Mapping the basement of the Cerdanya Basin (Eastern Pyrenees) using seismic ambient noise.

Jordi Díaz[1], Sergi Ventosa[1], Martin Schimmel[1], Mario Ruiz[1], Albert Macau[2], Anna Gabàs[2], David Martí[1], Özgenç Akin[1,3] and Jaume Vergés[1]

[1]GeoSciences Barcelona, Geo3Bcn, CSIC c/ Solé Sabarís sn, 08028 Barcelona, Spain
[2]Institut Cartogràfic i Geològic de Catalunya, Barcelona, Spain
[3]Karadeniz Technical University, Trabzon, Turkey

Correspondence to: Jordi Díaz (jdiaz@geo3bcn.csic.es)

Abstract. Ambient seismic noise acquired in the Cerdanya Basin (Eastern Pyrenees) is used to assess the capability of different methodologies to map the geometry of a small-scale sedimentary basin. We present results based on a 1-year long broad-band deployment covering a large part of the Eastern Pyrenees and a 2-months long high-density deployment covering the basin with interstation distances around 1.5 km. The explored techniques include autocorrelations, horizontal to vertical spectra ratio, ambient noise Rayleigh wave tomography and band-pass filtered ambient noise amplitude mapping. The basement depth estimations retrieved from each of these approaches, based on independent datasets and different implicit assumptions, are consistent, showing that the deeper part of the basin is located in its central part, reaching depths of 600-700 m close to the Têt Fault trace bounding the Cerdanya Basin to the NE. The consistency between the results from all the methodologies provides solid constraints to our basement depth estimation. The results show also that when high-density seismic data are available, mapping the ambient noise amplitude in a selected frequency band is a valid tool to quickly map the sedimentary 3D geometry. Beside this methodological aspect, our results help to improve the geological characterization of the Cerdanya Basin and will provide further constraints to refine the seismic risk maps of an area of relevant touristic and economic activity.

1 Introduction

The objective of this contribution is to evaluate the potential of several methodologies based on the analysis of the seismic signals recorded in the absence of earthquake-generated waves, as autocorrelations, Horizontal to Vertical Spectra Ratio (HVSR), Ambient Noise Tomography (ANT) or noise amplitude maps to define the geometry of the Cerdanya Basin (CB), a relatively small Neogene sedimentary basin located in the eastern part of the Pyrenees Axial Zone. The basin extends along 35 km, has a maximum width of 5-7 km and is crossed by the Segre River, one of the main tributaries of the Ebro River. The mean altitude of the CB is of 1100 m, with surrounding mountain ranges reaching 2500-2900 m. It is limited to the East by the Alp-Têt Fault, interpreted as a major structural feature in the Eastern Pyrenees formed by NE-SW right-stepping en-echelon faults and E-W oriented faults.

Some geological and geophysical studies have provided information on the structure of the subsoil in the first hundred meters depth in the CB, using vertical electric sounding (Pous et al. 1986), seismic (Macau et al. 2006), gravimetric (Rivero et al. 2002) methods or geological studies (Cabrera et al. 1988). However, the results obtained were rather unclear and difficult to correlate with lithological information. The most relevant contribution to the knowledge of the basin geometry was published by Gabàs et al. (2016) and included the joint use of magnetotelluric and passive seismic data along a high-density 2D profile across the basin, between the villages of Ger and Alp. The obtained models show an average value of the electrical resistivity overburden close to 40 Ohm·m and can be correlated with Quaternary and Neogene deposits. The derived bedrock profile has a maximum thickness of 500 m near its SE termination and an asymmetric geometry, with a smooth increase in depth to the NW and a more abrupt change in the SE termination. This layer is a resistive zone with electrical resistivity values between 1000 Ohm·m and 3000 Ohm·m and could be correlated with the top of the Palaeozoic rocks constituting the basement (limestones and slates) (Roca 1996; ICGC 2016).

We use the seismic data acquired in the framework of the SANIMS project (Spanish M. of Science, Innovation and Universities, Ref.: RTI2018-095594-B-I00), which includes two different deployments. Firstly, we deployed 24 broad-band period stations covering the CB and the surrounding areas with a twofold objective; investigating the basin and providing data for regional-scale tomographic studies (Fig. 1b). Ten of the stations were deployed along an EW profile crossing the CB with an interstation spacing of 4-6 km. The rest of instruments were deployed forming an outer ring at approximately 35 km of the basin. These instruments were active between September 2019 and November 2020. Secondly, we deployed a high-resolution network covering the basin using 140 Rau-Sercel nodes equipped with 3 component high-frequency geophones (Fig. 1c). The network had an interstation spacing of 1.5 km, covered an area of about 300 km^2 and has been active for two months, between April and June 2021. Additionally, a high-density node profile, crossing the basin along a NW-SE line, was designed with an interstation spacing of 700 m. Although the two deployments were planned to be operative during the same time period, the logistical constrains related to the COVID19 mitigation measures delayed the high-density station deployment by one year.

Before discussing the results provided by noise-based methodologies, we want to note that a first insight of which is the area with thicker sediments arises from the inspection of the Receiver Functions (RF) calculated with the main objective of mapping the bottom of the crust. The RF method uses the P-to-S wave conversion at large velocity discontinuities to map subsurface structures, typically the Moho, and is widely used to explore crustal structure. (Zelt and Ellis 1999) have described the effect of sedimentary basins on RFs, which include a phase lag in the direct P arrival resulting from the increased Poisson's ratio in the sedimentary layer and the presence of large reverberating phases that can overprint the arrival of the phase converted at the Moho. Figure 2 shows the RF stack at the broad-band stations installed along the CB. It is easy to observe that stacks corresponding to stations CN02 to CN10 show late arrivals of the direct P wave, with maximum time lags for stations CN07 and CN08. These two sites show also large reverberations between 2 and 6 s, hence suggesting the presence of a significant

65 sedimentary cover in the central part of the CB. Further modeling, out of the scope of this contribution centered in the use of ambient noise, can provide additional information on the properties of the basin (Yu et al. 2015).

2 Autocorrelation methods

Autocorrelation methods are based on the evaluation of the similarity of a seismic trace with a delayed version of itself, as this similarity depends on the subsurface structure. Claerbout (1968) showed that the zero-offset Green's Function of a one-
70 dimensional medium can be recovered from the autocorrelation of transmitted plane waves originated in the subsurface. For 2D and 3D media, Wapenaar (2004) has proved that this approach is still valid, although presence of wave fields which are not diffuse does not allow to recover the exact function. However, the obtained result, usually referred to as the empirical Green's function (EGF) to express its approximative character, is now widely used to characterize the subsurface structure.

75 Autocorrelation and cross correlation of ambient noise have been applied to dense station deployments to retrieve P wave reflections for crustal-scale imaging (e.g., Ruigrok et al. 2012). More recently, this approach has been used to map the Paleozoic basement in areas as the Ebro Basin (Romero and Schimmel 2018), as it provides a fast and consistent imaging of the basement structure. However, mapping such shallow structures demands to work in frequency bands between 3 and 18 Hz, a point that may hamper the applicability of the method due to the dominance of local noise overprinting the weak amplitude
80 body wave reflections. Further, the presence of structural complexities complicates the EGFs and often results in ambiguities in the interpretation of the autocorrelations. These ambiguities, nevertheless, can be reduced by using dense station deployments and a priori information arising from well logs.

In this contribution we have calculated the autocorrelations for all the broad-band stations located along the CB. The pre-
85 processing consists in correcting the raw data to ground velocity, band-pass filtering from 8 to 20 Hz, dividing into one-hour-long non-overlapping sequences, and rejecting sequences containing gaps or high-amplitude signals. We compute autocorrelations up to a maximum lag time of 20 s using wavelet phase cross-correlations with a complex Mexican-hat wavelet with 2 voices per octave and no decimation. Then, we smooth the hourly autocorrelations, stacking one-day-long consecutive cross-correlations separated by 12 hours and weighting them by the inverse of the norm between 2 and 3 seconds.

