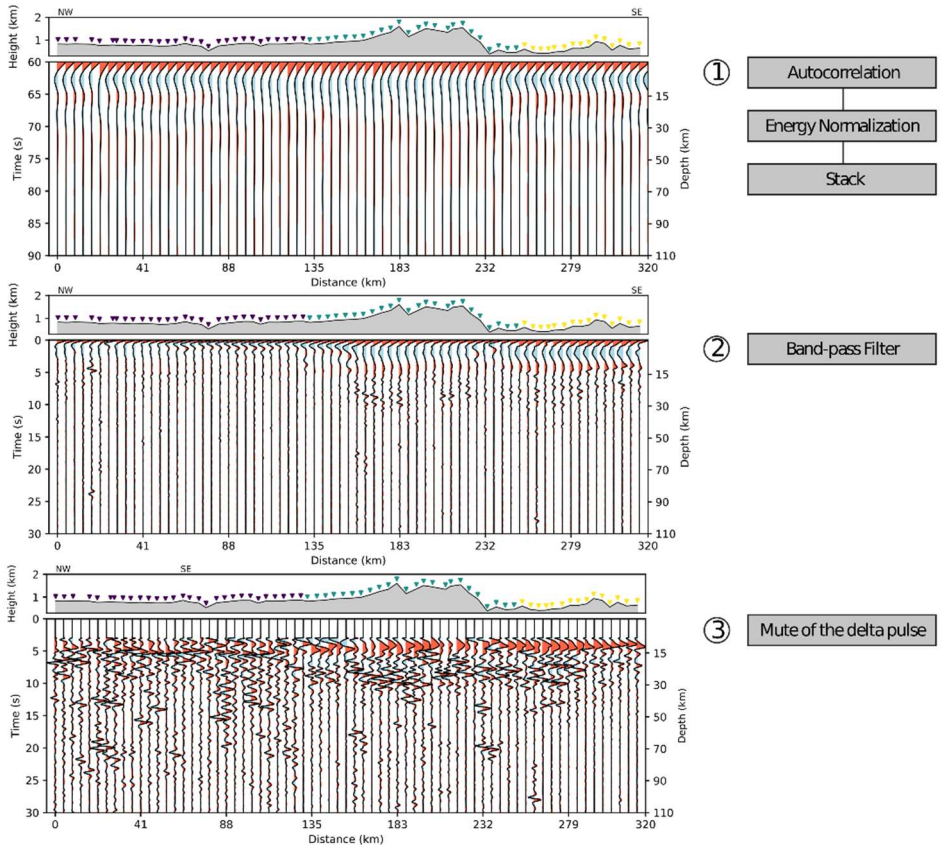


Figure S2. Stages of the processing zero-offset reflection data to construct the lithospheric image across the Spanish Central System

The following explanation has been added to the supplementary material,

The post stack processing applied to generate the zero-offset reflection profile is shown in figure S2.
Step 1 represents the raw stack of the autocorrelations of events for the three deployments, where the delta pulse around t=0 s dominates the image, as its amplitude is much higher than the amplitudes of the rest of the profile.

Step 2 comprises band-pass filtering of the raw stack to enhance higher frequencies and reject the frequencies dominated by the microseism. Therefore, a band-pass filter of 0.7-2 Hz is applied.

In step 3, the dominating delta pulse is muted. The selection of the muting window is based on the entire wavelet (including the positive and negative times of the autocorrelation). The length of the wavelet is 6 s, 3 s positive and 3 s negative times. Therefore, the muting window applied is 3 s.

We have included in the supplementary material an image comparing the three approaches to eliminate the delta pulse around t=0 s, along with a brief explanation of each strategy.

