Supplement

S1 Additional information on data

In the section 'Data processing and decay properties of coda waves', we introduced the data of the Dense Array of North Anatolia, which has 73 three-component stations. In both our Q_c measurements and the MLTW analysis we average the horizontal component pairs before we analyse the data. This is because the vertical components look rather noisy, which we can observe in Fig. S1. The vertical components (last column and last row) show a strong band of energy near t = 0, which is most likely related to tele-seismic body waves.

Analysis of the power spectra of the horizontal components in all four sub-regions show that in the frequency band 0.1-0.5 Hz, most energy is around 0.3 Hz (Fig. S2). With a group velocity of $\sim$ 2.1 km/s, the dominant wavelength in our data is therefore $\sim$ 7 km. The scattering mean free path in the fault zone, in the order of 11 km, is thus larger than the dominant wavelength.

S2 Additional information on Q_c^{-1}

Our measurements of the coda decay are done on the envelopes of the averaged horizontal components. Fig. S3 and S4 show the Q_c measurements of the individual component pairs for the anti-causal and causal parts, respectively. The stations are alphabetically ordered, which implies a geographical ordering (see Fig. 1 of the manuscript). We observe that the maps show similar results, implying no obvious bias in the measurements due to directivity of the noise source. This is as expected, because the waves that form the coda are theoretically, especially in the later parts, less affected by the directivity of propagation and thus by the component pair.

S3 Additional information on scattering properties Q_i and ℓ

S3.1 Effect of Velocity

This section explores the correct surface wave group velocity per sub-region. The initial value is derived from the first arrivals on the CCFs. In these tests we change the velocity in the RTE and for the estimates of the NEDs and compare the misfits with the misfits we have for the initial velocities. The range of velocities we test are -0.7 to +0.3 km/s with steps of 0.1 km/s w.r.t. the initial velocity. For the northern sub-region (Fig. S5) using a lower velocity minimises the minimum misfit slightly and the resulting ℓ is higher, but still not better constrained. No change for Q_i compared to the initial velocity. For the central region (Fig. S6), using a 0.6 km/s lower velocity the optimal Q_i is slightly lower and ℓ is higher but still not better constrained. Lastly, for the southern sub-region (Fig. S7), we find a smaller minimal misfit using a velocity that is 0.2 km/s slower. However, the found Q_i and ℓ do not change. Fig. S8 - S10 show the velocities derived from the first arrivals of the CCFs and the ones that minimizes the misfit.

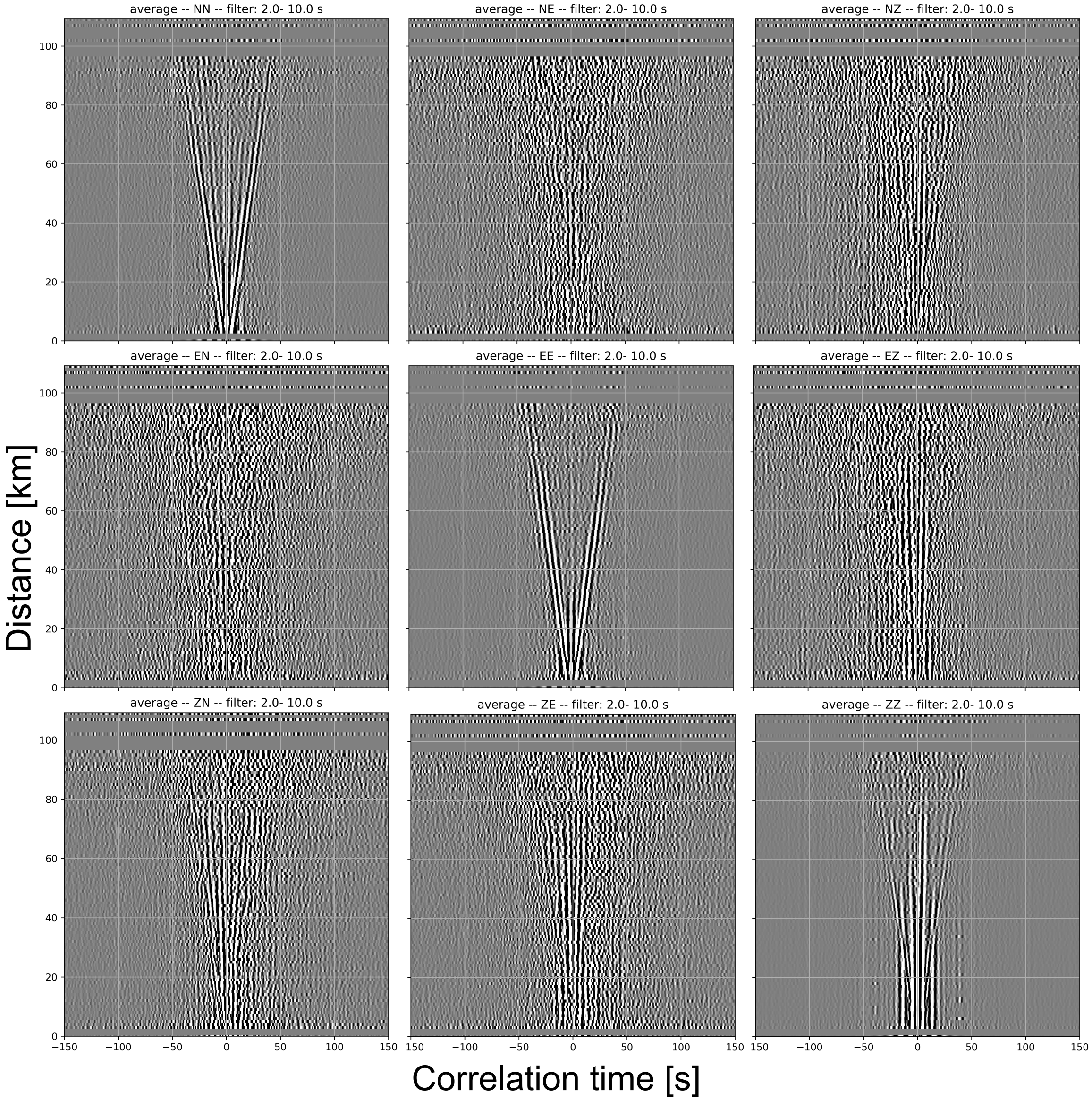

Figure S1: Bin-stacked cross-correlation functions (CCFs) calculated for all inter-station paths of the DANA network, for all three-component combinations. The data used to compose the CCFs were filtered between 0.1 and 0.5 Hz; the resulting CCFs are binned, stacked and normalised for display purposes.