90

Most of the broad-band stations located along the CB show reflectors related to the bottom of the basement (Fig. 3), arriving at two-way travel times ranging between 0.4 s and 0.6 s. The Vs models obtained by Gabàs et al. (2016) show velocities between 0.5 and 1.0 km/s in the uppermost layers. From the Vs/Vp relationship proposed by Brocher (2005), we assume a Vp value around 2.0-2.25 km/s, resulting in basement depths range between 420 m for CN03 and the 670 m for station CN07Click
95 or tap here to enter text., hence providing our first quantitative estimation of sediment thicknesses in the basin, which are consistent with the insights provided by RFs.

3 Ambient noise tomography (ANT)

Ambient noise tomography is based on the extraction of the fundamental mode Rayleigh waves to measure inter-station group
100   and phase velocity dispersion curves. The obtained dispersion curves are then inverted following a hybrid l_1-l_2 norm criterion
using the fast marching method on the forward problem to produce velocity maps for a set of periods.

The data gathered with both the broad-band and the nodes deployments, altogether with the data at the permanent stations
covering the area have been used to obtain ANT models at two different scales; a regional one, covering most of the NE section
105   of Iberia, and a more local one, centered in the CB. We focus here to the local tomography study. The data processing includes
correcting the raw data to ground velocity from 0.05 to 20 Hz, band-pass filtering from 0.1 to 5 Hz, decimating to 20 samples
per second, dividing into one-hour-long non-overlapping sequences, and rejecting sequences containing gaps or high-
amplitude signals. We compute symmetric cross-correlations up to a maximum lag time of 90 s using the wavelet phase cross-
correlation, and estimated the average and the confidence interval of the phase velocity extracted from the cross-correlation
110   ensemble per station pair through resampling strategies (Schimmel et al. 2017). Basically, we stack each jackknife sample
(Efron and Stein 1981) with the two-stage time-scale PWS (Ventosa et al. 2017) and then measure Rayleigh phase-velocity
dispersion curves following (Ekström et al. 2009). Finally, we construct Rayleigh phase-velocity maps solving an inverse
problem with l_1-norm misfit function on the data space and a l_2-norm on the model space using the steepest-decent method,
and applying the fast marching method (Rawlinson and Sambridge 2005) to solve the forward problem.

115

The map obtained for the shortest period available, 1.035 s, shows a clearly defined low velocity zone covering the central part
of the basin, the same area where RFs and autocorrelation methods have already pointed to a significant sedimentary cover
(Fig. 4a). The low velocity zone in ANT maps extends to the NE following the direction of the Têt Fault, including the area
near Puigcerdà, although with slightly higher Vs values, around 2.0 km/s. To the southwest, the end of the CB is delineated by
120   velocities around 2.2 km/s, still lower than in the surrounding areas. For the next period investigated, 1.414 s, the map is similar
although the central and eastern part show a more uniform velocity distribution and the southwestern section, with slower
velocities, still reflects the geometry of the basin (Fig. 4b). The sensitivity kernels of the periods analyzed are mostly sensitive
to depths ranging between 0.2-0.8 km. Therefore, the low velocity zones can be interpreted as corresponding to sediments in
the uppermost 500m, with significant variations along the basin.

4 Horizontal to vertical spectral ratio (HVSR)

One of the most usual methods to characterize shallow structure using seismic data is the Horizontal to Vertical Spectral Ratio (HVSR) method (Nakamura 1989; Bard 2004), as it provides a reliable, fast and low-cost tool to estimate site characterizations. Analyzing the seismic background noise during different time intervals, this method allows to obtain the soil fundamental frequency (f0), related to the strong impedance contrast at the soil-bedrock interface (Field and Jacob 1993). This fundamental frequency can then be used to estimate the depth of the soil-bedrock interface by means of empirical relations with borehole stratigraphies or velocity-depth profiles in places where this kind of results are available (Ibs-Von Seht and Wohlenberg 1999; Delgado et al. 2000; Benjumea et al. 2011; Akin and Sayil 2016). HVSR methods were applied by Gabàs et al. (2016) to define the geometry of the CB along a 2D-profile. The f0 values obtained along the profile range between 0.3 and 1.7 Hz. The use of seismic array methods has allowed the authors to obtain shear-wave velocity-depth profiles and to then infer a scaling law between f0 and basement depths.

The processing of the data acquired by the broad-band stations and the seismic nodes to calculate the HVSR starts with correcting the instrument response to obtain velocity values, filtered between 0.05 and 20 Hz. All available data, spanning over a year for the broad-band stations and 2 months for the node deployment, are sliced into sequences of 240 s with a 50% of overlap and windowed with a Hann taper. Then, their spectra are smoothed, applying a bank of 261 filters with a quality factor (i.e., bandwidth divided by central frequency) of 20 and subsequently averaging their outputs. Finally, the optimum HVSR is measured in the least square sense, i.e., HVSR = sqrt(E{H·V}/E{V·V}) where H and V are the horizontal and vertical spectra, and the expectations E(H) and E(V) are measured as the mean value after removal of the aberrant values. Supplemental Figure 1 shows some examples of the obtained HVSR.

The geophones used by the high-density array have a characteristic frequency of 10 Hz, which means that the recording sensitivity decreases for frequencies below this value. As the f0 values expected in a sedimentary basin are clearly below 10 Hz, these instruments are not the most suitable for this type of study. However, a significant number of the instruments have provided useful f0 values, even below 1.0 Hz. For the broad-band stations along the basin, the lower f0 values (0.36-0.38 Hz) are observed at stations CN07 and CN08, with values between 0.40 and 0.60 Hz for stations CN04, CN06, CN09, and CN10. Stations outside the basin, mostly located over the Paleozoic massif, do not show clear frequency peaks for frequencies below 20 Hz. Regarding the nodes, we have retained 59 valid measurements from a total of 143 (40%). A large number of the nodes installed in the central part of the basin, near broad-band stations CN06-CN08, have not provided useful HVSR measurements. We interpret that the frequency range of these instruments was not sensitive enough to the low-frequency f0 values expected for the sites located in this area. Figure 5a show the retained f0 measurements over the network.

As discussed above, Gabàs et al. (2016) have adjusted an exponential law to relate f0 and basement depth, based on their velocity/depth models. Even if this relationship was inferred for the narrow zone covered by their experiment, it still seems to be better suited for application to our case, located in the same sedimentary basin, than experimental laws published for other

160 basins and has therefore been used to translate the new f0 measurements to basement depth estimations (Fig. 5b). Just four broad-band stations, all of them in the central part of the basin, have depths exceeding 450 m, reaching a maximum value of 530 m below station CN07. Stations located near the Puigcerdà area (1.95º, 42.42º) show depth values above 400 m, while large parts of the basin show basement depths ranging between 350 and 450 m. As expected, the lower values are found at the locations close to the borders of the basin.

165

We have interpolated a gridded surface using the nearest neighbor algorithm included in the GMT software package, using a search radius of 2 minutes of arc and requiring 3 out of 8 sectors providing data. These values have been selected by trial-and-error in order to avoid artifacts related to the interpolation and to keep a good spatial resolution. Although this surface must be taken with caution, as the interpolation is based on a strongly uneven point distribution, it provides our first quantitative map

170 of the basement.