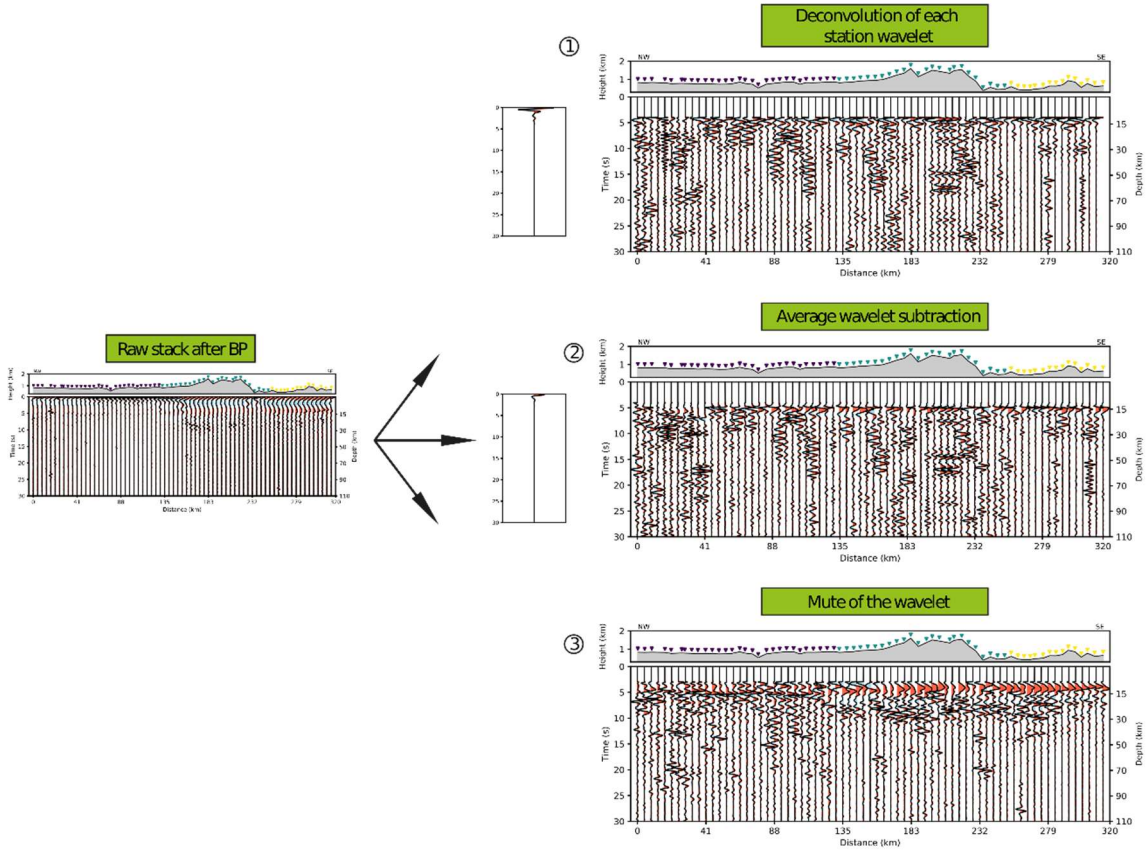


Figure S3. Approaches to eliminate the influence of the delta pulse at t=0. 1) deconvolution for each station of the wavelet, 2) subtraction of the average wavelet of all the station, 3) muting of the wavelet.

Three different approaches were used in order to eliminate the influence of the delta pulse at time t=0 (Fig. S3).

First deconvolution of the wavelet for each station was tested. For each station the wavelet dominating the trace around t=0 s is extracted and used for deconvolution. The construction of the wavelet used the full autocorrelation stack (i.e., positive and negative times). The time window of the wavelet was selected by visual inspection and it was selected to be 5s. This approach did not yield good results as it suppressed most of the reflectivity throughout the profile.

The second approach is based on the subtraction of the average delta pulse. To construct the wavelet, all the stations are stacked together. The selection of the time-window to extract the wavelet followed the same procedure as in the deconvolution approach. Then the wavelet is subtracted from each station stack. This approach seems to produce similar results as the deconvolution, except it preserves more reflectivity earlier than 5 s. Still, most of the coherent reflectivity was suppressed.

The selected technique to eliminate the delta pulse was muting. We selected the time window to be muted as in the other two procedures, but to keep it as possible. A window of 3 s was selected, even though the wavelet is probably slightly longer at the southern stations compared to the northern ones. This procedure preserved the reflectivity of the profile and was selected for that reason.