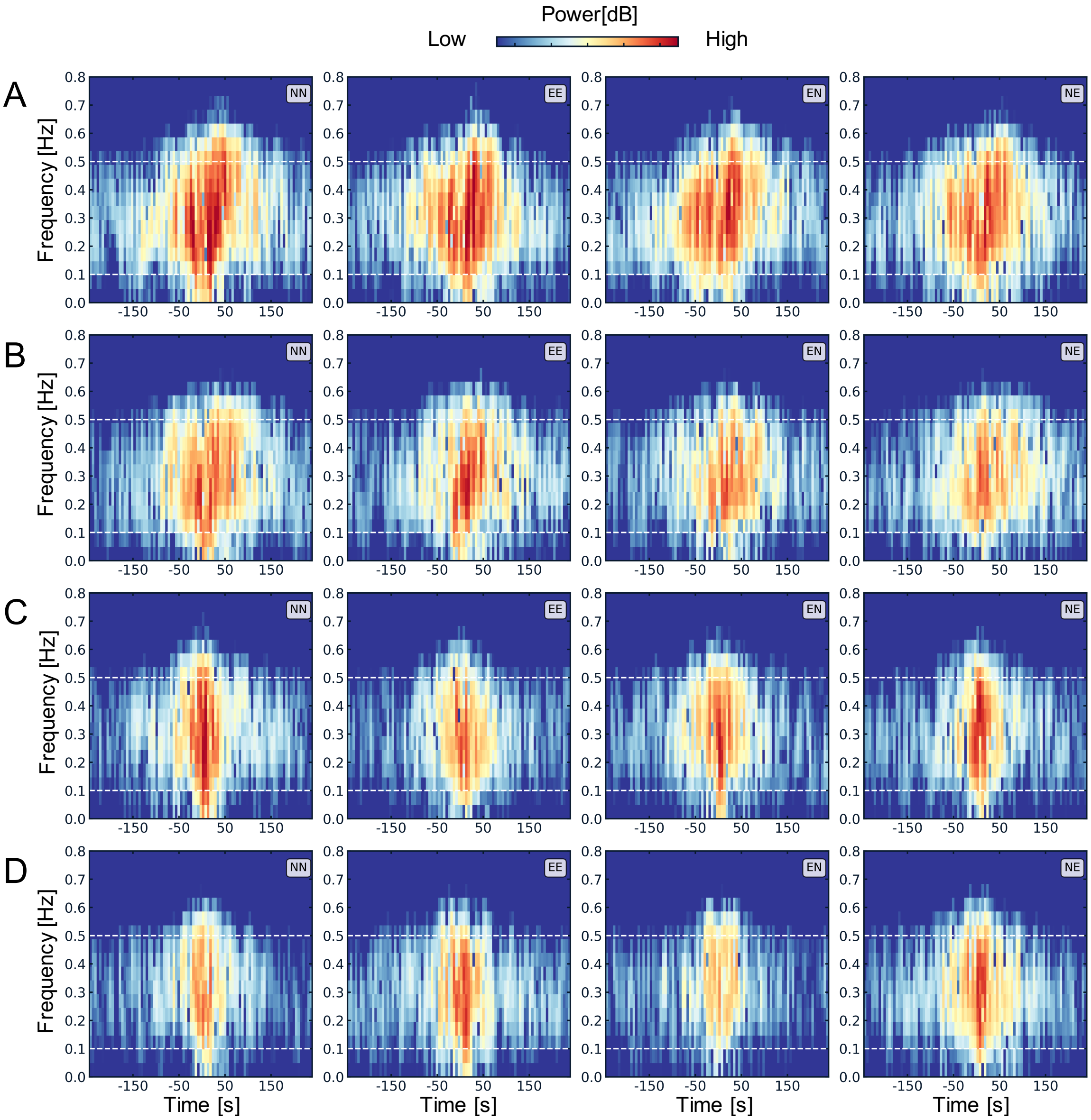

Figure S2: Spectrograms of the horizontal component pairs for the four sub-regions: A) north, B) fault zone, C) centre and D) south. Colder colours for less energy. The white dotted vertical lines show the frequency band for which we filtered the data (0.1 and 0.5 Hz).

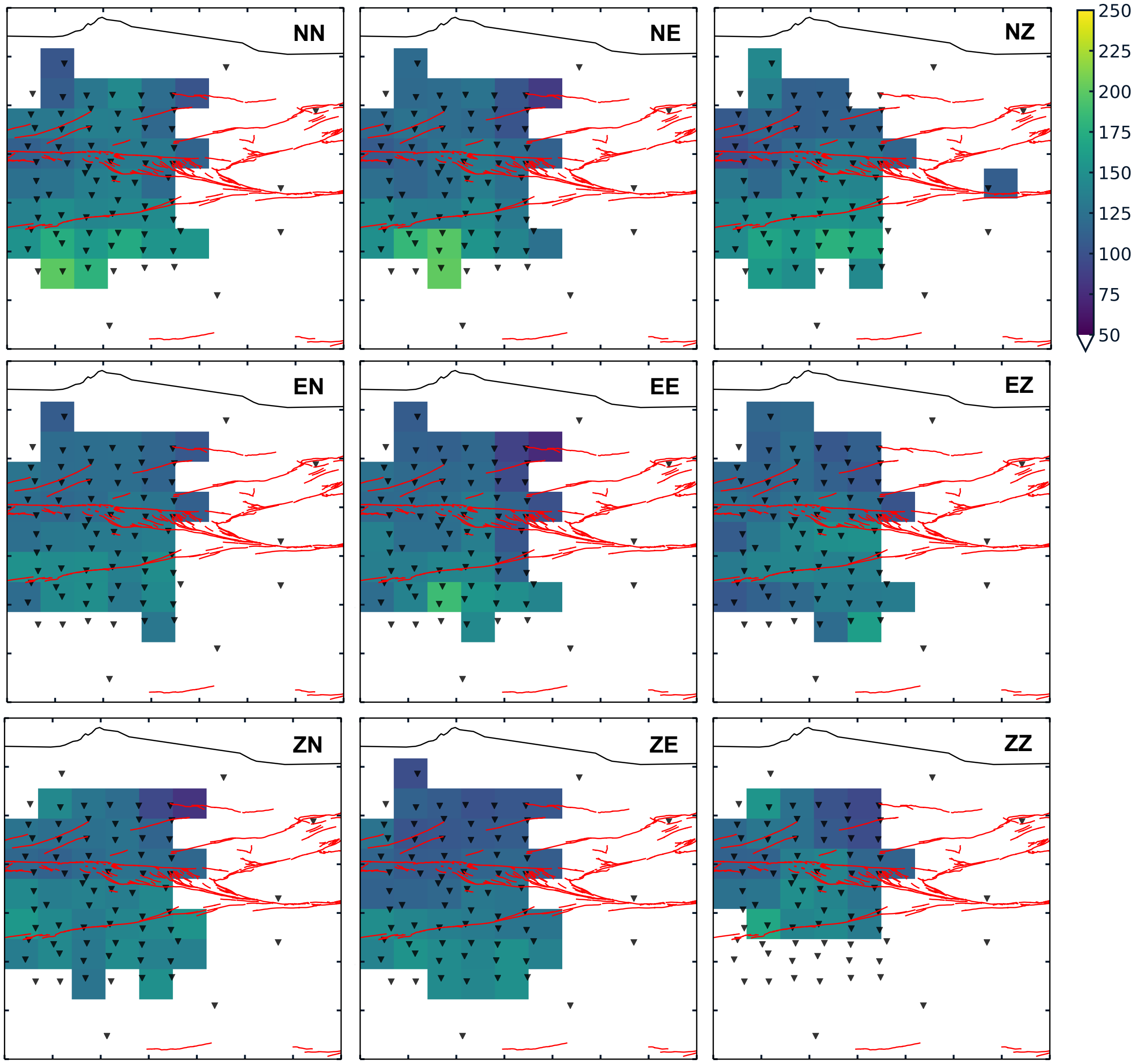

Figure S3: Q_c maps for anti-causal part of CCFs for all component combinations. Each cell shows the inverse of the arithmetic mean of the measured Q_c^{-1} values. Only cells with a minimum number of 5 rays per cell are shown. Other selection parameters are the interstation distance < 35 km and a correlation coefficient of the linear regression, $R^2 > 0.75$.