5 Seismic amplitude mapping

Between 0.1 and 1 Hz, in the frequency range commonly known as the microseismic peak, the origin of the ground vibration is mainly related to the interaction of oceanic waves (Díaz 2016). This explains the great similarity of the spectrograms in this range for all the analyzed stations. Background seismic vibrations at frequencies above 2 Hz in stations located near populated

175 areas are dominated by human activities. The seismic signals show typically a large daytime/nighttime variation, with large amplitudes during working hours and much smaller ones during nighttime and weekends. This point has been evidenced during the recent COVID19 lockdown, when seismic data in the 2-20 Hz has been used as a proxy of human activity, both at local scale (Diaz et al. 2021; Maciel et al. 2021) or at global scale (Lecocq et al. 2020a). The frequency range between 1 and 10 Hz, located between the microseismic peak and the band dominated by anthropogenic noise, provides the best opportunity to

180 explore the eventual relationship between seismic amplification and geological structure. Other processes, as rainfall or wind bursts can contribute to the observed amplitudes, but their effect tend to be limited in time, while the amplification effects due to sediments should be observed continuously.

In order to analyze the amplitude variations as a function of time, the instrumental response is removed following standard

185 procedures. The Power Spectra Density (PSD) is then calculated to quantify the energy levels at each frequency, using an Obspy implementation (Krischer et al. 2015) of the classical PQLX ("IRIS- PASSCAL Quick Look eXtended") software (Mcnamara et al. 2009), based on the open-access "SeismoRMS" software package (Lecocq et al. 2020b). The data is divided into 30-minute windows with 50% of overlap and the PSD of each window is computed using the Welch method. The

spectrograms retrieved from the PSD analysis show the power of the seismic acceleration, expressed in decibels (dB) referred
to 1 m^2/s^4/Hz. The inspection of these spectrograms (Fig. 6) confirms that the day/night variations typically related to human
activity dominates the spectra at frequencies above 10-15 Hz. It can also be observed that, for frequencies above 40 Hz,
episodes of increased amplitudes can be recognized at many of the stations. Recently, Diaz et al. (2022) have shown that the
seismic signals at this frequency range are dominated by rainfall episodes and proposed to use seismic data as a proxy of
rainfall. These observations confirm that the 1-10 Hz band is the best choice to analyze a possible relationship with the subsoil
geology. The effect of anthropogenic noise is still visible in this band, in particular for nodes located close to villages or main
roads, but, as shown in Fig. 6, its energy is much lower than for frequencies above 10 Hz.

From the calculated PSD spectrograms, we extract the median value of the power spectra in the 1-10 Hz band for the whole
available records (12 months for the broad-band stations, 2 months for the nodes). The amplitude of the power spectra is
calculated at intervals of 30 minutes and expressed in dB relative to 1 m^2/s^4/Hz. In a first stage, this procedure is applied to the
broad-band stations installed along the basin (Fig. 7a). As observed, the largest values correspond in an area coincident with
that derived from RFs, autocorrelation and HVSR studies.

The same approach has then been applied to the dense seismic network available in the CB. Fig. 7b shows the median values
for all the nodes, clearly showing a distribution with low-level sites around the borders of the basin and large amplitude sites
in the center of the basin, over the same areas previously identified as showing the largest sedimentary thicknesses. As the
results derived from the nodal network have a high-density distribution, with a site located every 1.5 km approximately, it is
possible to interpolate a continuous grid covering the area. As for the HVSR case, we have used the nearest neighbor algorithm
included in the GMT package (Wessel et al. 2013), to obtain a continuous grid, using the same parameters than in the previous
case.

A visual comparison between the seismic noise amplitude values and the basement depths estimated by Gabàs et al. (2016)
along the Ger-Das profile, suggest that it is possible to obtain a scaling law between both datasets (Suppl Fig 2a). Although a
linear relationship can provide a general good adjustment, the better results are obtained using a degree two polynomial
adjustment, following the expression:

$$depth = 3.97 *dB^2 + 1062.89*dB + 71118.22$$

Suppl. Fig. 2b shows the quality of the adjustment. It can be observed that for depths exceeding 400 m, the power amplitude
increases very slowly, suggesting that there is a threshold effect in the relationship between sediment thickness and amplitude
amplification. In order to avoid extrapolation effects, we have limited the application of this law to noise values ranging
between -133 and -120 dB.

Following this law, power amplitude values are converted to basement depth estimations and represented in a map (Fig. 8a). The dense network facilitates to interpolate a continuous grid using the nearest neighbor algorithm, with a search radius of 3 km, an interval of 0.3 km and requiring 6 of 8 sectors with data. Finally, in order to check the consistency of our results, we

225 have interpolated, using the same parametrization, a new grid using as input the basement depth estimations arising from HVSR and seismic noise analysis (Fig. 8b). As observed, both models are very similar, with the deeper values, reaching values exceeding 600 m, located in the central part of the basin and a secondary maximum with depths around 300 m in the western part of the basin. The model is very similar to that obtained from just the seismic noise data.

6 Discussion and conclusions

230 Ambient seismic noise data acquired in the Cerdanya Basin has been used to assess the suitability of different methodologies to investigate the 3D geometry of this sedimentary basin located in the Eastern Pyrenees. Autocorrelation relies on the reflection of body waves of unconstrained origin, ambient noise tomography is based in the propagation of surface waves of unknown origin between the receivers, HVSR and seismic amplitudes methods are both based on resonances, but can still be considered as independent, as HVSR consider the horizontal and vertical components and provides measurements in terms of

235 frequency, while seismic amplitude is based on vertical data and provides measurements in terms of energy. Therefore, all approaches are complementary, reduce ambiguities and provide a more complete picture of the basin depth. Figure 9 summarizes the results obtained from each methodology and compares the previous knowledge on the depth of the CB, limited to a narrow profile (Fig. 9a), and the present-day estimations, resulting in a 3D regional scale map of the depth of the basement beneath the Neogene-Quaternary sedimentary deposits (Fig. 9d).

240

The analysis of the autocorrelations of the broad-band stations within the basin has shown that reflectors associated to the sediment/basement discontinuity can be identified in most of these sites. Reflections in autocorrelations need to be calibrated using boreholes for their correct interpretations. Unfortunately, no borehole data is available in the CB, but using realistic Vs values, we have estimated that the vertical sedimentary thickness ranges between 420 and 670 m, which agrees with the

245 maximum preserved Miocene-Quaternary sedimentary infill accumulative thickness of almost 800 m (Cabrera et al. 1988) (Fig. 9b). A more quantitative result arises from the use of HVSR to the broad-band stations and the high-frequency geophones deployed with short interstation distances in the basin. Using the empirical formula proposed by (Gabàs et al. 2016), the f0 measurements have been translated to depths. The high frequency geophones used in the node deployment, with a natural frequency cut-off of 10 Hz, are not able to recover the f0 frequencies in the deeper part of the basin, but provide interesting

250 values for the thinner parts, hence providing a first 3D vision of the basin geometry (Fig. 9c). The results clearly evidence that the thicker sedimentary successions are those already identified from the autocorrelation analysis, although the obtained thicknesses are systematically lower. We interpret that this difference arises from the assumptions needed to carry out the time-depth conversion of the identified reflections. Finally, the analysis of the seismic noise amplitude in the 1-10 Hz band, can be

interpreted as an excellent proxy of the sediment thickness along the basin. Passing from amplitudes measured in dB to
255   basement depth is possible by correlating the noise measures at the locations studied by Gabàs et al. (2016) with the
corresponding basement depths. The polynomial correlation obtained allows us to determine a 3D map of the basin, which is
fully consistent with the HVSR and autocorrelation estimation (Fig. 9d). Additional confirmation of the results arises from the
ANT obtained using the high-density dataset (Fig 9e). Although the depth inversion of this dataset is a problem still not solved,
the velocity maps at short periods, sensitive to the uppermost parts of the crust, show low velocity areas clearly consistent with
260   the previous results.

