Geophysical Journal International

Geophys. J. Int. (2022) **230,** 776–795
Advance Access publication 2022 February 18
GJI Seismology

https://doi.org/10.1093/gji/ggac069

Derivation of a near-surface damping model for the Groningen gas field

E. Ruigrok [1,2] A. Rodriguez-Marek,[3] B. Edwards,[4] P.P. Kruiver,[1] B. Dost [1] and J. Bommer[5]

[1]*R&D Seismology and Acoustics, Royal Netherlands Meteorological Institute (KNMI), Utrechtseweg* 297, 3731 *GA De Bilt, the Netherlands.*
E-mail: e.n.ruigrok@uu.nl
[2]*Department of Earth Sciences, Utrecht University, Princetonlaan* 8*a,* 3584 *CB Utrecht, the Netherlands*
[3]*Charles E Via, Jr., Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA* 24061*, USA*
[4]*Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Liverpool L*693*GP, UK*
[5]*Department of Civil & Environmental Engineering, Imperial College London, London SW*72*AZ, UK*

Accepted 2022 February 16. Received 2022 January 16; in original form 2021 August 17

SUMMARY

Seismic damping of near-surface deposits is an important input to site-response analysis for seismic hazard assessment. In Groningen, the Netherlands, gas production from a reservoir at 3 km depth causes seismicity. Above the gas field, an 800 m thick layer of unconsolidated sediments exist, which consists of a mixture of sand, gravel, clay and peat strata. Shear waves induced at 3 km depth experience most of their anelastic attenuation in these loose sediments. A good estimate of damping is therefore crucial for modelling realistic ground-motion levels. In Groningen, we take advantage of a large network of 200 m deep vertical arrays to estimate damping from recordings of the induced events. As a first step, we apply seismic interferometry by deconvolution to estimate local transfer functions over these vertical arrays. Subsequently, two different methods are employed. The first is the 'upgoing' method, where the amplitude decay of the retrieved upgoing wave is used. The second is the 'up-down' method, where the amplitude difference between retrieved up- and downgoing waves is utilized. For the upgoing method, the amplitude of the upgoing direct wave is affected by both elastic and anelastic effects. In order to estimate the anelastic attenuation, it is necessary to remove the elastic amplification first. Despite the fact that elastic compensation could be determined quite accurately, non-physical damping values were estimated for a number of boreholes. Likely, the underlying cause was small differences in effective response functions of geophones at different depths. It was found that the up-down method is more robust. With this method, elastic propagation corrections are not needed. In addition, small differences in *in situ* geophone response are irrelevant because the up- and downgoing waves retrieved at the same geophone are used. For the 1-D case, we showed that for estimating the local transfer function, the complex reverberations need to be included in the interferometric process. Only when this is done, the transfer function does not contain elastic transmission loss and Q estimation can be made without knowing the soil profile in detail. Uncertainty in the estimated damping was found from the signal-to-noise ratio of the estimated transfer function. The Q profiles estimated with the up-down method were used to derive a damping model for the top 200 m of the entire Groningen field. A scaling relation was derived by comparing estimated Q profiles with low-strain damping profiles that were constructed using published models for low-strain damping linked to soil properties. This scaling relation, together with the soil-property-based damping model, allowed up-scaling of the model to each grid-cell in the Groningen field. For depths below 200 m, damping was derived from the attenuation of the microseism over Groningen. The mean damping model, over a frequency band between 2 and 20 Hz, was estimated to be 2.0 per cent (0–50 m depth), 1.3 per cent (50–100 m), 0.66 per cent (100–150 m), 0.57 per cent (150–200 m) and 0.5 per cent (200–580 m).

Key words: Downhole methods; Induced seismicity; Seismic attenuation; Seismic interferometry; Site effects; Wave scattering and diffraction.

1 INTRODUCTION

Seismic monitoring is typically undertaken with an array of seismic sensors installed at the Earth's surface. In areas with high levels of seismic noise, however, the monitoring is more effective with sensors installed at depth. This is for example the case for Japan's Hi-net, where sensors are installed at depths of 100 m or more (Aoi *et al.* 2020). This location shields the sensors from large amounts of surface waves induced by anthropogenic activities. When making a hole to place a sensor at depth, in many cases additional sensors are added with little additional effort. Sensors installed at multiple depth levels form local vertical arrays (VAs). These arrays provide the opportunity to obtain near-surface seismic properties from the recordings (e.g. Trampert *et al.* 1993; Parolai *et al.* 2010; Nakata & Snieder 2012; Matsushima *et al.* 2016; Hofman *et al.* 2017).