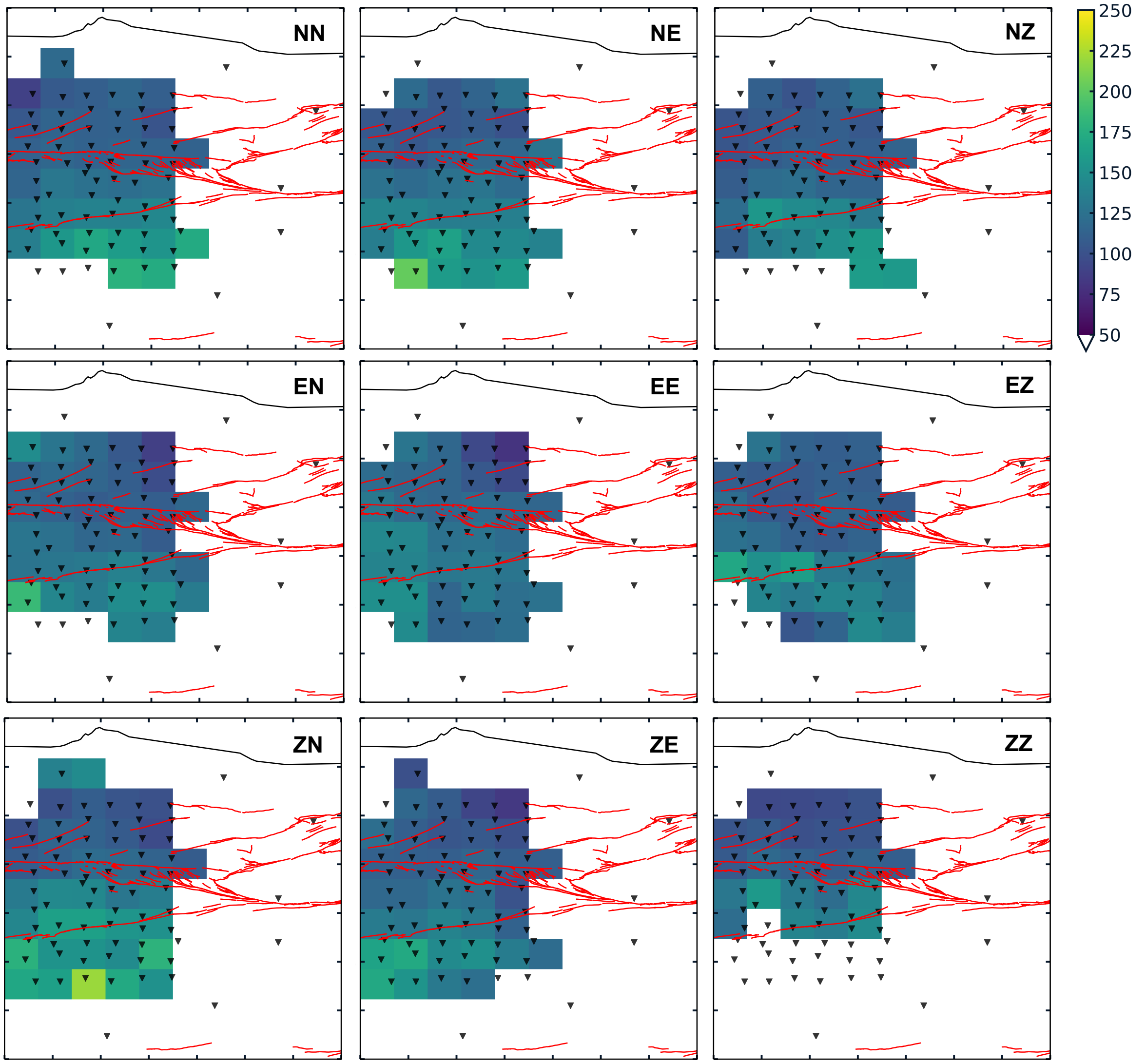

Figure S4: Q_c maps for causal part of CCFs for all component combinations. Each cell shows the inverse of the arithmetic mean of the measured Q_c^{-1} values. Only cells with a minimum number of 5 rays per cell are shown. Other selection parameters are the interstation distance < 35 km and a correlation coefficient of the linear regression, $R^2 > 0.75$.

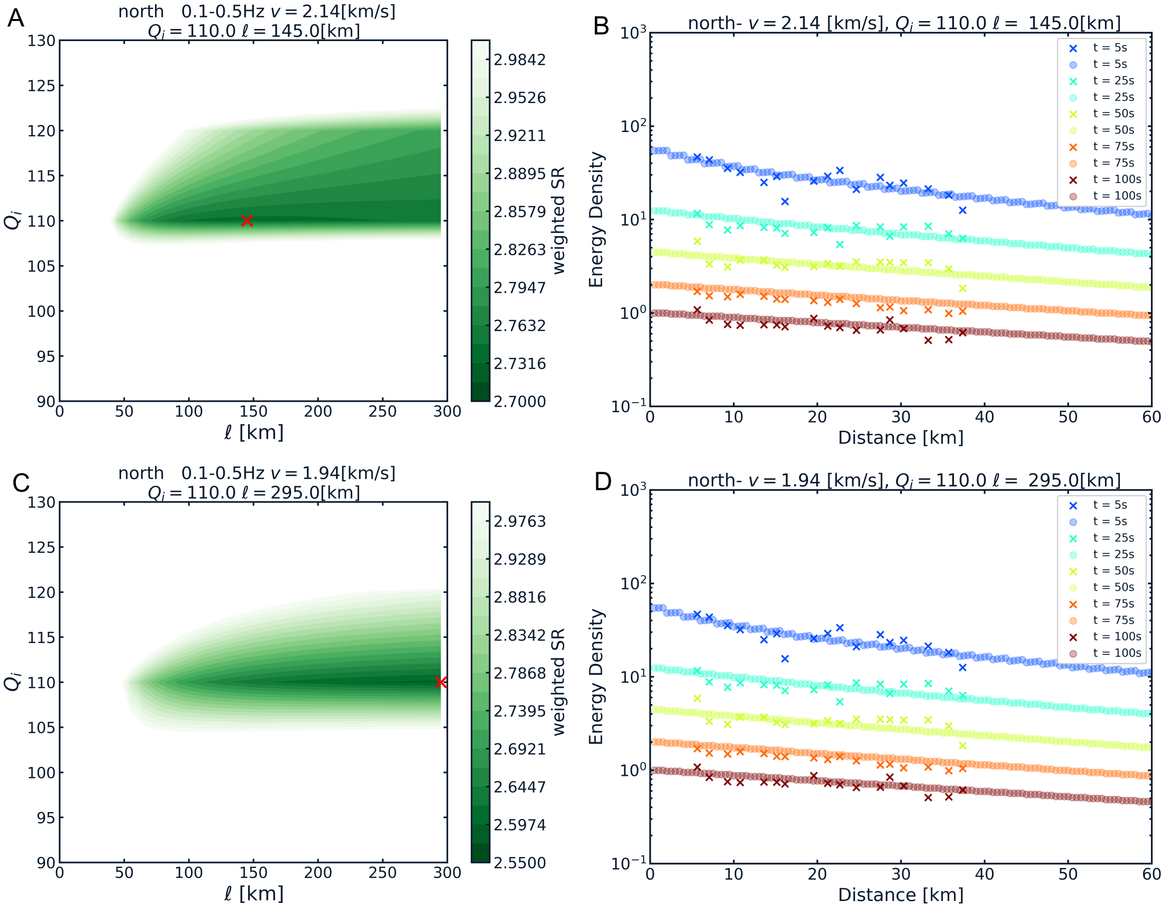

Figure S5: Velocity test results for northern sub-region. The left column shows the SM and the right column the observations ('X') with solution for the 2D-RTE ('o') using the optimal parameters. The colours indicate the time windows. A)-B) are the misfit using the original velocity derived from the arrivals on the data. C)-D) for a velocity that is 0.2 km/s less, showing the smallest minimum weighted misfit between observations and 2D RTE.