We are not applying any multiple suppression technique as we are relaying on the reflectivity patterns rather than in individual and clear reflections. Nevertheless, multiples in the image

would be related to sedimentary basins, and crustal interfaces such as the upper-lower crust boundary and the Moho.

The first ones would be associated with the Duero Basin, the Tajo Basin, and the Tietar Basin. The depth of these basins is rather unknown making it difficult to apply predictive-deconvolution. Nevertheless, if we assume a depth in the range of 1-2 km depth and standard velocities for sedimentary basins of 3500 m/s, the multiples should be restricted to the upper crust. Furthermore, considering the frequencies used, the top and bottom of shallow basins could not be even resolved.

Multiples related to crustal interfaces such as the Moho should be present at times greater than 20 s. Little reflectivity is found below 20 s TWT allowing us to think that, if present, multiples are not disturbing our interpretation.

I understand that the main argument for discerning between crust and mantle is the different reflectivity of the two, while concerning the boundary between upper and lower crust the authors are referring to the different wavelength of the reflections. Honestly, as a not-expert eye, the choice of which of the reflectors has been interpreted seems arbitrary. The different amplitudes might be due to the different events used for imaging the three segments of the profiles. The authors have to clearly show that what they interpret is not an artifact but rather a robust feature, moreover the authors have to show clearly if they have used prior information on the local structure for their interpretation.

As we state in the manuscript, the interpretation is based on changes in reflectivity. We also state that we base our model with previous knowledge of the area and surrounding areas, and therefore, none of the proposed features is completely new, or to some instance random.

The boundary between the upper and lower crust is evidenced by a drastic change in the reflectivity pattern between the upper and lower crust where the frequency content and amplitudes clearly changes. This interpretation is supported by the fact that normal-incidence and wide-angle seismic data within the Iberian Massif reveal the same feature to the south of the profile, in the ALCUDIA dataset. In fact, in the crossing point between the presented dataset with the Alcudia-WA dataset, both mid-crustal interfaces appear at the same, travel time. So we suggest that the boundary we see at mid-crustal depths is the northward continuation of a feature seen along 600 km to the S of our area of study. The discontinuity can be seen in Martínez-Poyatos et al. (2012), and how it extends southwards in Simancas, et al. (2013).

Regarding the difference in amplitudes, one of the steps in our processing workflow was to normalize by their energy the autocorrelations before stacking, which make the final stacks not to be affected by different amplitude of the events.

Finally, there are several interpreted features in the mantel (outlined by green dashed lined in Figure 5). As far as I understand the GIoPSI can be applied to very long arrays, due to the fact that this method is valid for interpreting phases that are constantly recognized along the array itself; therefore, such short (in space) reflectors might be artifacts and their interpretation misleading. The authors have to show how they can distinguish between real signals and artifacts in such an image.

Identification of phases with GloPSI follow the same principles as any seismic image and, therefore, reflectivity will be retrieved whenever there is a detectable change in the physical properties of the underlying materials. Because of that, GloPSI could be applied even to single station. However, using arrays always facilitates interpretation, so having several stations close to each other is always preferred. Also, the imaging with GloPSI is conditioned by the inclination of the identified structures. If they have a high dip, they will not be interpretable by their presence, but by their absence in an image from an array of stations.

The reflections that we are seeing within the upper mantle have lengths between a few kilometres and 50 km. We correlate the depth of this reflectivity with similar features seen in wide-angle data for southern Spain (Ayarza, et al., 2010). A recent study (Palomeras et al., submitted) explores these mantle reflectors further north but still in the Iberian Massif by means of a compilation of wide-angle profiles. These authors find the same features in the mantle in the western part of the Iberian Peninsula.

We agree that the lack of control over possible artifacts should be noted, especially in connection to features of shorter lateral continuity. We have included the following lines in line 465,

However, the lack of control on possible artifacts within the upper mantle should be noted and these results should be taken carefully.