The CB can be divided in two main sections, the eastern NE-SW trending section and the E-W trending western part (Fig. 9f).
This geometry is clearly related to the position of the Têt Fault, which shows a similar change in orientation (Calvet et al.
2022). The thicker sedimentary succession of the basin is located in its eastern part, extending between 1.85ºE and 2.95ºE.
265   This area has a sedimentary thickness exceeding 600 m and seems to thicken to the SE,
towards the Têt Fault. The E-W oriented western part of the section shows a sedimentary depocenter reaching 300 m that
seems to be disconnected from the one in the eastern sector.

Figure 10 shows the basement depth transect along the high-density node profile crossing the central domain of the Cerdanya
270   Basin, altogether with two profiles crossing its northern and southern domains, following the geological models shown in
Calvet et al. (2022). The northern profile shows the base of the basin displaying a gentle deepening towards the SE with its
deepest part located in its central part close to the town of Puigcerdà. This deepest part of the basin nearly coincides with the
limit between the areas covered by alluvial fans of granitic and non-granitic sources. As reflected in the geological maps (ICGC
2016), the western part is covered by alluvial materials of granitic origin transported from the north. To the east, the basin is
275   covered by alluvial fans from slate or limestone source reaching the basin from the ENE (Cabrera et al. 1988). The seismic
data in this eastern zone do not show amplification related to the presence of sediments. Further studies will be needed to
reconcile this local difference between the geological and seismological models.
The central profile shows a sedimentary thickness reaching 700 m and a steep margin in the location of the Têt Fault limiting
the basin along its SE side (ICGC 2016). Although the depth of the basin in the central part does not fully agree with the
280   geological model along the central profile in Calvet et al. (2022), our results are consistent with the en-echelon geometry
proposed in the model to accommodate the total sedimentary thickness. However, the seismic data suggest that the sedimentary
basin reaches the Segre River, a point not documented in the geological profile.
The profile crossing the westernmost part of the CB along a north-south direction shows a basin thickening to the south and a
step discontinuity beneath the trace of the Têt Fault, an image consistent with previous models based on geological observations
285   (Cabrera et al. 1988; Calvet et al. 2022). Further analysis considering additional geological data will be needed to fully exploit
our new results, that can provide additional constraint to understand the role of the Têt Fault in defining the geometry of the
Cerdanya Basin and its present-day degree of activity.

The consistency between the results from all the methodologies used provides solid constraints to our basement depth
estimation and confirms that these approaches, quicker to obtain than tomography inversions, are a good option to assess the
geometry of sedimentary basins, especially in lightly eroded to non-eroded basins, or with dense forestation. In this sense, both
autocorrelation and seismic amplitude analysis can easily provide information on the relative thickness of the investigated
area. In order to get absolute depths, it its necessary to have some kind of additional data, as boreholes or Vs profiles. We want
to highlight that to obtain detailed 3D information of a sedimentary basin using seismic data is only possible if dense temporary
(large-N) arrays are deployed over the area as, with exception of tomography, all used approaches provide results which are
refer to the structure beneath each station. As a good knowledge of sedimentary basins is a key point to refine seismic risk
maps we conclude that the cost-benefit ratio of using large-N deployment recording background seismic noise is very positive.
We are convinced that these results will provide additional constrains to the geomorphological characterization of the basin,
as they increase significantly the knowledge of its 3D geometry.

Acknowledgments

We acknowledge the Geo3Bcn-CSIC LabSis Laboratory (http://labsis.geo3bcn.csic.es) for making their seismic stations
available for this experiment. We want to thank the local authorities that help to find appropriate and secure locations for the
instruments.

This work has benefited from open source initiatives such as Obspy (Krischer et al., 2015), SeismoRMS (Lecocq et al., 2020b)
and GMT (Wessel et al., 2013). The codes implementing WPCC and two-stage ts-PWS (Ventosa et al., 2017, 2019) are open
source under the LGPL v3 license and available at https://github.com/sergiventosa. Geological maps have been recovered
using the Macrostat platform for the aggregation and distribution of geological data (http://macrostat.org).

This is a contribution from the SANIMS project (RTI2018-095594-B-I00), funded by the Ministry of Science, Innovation and
Universities of Spain. JV has benefited from complementary funding of the ALORBE project (PIE-CSIC-202030E10).

References

Akin, Ö. and Sayil, N.: Site characterization using surface wave methods in the Arsin-Trabzon province, NE Turkey, Environ
Earth Sci, 75, 1–17, https://doi.org/10.1007/s12665-015-4840-6, 2016.

Bard, P. Y.: The SESAME project: an overview and main results, 13th World Conference on Earthquake Engineering,
Vancouver, 2004.

Benjumea, B., Macau, A., Gabàs, A., Bellmunt, F., Figueras, S., and Cirés, J.: Integrated geophysical profiles and H/V
microtremor measurements for subsoil characterization, Near Surface Geophysics, 9, 413–425,
https://doi.org/https://doi.org/10.3997/1873-0604.2011021, 2011.

320   Brocher, T. M.: Empirical Relations between Elastic Wavespeeds and Density in the Earth ' s Crust, 95, 2081–2092, https://doi.org/10.1785/0120050077, 2005.

Cabrera, L., Roca, E., and Santanach, P.: Basin formation at the end of a strike-slip fault: the Cerdanya Basin (eastern Pyrenees), Journal of the Geological Sociery, 261–268 pp., 1988.

Calvet, M., Delmas, M., Gunnell, Y., and Laumonier, B.: Geology and Landscapes of the Eastern Pyrenees, Springer

325   International Publishing, Cham, https://doi.org/10.1007/978-3-030-84266-6, 2022.

Claerbout, J. F.: Synthesis of a layered medium from its acoustic transmission response, Geophysics, 33, 264–269, https://doi.org/10.1190/1.1439927, 1968.

Delgado, J., López Casado, C., Estévez, A., Giner, J., Cuenca, A., and Molina, S.: Mapping soft soils in the Segura river valley (SE Spain): a case study of microtremors as an exploration tool, J Appl Geophy, 45, 19–32,

330   https://doi.org/https://doi.org/10.1016/S0926-9851(00)00016-1, 2000.

Díaz, J.: On the origin of the signals observed across the seismic spectrum, Earth Sci Rev, 161, https://doi.org/10.1016/j.earscirev.2016.07.006, 2016.

Diaz, J., Ruiz, M., and Jara, J. A.: Seismic monitoring of urban activity in Barcelona during the COVID-19 lockdown, Solid Earth, 12, 725–739, https://doi.org/10.5194/se-12-725-2021, 2021.

335   Diaz, J., Ruiz, M., Udina, M., Polls, F., Martí, D., and Bech, J.: Monitoring storm evolution using high density seismic arrays, PREPRINT (Version 1) available at Research Square , https://doi.org/10.21203/rs.3.rs-1803804/v1, 2022.

Efron, B. and Stein, C.: The Jackknife Estimate of Variance, The Annals of Statistics, 9, 586–596, 1981.

Ekström, G., Abers, G. A., and Webb, S. C.: Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, Geophys Res Lett, 36, https://doi.org/https://doi.org/10.1029/2009GL039131, 2009.

340   Field, E. and Jacob, K.: The theoretical response of sedimentary layers to ambient seismic noise, Geophys Res Lett, 20, 2925–2928, https://doi.org/10.1029/93GL03054, 1993.

Gabàs, A., Macau, A., Benjumea, B., Queralt, P., Ledo, J., Figueras, S., and Marcuello, A.: Joint Audio-Magnetotelluric and Passive Seismic Imaging of the Cerdanya Basin, Surv Geophys, 37, 897–921, https://doi.org/10.1007/s10712-016-9372-4, 2016.

345   Ibs-Von Seht, M. and Wohlenberg, J.: Microtremor Measurements Used to Map Thickness of Soft Sediments, Bulletin of the Seismological Society of America, 250–259 pp., 1999.