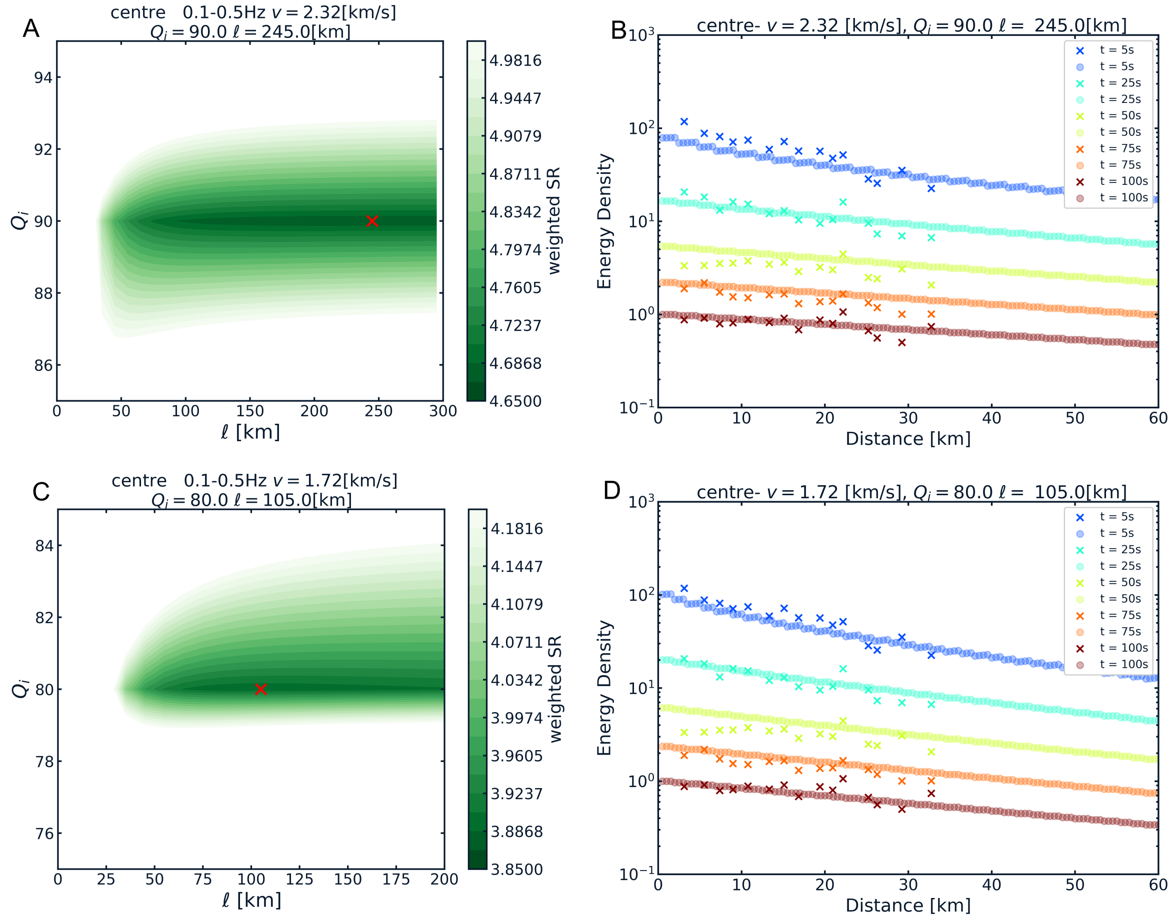

Figure S6: Velocity test results for the central sub-region. The left column shows the SM and the right column the observations ('X') with solution for the 2D-RTE ('o') using the optimal parameters. The colours indicate the time windows. A)-B) are the misfit using the original velocity derived from the arrivals on the data. C)-D) for a velocity that is 0.6 km/s less, showing the smallest minimum weighted misfit between observations and 2D RTE.

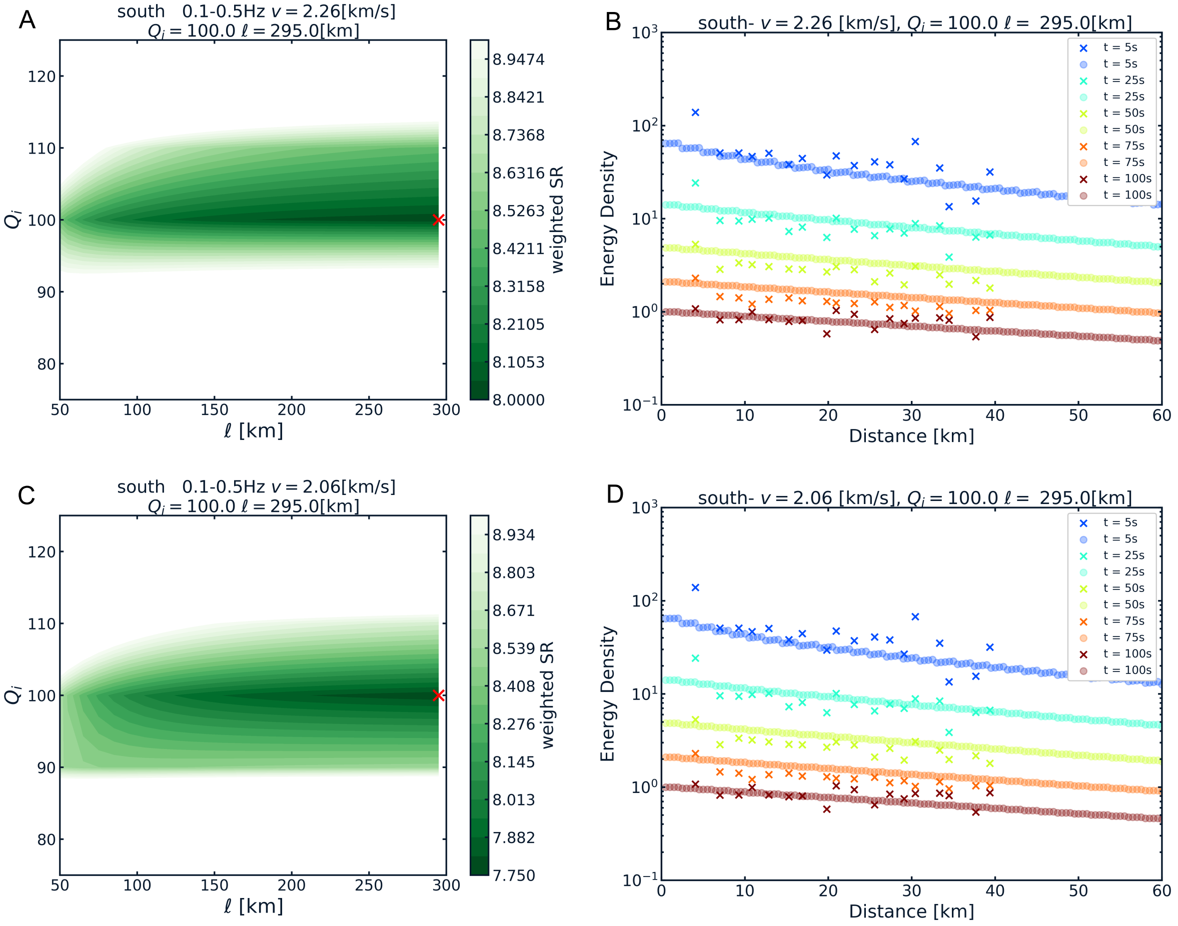

Figure S7: Velocity test results for the southern sub-region. The left column shows the SM and the right column the observations (‘X’) with solution for the 2D-RTE (‘o’) using the optimal parameters. The colours indicate the time windows. A)-B) are the misfit using the original velocity derived from the arrivals on the data. C)-D) for a velocity that is 0.2 km/s less.