Geological map of Catalonia, 1:50,000:

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J.: ObsPy: A bridge for seismology into the scientific Python ecosystem, Comput Sci Discov, 8, 1–17, https://doi.org/10.1088/1749-4699/8/1/014003,

350   2015.

Lecocq, T., Hicks, S. P., Noten, K. van, Wijk, K. van, and Koelemeijer, P.: Global quieting of high-frequency seismic noise due to COVID-19 pandemic lockdown measures, Science (1979), 369, 1338–1343, 2020a.

Lecocq, T., Massin, F., Satriano, C., Vanstone, M., and T., M.: SeismoRMS - A simple Python/Jupyter Notebook package for studying seismic noise changes, Version 1.0, https://doi.org/10.5281/zenodo.3820046, 2020b.

355 Macau, A., Figueras, S., Colas, B., le Brun, B., Bitri, A., Susagna, T., Cirés, J., González, M., and Roullé, A.: Seismic microzonation in two valleys of the eastern Pyrenees: Andorra and the Cerdanya, 3–8 pp., 2006.

Maciel, S. T. R., Rocha, M. P., and Schimmel, M.: Urban seismic monitoring in Brasilia, Brazil, PLoS One, 16, 1–14, https://doi.org/10.1371/journal.pone.0253610, 2021.

Mcnamara, D. E., Hutt, C. R., Gee, L. S., Benz, H. M., and Buland, R. P.: A Method to Establish Seismic Noise Baselines for
360 Automated Station Assessment, Seismol. Res. Lett., 80, 628–637, https://doi.org/10.1785/gssrl.80.4.628, 2009.

Nakamura, Y.: A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface, Quarterly Report of RTI, 30, 25–33, 1989.

Pous, J., Julia, R., Sole Sugranes, L., and Frunqub, M.: Cerdanya basin geometry and its implication on the neogene evolution of the eastern Pyrenees, 355–365 pp., 1986.

365 Rawlinson, N. and Sambridge, M.: The Fast Marching Method: An Effective Tool for Tomographic Imaging and Tracking Multiple Phases in Complex Layered Media, Exploration Geophysics, 36, 341–350, https://doi.org/10.1071/EG05341, 2005.

Rivero, L., Pinto, V., and Casas, A.: Moho depth structure of the eastern part of the Pyrenean belt derived from gravity data, 33, 315–332, 2002.

Roca, E.: The Neogene Cerdanya and Seu d'Urgell intramontane basins (Eastern Pyrenees)., in: Tertiary basins of Spain: the
370 stratigraphic record of crustal kinematics, edited by: Friend, P. F. and Dabrio, C. J., Cambridge University Press, Cambridge, 114–119, 1996.

Romero, P. and Schimmel, M.: Mapping the Basement of the Ebro Basin in Spain With Seismic Ambient Noise Autocorrelations, J Geophys Res Solid Earth, 123, 5052–5067, https://doi.org/10.1029/2018JB015498, 2018.

Ruigrok, E., Campman, X., and Wapenaar, K.: Basin delineation with a 40-hour passive seismic record, Bulletin of the
375 Seismological Society of America, 102, 2165–2176, https://doi.org/10.1785/0120110242, 2012.

Schimmel, M., Stutzmann, E., and Ventosa, S.: Measuring Group Velocity in Seismic Noise Correlation Studies Based on Phase Coherence and Resampling Strategies, IEEE Transactions on Geoscience and Remote Sensing, 55, https://doi.org/10.1109/TGRS.2016.2631445, 2017.

Ventosa, S., Schimmel, M., and Stutzmann, E.: Extracting surface waves, hum and normal modes: time-scale phase-weighted
380 stack and beyond, Geophys J Int, 30–44, https://doi.org/10.1093/gji/ggx284, 2017.

Ventosa, S., Schimmel, M., and Stutzmann, E.: Towards the processing of large data volumes with phase cross-correlation, Seismological Research Letters, 90, 1663–1669, https://doi.org/10.1785/0220190022, 2019.

Vergés, J., Kullberg, J. C., Casas-Sainz, A., de Vicente, G., Duarte, L. V., Fernàndez, M., Gómez, J. J., Gómez-Pugnaire, M. T., Jabaloy Sánchez, A., López-Gómez, J., Macchiavelli, C., Martín-Algarra, A., Martín-Chivelet, J., Muñoz, J. A., Quesada,
385 C., Terrinha, P., Torné, M., and Vegas, R.: An Introduction to the Alpine Cycle in Iberia, in: The Geology of Iberia: A

Geodynamic Approach: Volume 3: The Alpine Cycle, edited by: Quesada, C. and Oliveira, J. T., Springer International Publishing, Cham, 1–14, https://doi.org/10.1007/978-3-030-11295-0_1, 2019.

Wapenaar, K.: Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, Phys Rev Lett, 93, https://doi.org/10.1103/PhysRevLett.93.254301, 2004.

390 Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., and Wobb, F.: Generic Mapping Tools ., Eos (Washington DC), 94, 2013.

Yu, Y., Song, J., Liu, K. H., and Gao, S. S.: Determining crustal structure beneath seismic stations overlying a low-velocity sedimentary layer using receiver functions, J Geophys Res Solid Earth, 120, 3208–3218, https://doi.org/10.1002/2014JB011610, 2015.

Zelt, B. C. and Ellis, R. M.: Receiver-function studies in the Trans-Hudson Orogen, Saskatchewan, Can. J. Earth Sci, 585–603 395 pp., 1999.

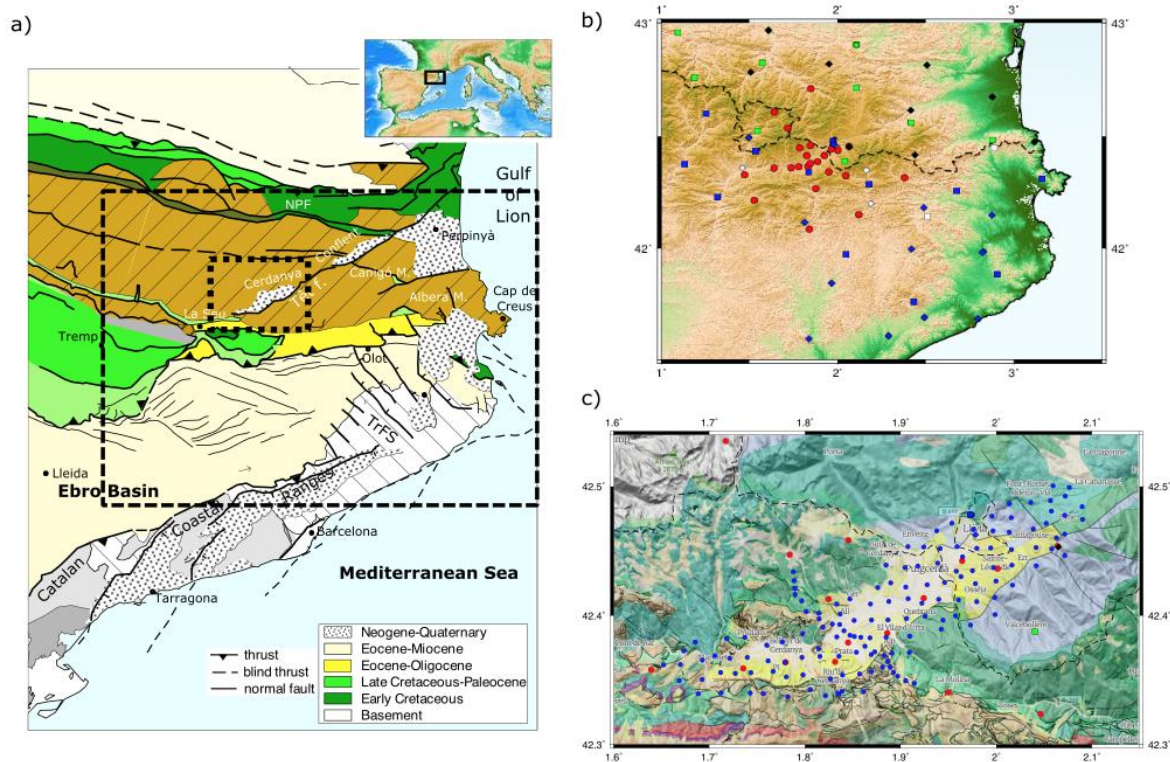



Figure 1: a) Simplified tectonic map of the Eastern Pyrenees, adapted from Vergés et al. (2019) including the main Pyrenean thrusts and Neogene extensional faults. Dashed squares show the location of the maps in b) and c). b) Deployment of the SANIMS broad-band stations (red dots). Permanent broad-band (squares) and accelerometric (diamonds) are included for reference. Blue: CA network; White: ES network; Green: FR network; Black: RA network. The inset map shows the location of the investigated area. Background topography is derived from the Shuttle Radar Tomography Mission (SRTM) and has a resolution of 1 arc-second. c) High density seismic nodes (blue dots) deployed between April and June 2021. Red dots show the location of the previous broad-band deployment. The background shows the ICGC geological map around the Cerdanya Basin (ICGC 2016).

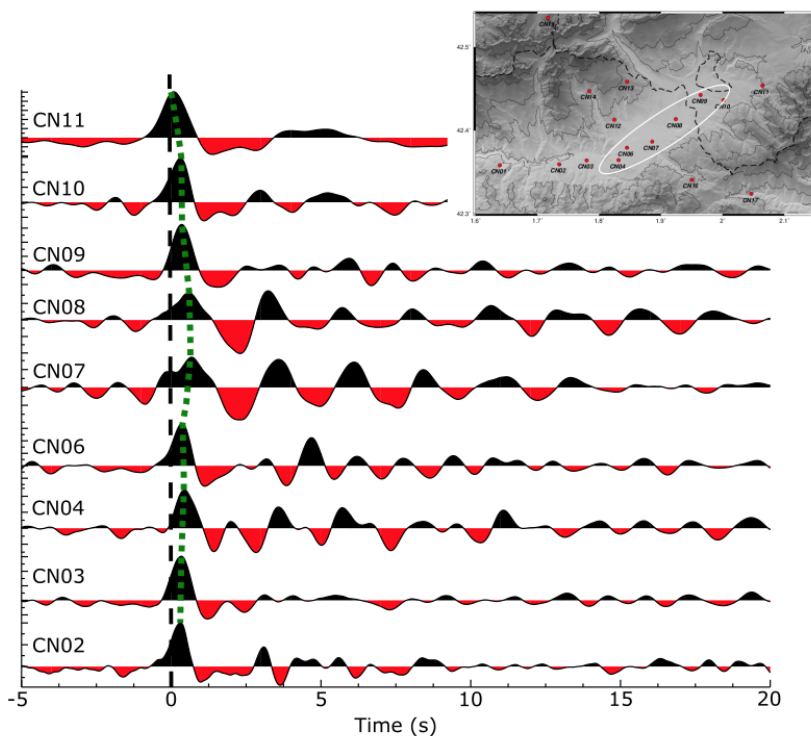



Figure 2: Location map and stacked RFs for the broad-band stations located along the Cerdanya Basin. Dotted green line show the direct P-wave time lag for stations along the basin. Large reverberations are clearly observed for stations CN07 and CN08. The inset map shows the area with anomalous RFs.

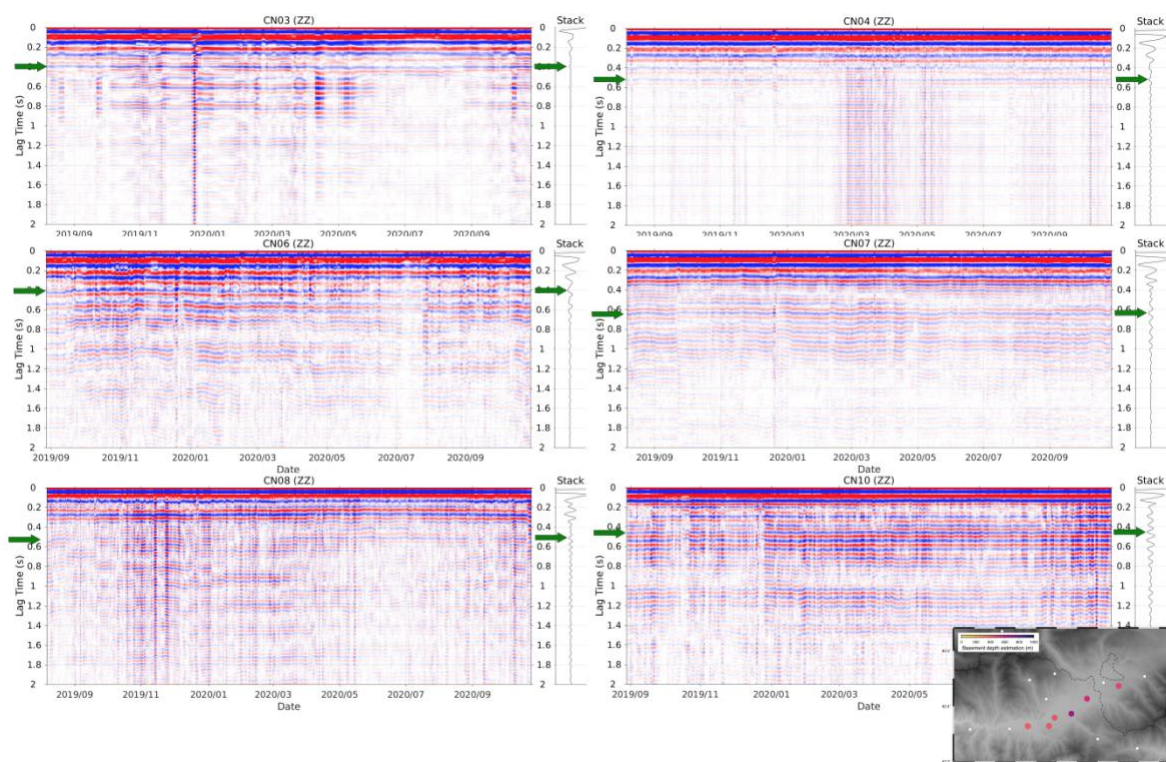



Figure 3: Daily autocorrelograms for broad-band stations CN03, CN04, CN06, CN07, CN08 and CN10, all located along the
415 Cerdanya Basin. Dark green arrows show the reflectors interpreted as corresponding to the basement. Vertical axis refers to
the two-way travel time (s). Traces are ordered by date, with the total stack shown beside each panel. The inset map shows the
basement depth estimations.