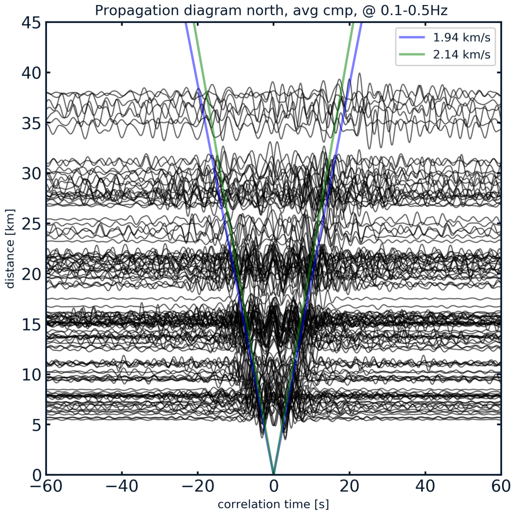

Figure S8: Cross-correlations versus distance diagram for the northern sub-region, with both the initial group velocity (green) and the velocity that gives the smallest misfit (blue) plotted on top.

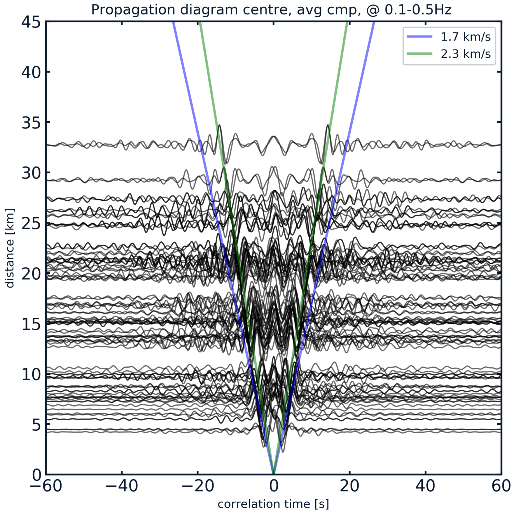

Figure S9: Cross-correlations versus distance diagram for the central sub-region, with both the initial group velocity (green) and the velocity that gives the smallest misfit (blue) plotted on top.

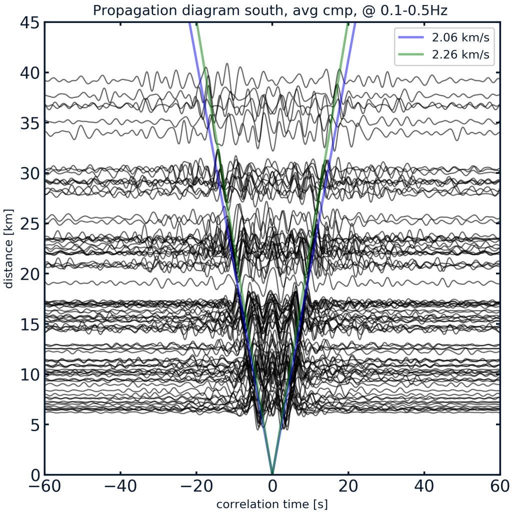

Figure S10: Cross-correlations versus distance diagram for the southern sub-region, with both the initial group velocity (green) and the velocity that gives the smallest misfit (blue) plotted on top.

S3.2 Effect of Noise

The effect of SNR of the CCFs (or fluctuations in the NEDs caused by e.g. back-scattering effects of the fault zone) on the optimisation is analysed by a synthetic inversion test. One group of tests is done in case of strong scattering attenuation, similar to the situation inside the fault zone. The other group of tests for weak scattering attenuation, simulating the situation outside the fault zone. The analytical solution of the RTE is compared with a 'noise-induced' solution of the RTE, in a similar way as the optimisation in the previous section where the observed data is compared to the modelled data. The 'noise' represents the fluctuations in the binned NED values. This 'noise' that we add to the solution, is within one standard deviation (STD) of the individually measured NEDs before binning. A random value within the range of ± 1 STD per distance bin of the FZ is used for the strong scattering attenuation tests. In a similar fashion, random values within the range of ± 1 STD of the central zone are taken as 'noise' for the test of weak scattering attenuation. Note that drawing a random value from the ± 1 STD range may underestimate the actual fluctuations we observe in the median NED values, because effectively it simulates a random value in the range of $\pm$ STD/$\sqrt{3}$.

For the strong scattering attenuation case, Q_i^{-1} and ℓ are well-constraint. This applies generally when ℓ is small w.r.t. the aperture of the array. As the scattering mean free path increases, the resolution on ℓ gets lower, i.e. only a range of potential values can be extracted. For 'too high' ℓ values, again w.r.t. the array size, the energy does not decay rapidly enough anymore. The distance dependence can therefore not be measured and ℓ cannot be resolved for. The former case seems to apply for ℓ less than 80 km with the current network, and the latter for $\ell > 150$ km. Fig. S11A shows an example of the resulting misfits and plots for small ℓ values.

In case of weak scattering attenuation, it is hard to constrain ℓ as can be observed in