420

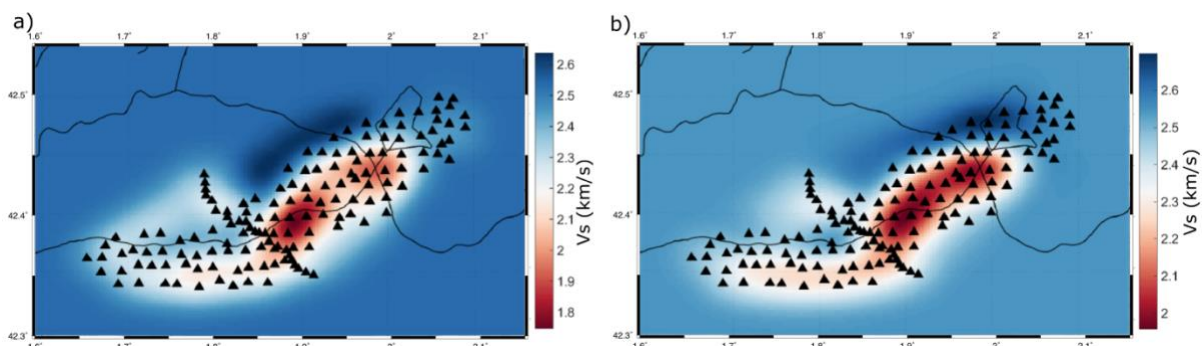
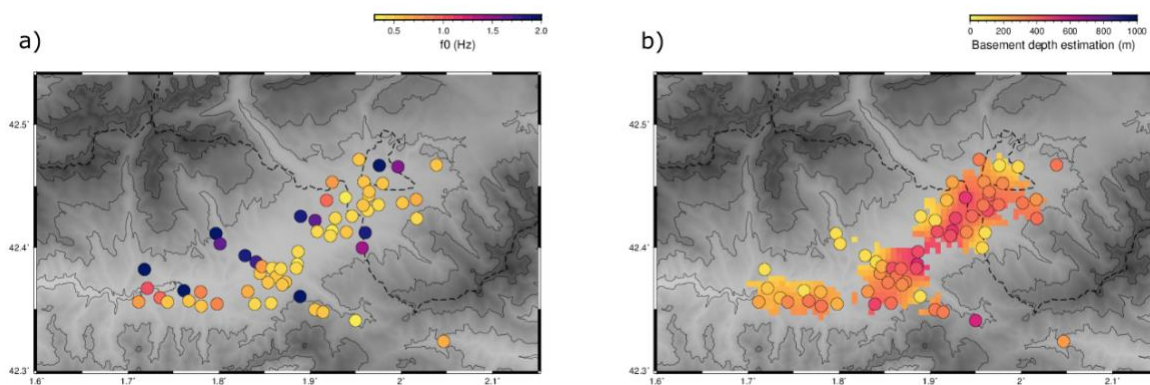



Figure 4: Rayleigh phase absolute velocity maps derived from ANT using the high-resolution array for the shorter periods available, 1.035 (a) and 1.414 s (b).

425 **Figure 5:** a) f0 values retrieved from broad-band stations and seismic nodes. b) Basement depths estimated from the f0 values using the scaling law proposed by Gabàs et al. (2014).

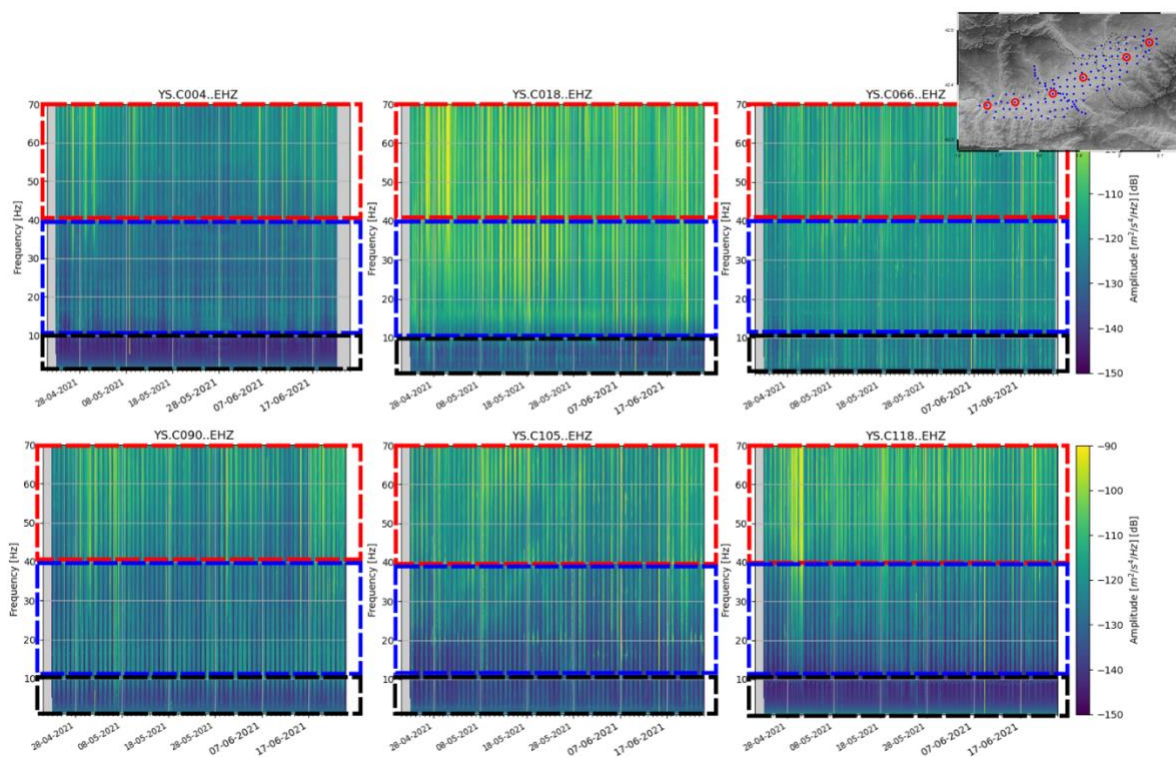



Figure 6: Spectrograms for stations distributed along the basin. Red and blue boxes show the frequency bands dominated by meteorologic and anthropogenic sources. Black boxes outline the frequency band related to site amplification

.

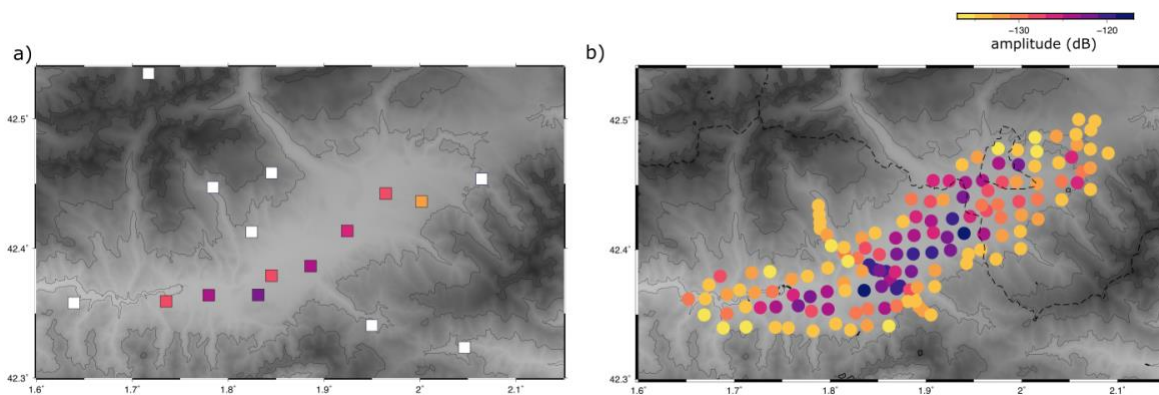



Figure 7: a) Seismic power amplitude for the BB stations b) seismic power amplitude for the dense nodal deployment. The color palette represents the median amplitude in the 1-10 Hz band, measured in dB. White squares are for broad-band stations with lower median amplitudes (out-of-scale).