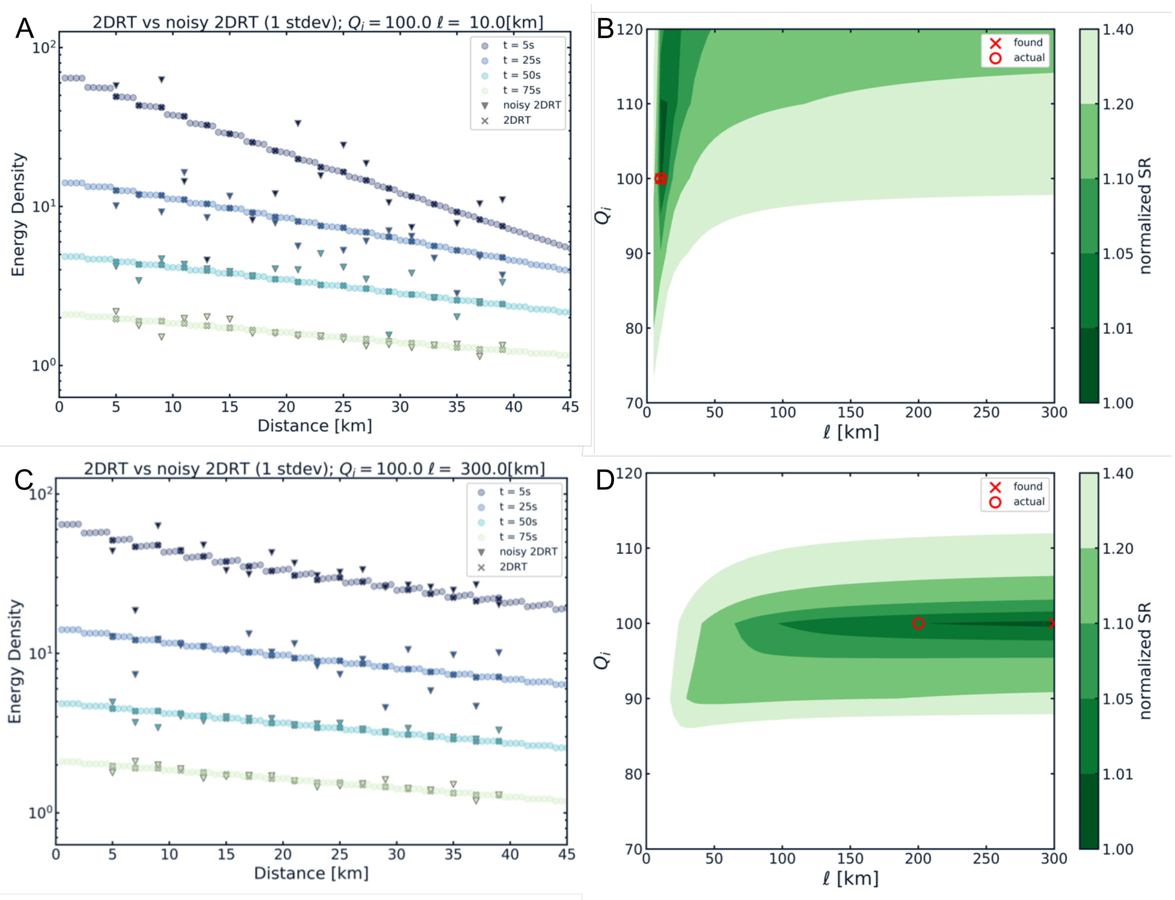

Figure S11: 'Noise' test to analyse the influence of fluctuations on the optimisation. A)-B) Case of rapid energy decay, simulating the situation inside the fault zone. C)-D) Case of slow energy decay, simulating the situation outside the fault zone. Panel A and C) show the NEDs for the 2D RTE as 'o' and the 'noisy' NEDs as triangles. The colours indicate the time windows; lighter colours for later times. Panel B) and D) show the normalised SR of the grid searches, with the 'O' indicating the value used in the non-noisy 2D RTE, and the 'X' shows the best fitting value between 'noisy' and non-noisy 2D RTE.

Fig. S11B. Q_i^{-1} on the other hand is still decently constrained. The rate of the energy decay with distance is small and therefore the 'noise' can be a dominant factor in the average spatio-temporal energy decay of the sub-region. Hence, for some regions it seems impossible to constrain the scattering mean free path properly with the current network. This is most plausibly the reason why no clear constraint on ℓ can be found for the sub-regions outside the fault zone.

S4 Additional information on Monte Carlo simulations

S4.1 Uniform models

In the first simulations we use a uniform model with Q_i and ℓ as derived in the optimisation on the actual data (Section 'Mapping of attenuation properties') to confirm the validity of the simulations. Since the output of these simulations is the total intensity (and after division by the seismic velocity, the energy density) at a chosen location with time, we performed a similar optimisation as we did on the actual observations. This provides us with the opportunity to compare not only the derived scattering properties, but also the measured energy densities per time and distance.

For all four simulations we were able to retrieve the Q_i value accurately and precisely (see Fig. S12). However, the scattering mean free path proves more difficult to retrieve outside the fault zone, as explained in section 'Mapping of attenuation properties'.

S4.2 Test on Uniform Q_i

To exclude the effect of Q_i in the simulations, we aim to find one single Q_i value for the entire model space, in order to simplify the other tests by taking one variable out of the equation. This is especially important when examining the influence of the ratio between the scattering mean free path and the width of the FZ. For these test we use a non-uniform model space, one that is divided into four sub-zones, and a E-W receiver line configuration (as used in the main text). The ℓ values in the model space are as derived from the observations. We test three different values for Q_i: 80, 90 and 100, and compare the results to the observed NEDs. The test results are shown in Fig. S13, S14, S15; panels A-D, correspond to the results in the north, FZ, centre and south respectively. The highest degree of agreement is when using a Q_i value of 100, we therefore chose to continue with this value for the tests from Section 'Constraints on the fault zone width' onward.

S4.3 Non-normalised simulations compared

redIn the section 'The signature of a finite width scattering zone', we discuss that the energies at absolute long times for the homogeneous and heterogeneous case are very close. This is shown by the results for the last time window in Fig. S16.

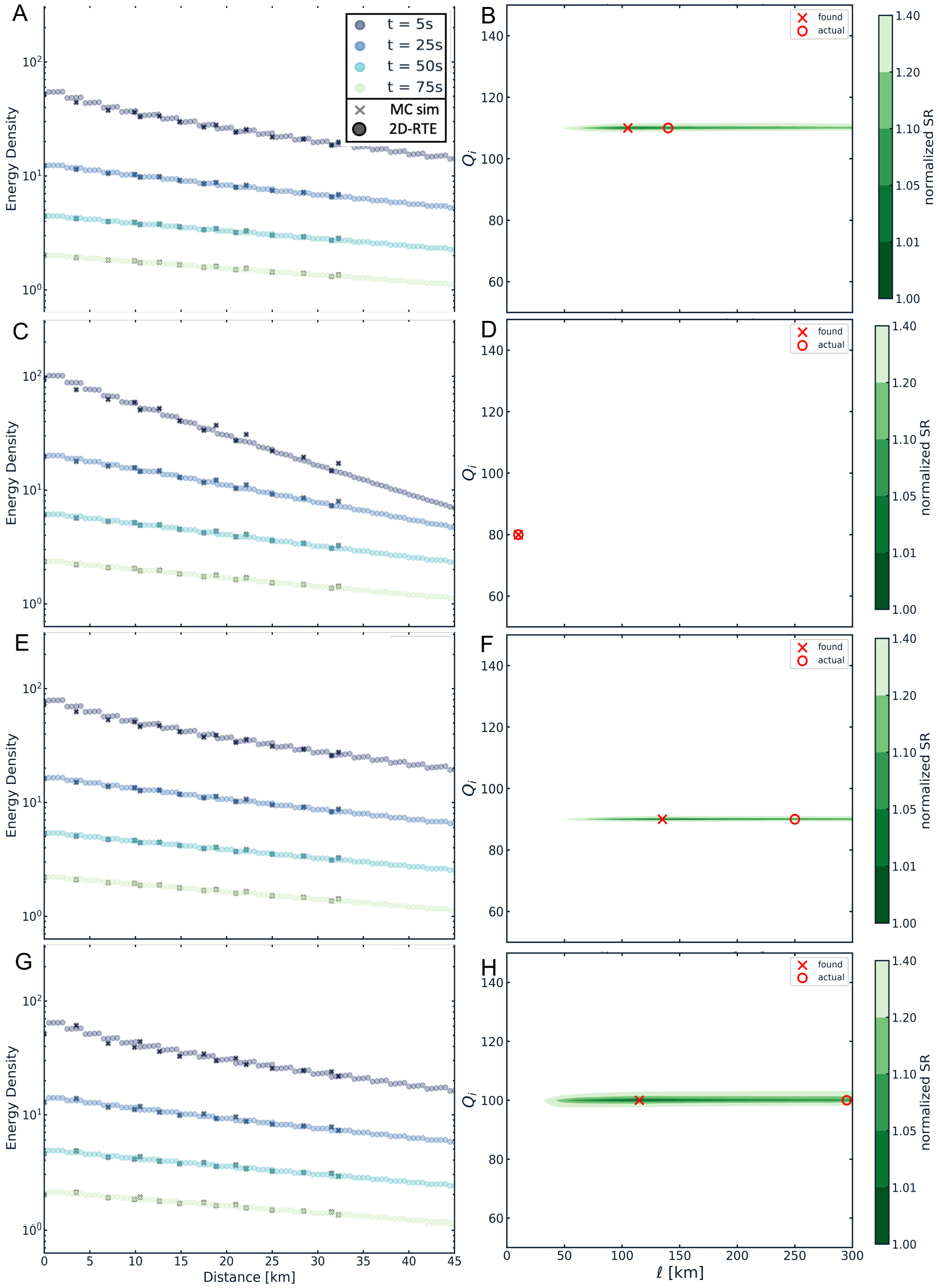

Figure S12: Optimisation results of simulated data for homogeneous models with scattering parameters from optimisation per sub-region. The figures in the left column show measured normalised energy density for simulation (‘x’) versus normalised energy density of 2D RTE (‘o’). The colour of both the X and O indicate the time window as specified in the legend. The right column shows the misfit and best ℓ and Q_i values found for simulations (‘X’), the red circle indicates the values used for the simulation. A-B are for the homogeneous model with values of the northern sub-region. C-D, E-F and G-H for fault zone, central and south respectively. NB: these values deviate slightly form the values shown in Fig. 7 of the main manuscript because these are for a grid search performed at lower resolution, so ΔQ_i= 10 instead of 2 and $\Delta\ell = 5$ instead of 2 km.

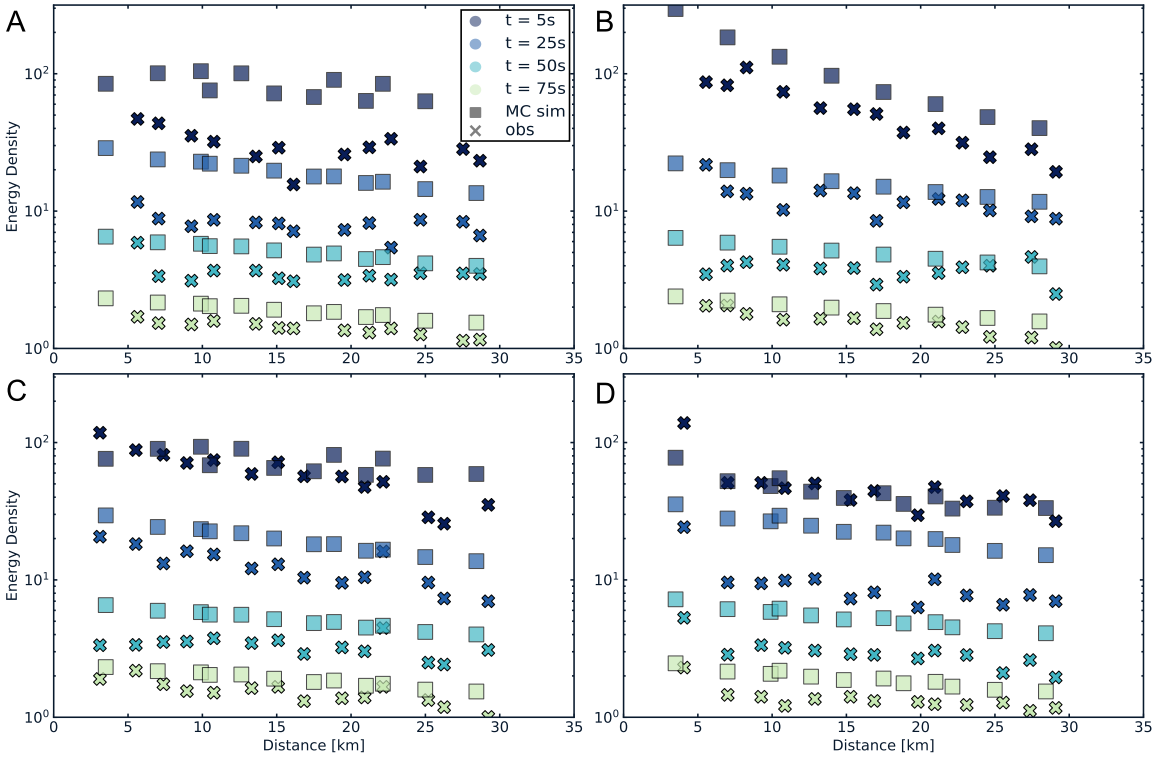

Figure S13: Normalised energy densities of simulation(squares) versus observations ('X') for a simple model using one Q_i=80 for the all four sub-zones and ℓ= 150 km outside the FZ (A,C,D) and ℓ= 10 km the FZ (B).

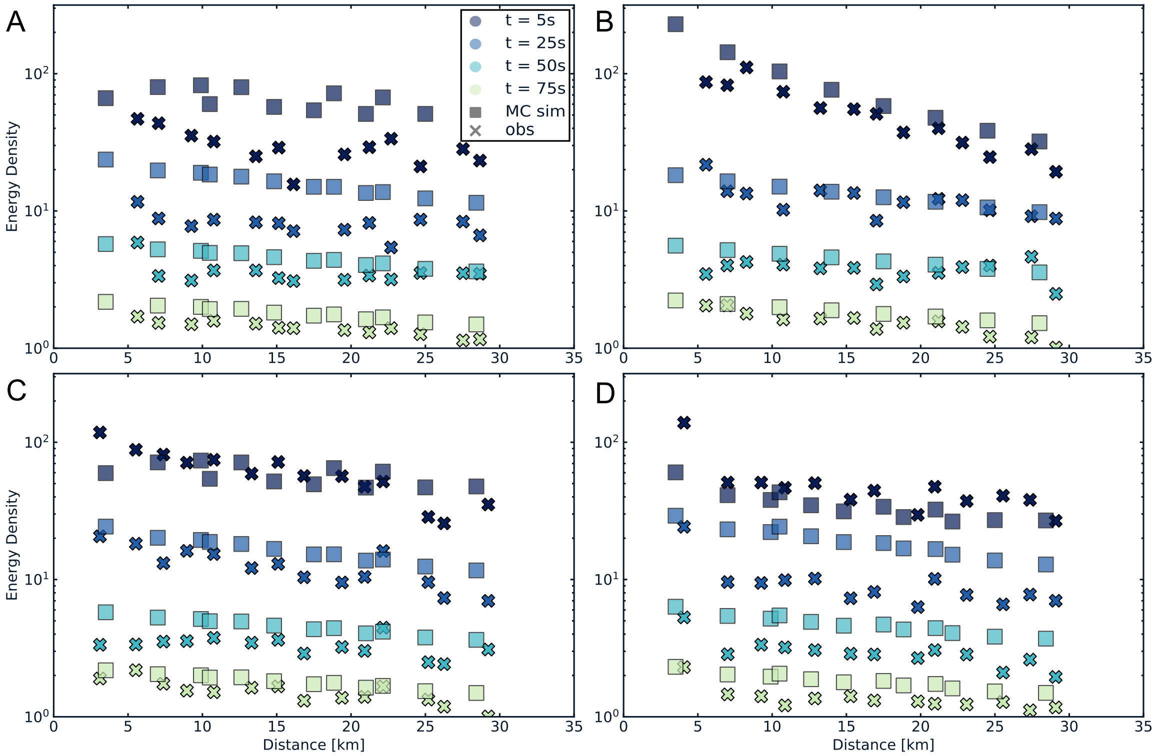

Figure S14: Normalised energy densities of simulation(squares) versus observations ('X') for a simple model using one Q_i=90 for the all four sub-zones and ℓ= 150 km outside the FZ (A,C,D) and ℓ= 10 km the FZ (B).