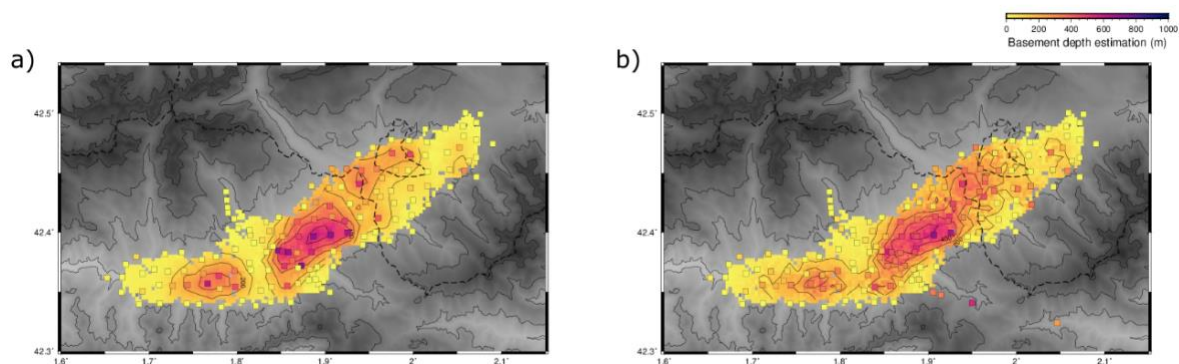



Figure 8: Basement depths inferred from seismic amplitudes in the 1-10 Hz band (a) and including also the HVSR-derived values (b).

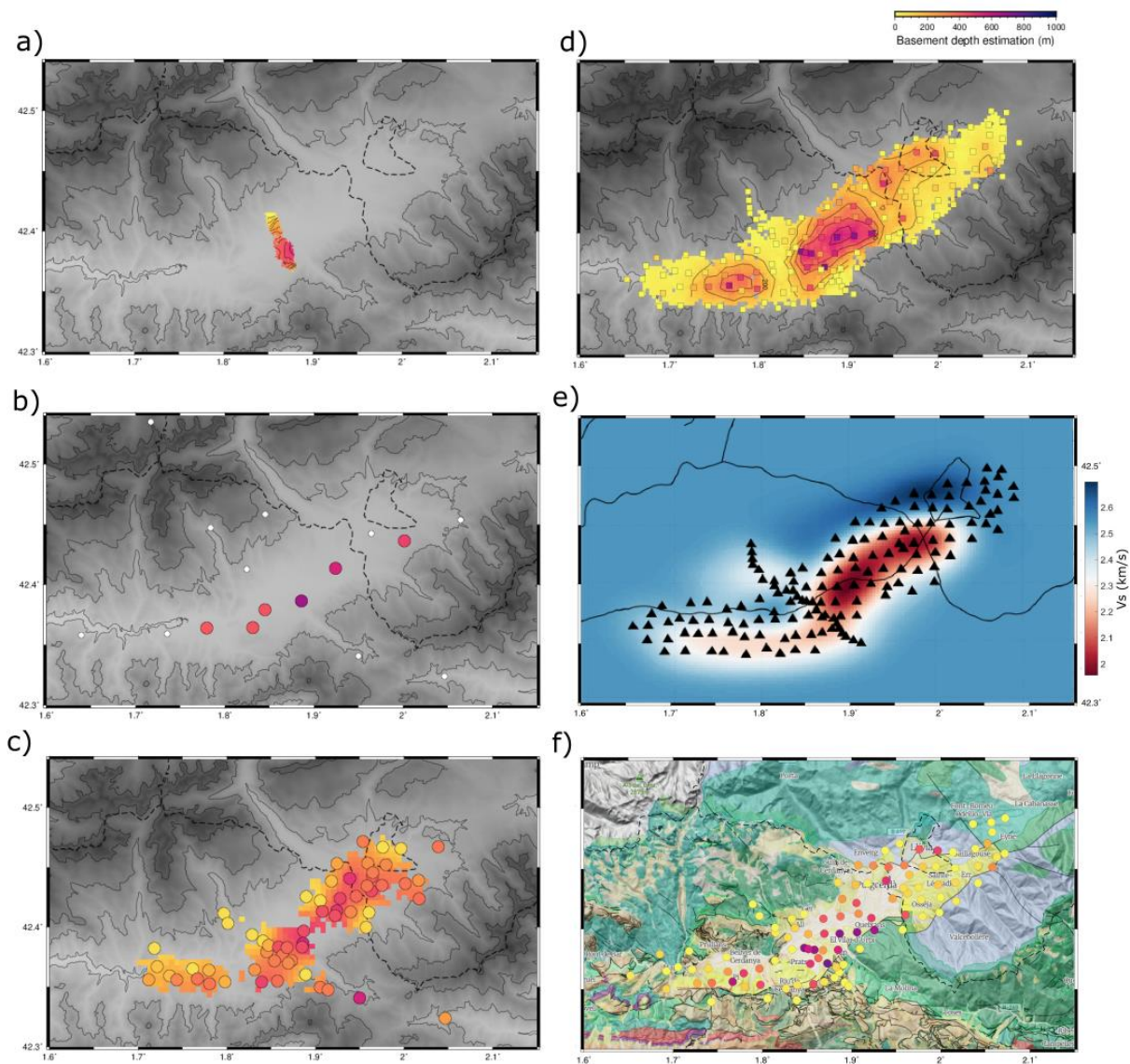



Figure 9: a) Basement depth estimations from Gabàs et al. 2016. b) Basement depth estimations from autocorrelations at these locations. c) Basement depth estimations from HVSR analysis d) Basement depth estimations from noise amplitude analysis. e) Shear-wave velocity map retrieved from ANT for a period of 1.414 s. f) Geological map (ICGC 2016) overprinted by the basin depth values derived from seismic noise. .

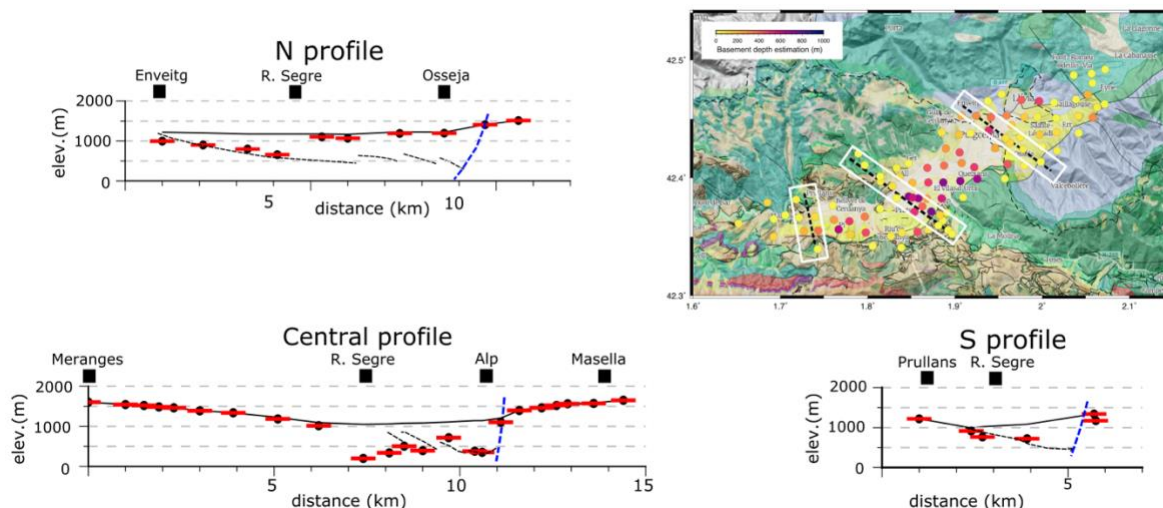



Figure 10: Profiles along the three main domains across the basin. X-axis is the distance along the profile and Y-axis show the topographic elevation (black lines) and the depth of the bottom of the sedimentary basin inferred from seismic noise amplitude at each site (red dashes). The location of the profiles and the depth estimations at the sites along the profiles are shown overprinting the accompanying geological map (ICGC 2016). Black dashed lines represent the base of the sedimentary basin in the models by Calvet et al. 2022. Blue dashed lines show the location of the Têt Fault, projected from the geological cross-section by Calvet et al. (2022).