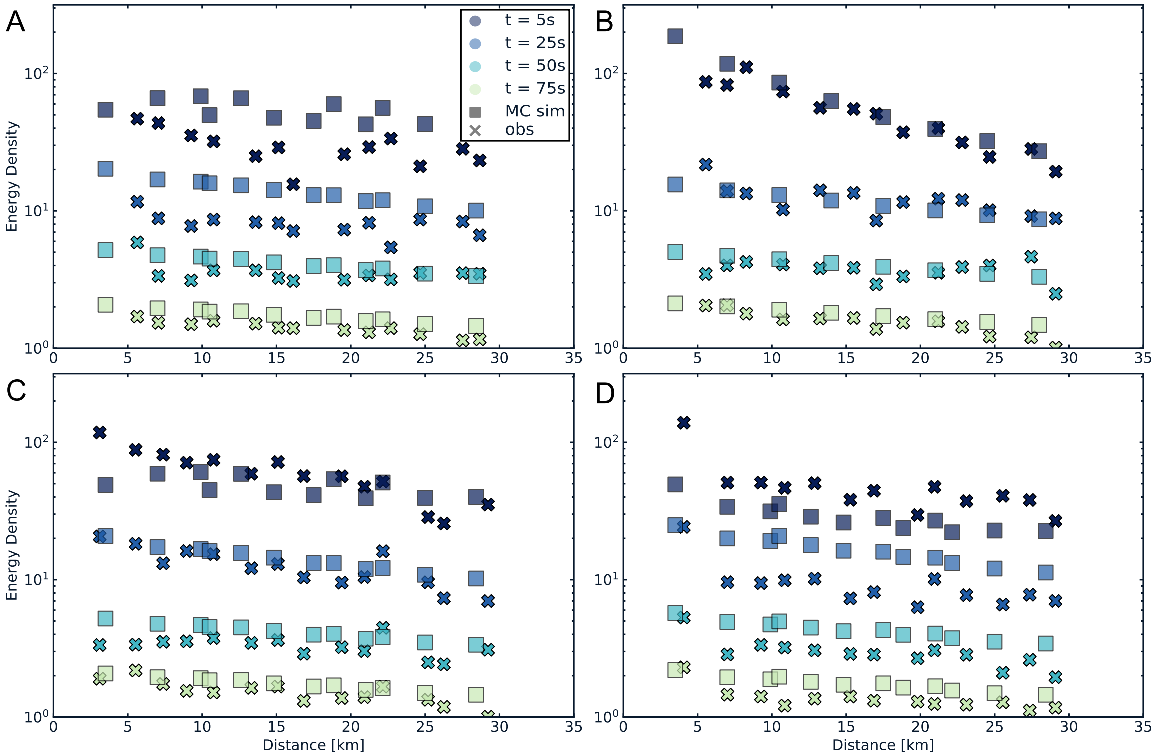

Figure S15: Normalised energy densities of simulation(squares) versus observations ('X') for a simple model using one Q_i=100 for the all four sub-zones and ℓ= 150 km outside the FZ (A,C,D) and ℓ= 10 km the FZ (B).

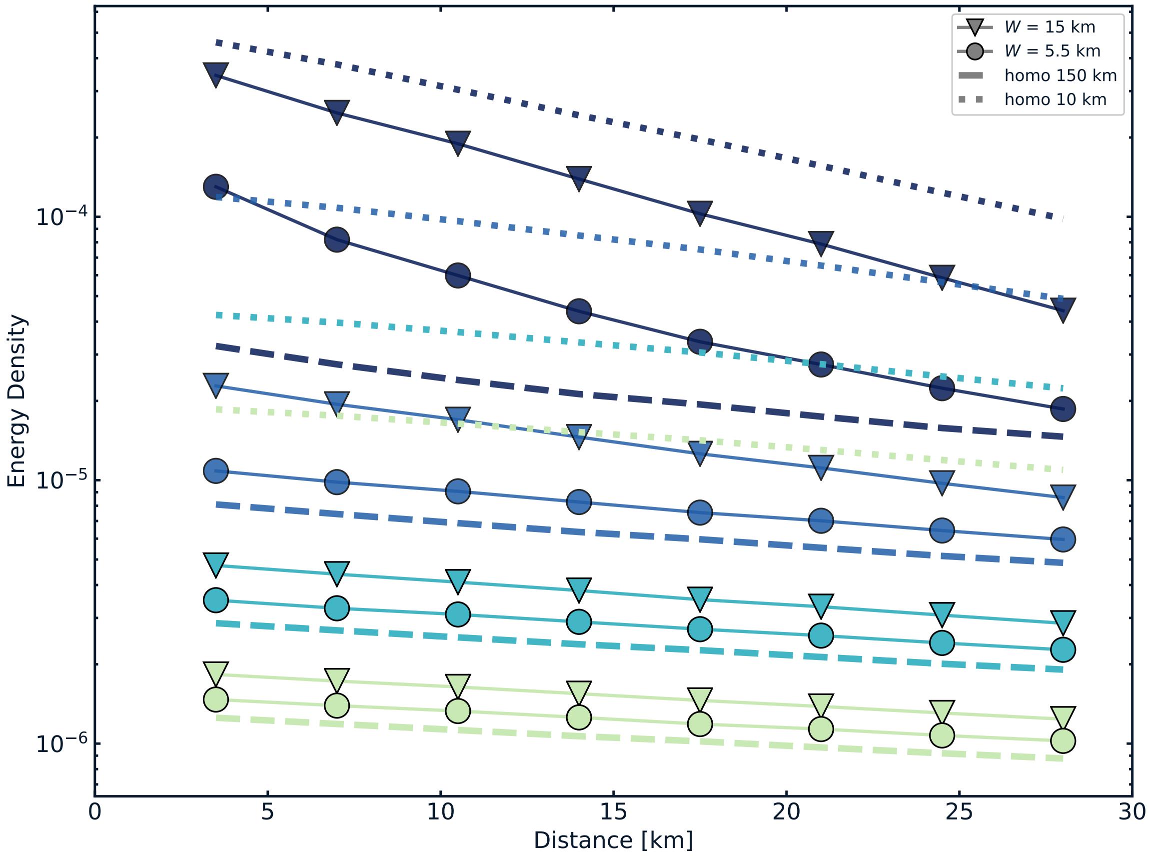

Figure S16: Comparison of spatio-temporal energy evolution of different simulations. The dashed and dotted lines show the non-normalised energy densities for the homogeneous model with ℓ= 150 km and ℓ= 10 km respectively. The non-normalised energy densities for the cases with a strong band of scattering with a width of 5.5 and 15 km are indicated by the lines with circles and triangles respectively. The colours indicate the different 15s-long time-windows, starting at 5 s, 25 s, 50 s and 75 s respectively.