One crucial near-surface parameter is damping. With seismic sources at depth, the amplitude decay of the wavefield is small to modest when waves propagate through the consolidated part of the Earth. The wavefield decay is typically much larger in the near surface with unconsolidated or weathered rock conditions. The damping can be described as a dimensionless Quality factor, Q, that describes the ratio between the energy stored and energy loss per cycle. Interchangeably, we use the low-strain damping, $d = 1/2Q$. In the near surface, large damping values together with low seismic velocities (thus many cycles) result in strong anelastic attenuation. An accurate estimation of this damping is critical in hazard models (e.g. Rodriguez-Marek *et al.* 2017; He *et al.* 2021). A poor estimate of the actual damping could result in an over- or underestimation of ground-motion levels, in particular at high frequencies where damping effects are dominant. The aim of this paper is twofold: (1) quest for a damping-estimation method that is robust in a high-scattering environment and (2) application of this method to derive a field-wide damping model for the Groningen gas field.

Damping is one of the most difficult seismic parameters to obtain. It can be estimated from the amplitude decay between two different depth levels. Yet, there are many other factors that control the amplitude development besides damping. Amplification occurs due to impedance reduction. Also, the wavefield is focused and defocused on near-surface features at various scales, such as former tidal channels or depositional basins. Moreover, the near-surface typically contains large impedance contrasts leading to conversions and reverberations, which might interfere with the waves that are used for estimating damping. Furthermore, it is paramount that the different sensors involved are well calibrated and have a similar coupling to the surrounding strata. To find a robust means of estimating damping in a complicated near-surface setting, we consider two methods with largely complimentary pros and cons.

One approach we apply builds on the work by Snieder & Safak (2006). They used earthquake recordings at different floors of a building. These recordings can be highly complicated due to a combination of source, path, site and building effects. By applying seismic interferometry to the ground-motion recordings, they isolated the building response. From the latter response, they extracted intrinsic attenuation. Their method can largely be imported to a near-surface setting. The main difference is that the building is made from the same material from top to bottom, whereas the near surface typically has a significant depth variation of seismic properties. The scheme that we apply deconvolves recordings at depth with recordings at the largest depth. This results in retrieving a strong upgoing direct wave. From the amplitude decay of this wave, the damping is estimated. We coin this the upgoing method.

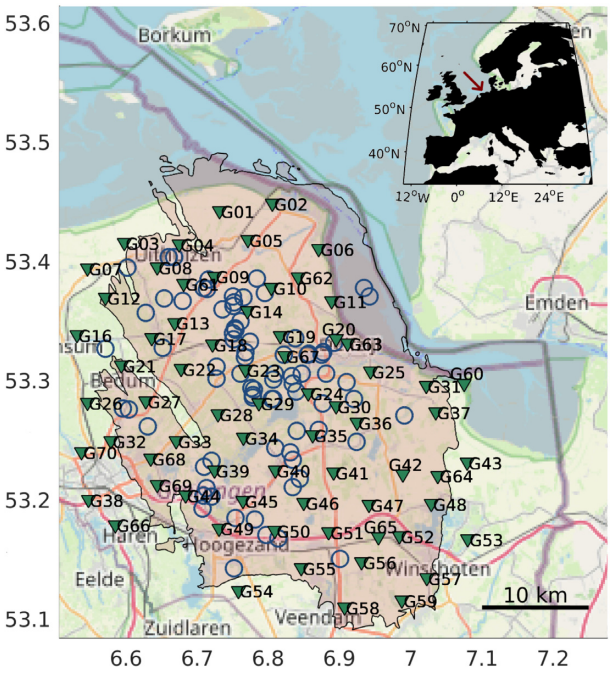

Figure 1. Seismic network of vertical arrays (green triangles) in the Groningen area (the G-network, Dost *et al.* 2017) installed to monitor the Groningen gas field (red-shaded area), located in the northeast of the Netherlands, in Western Europe (red arrow in inset). Blue circles are locations of events used in this study (between January 2015 and December 2019, with local magnitude of 1.5 and higher). The background map is from openstreetmap.org. Coordinates are in degrees latitude and longitude (WGS 84).

The other approach builds on work by Trampert *et al.* (1993), Fukushima *et al.* (2016) and Haendel *et al.* (2019). Earthquake records at depth are deconvolved by records at the Earth's surface. This yields a response in which the direct up- and downgoing waves stand out. Subsequently, in Fukushima *et al.* (2016) the spectral ratio of this up- and downgoing wave is used to estimate damping. We apply a small modification to this technique and coin it the up-down method.

The damping estimation is applied to the Groningen setting in the Netherlands. For an update of the hazard and risk model in this area (van Elk *et al.* 2019), the ground-motion model developed by Bommer *et al.* (2017) required refinement of the existing site-response model (Rodriguez-Marek *et al.* 2017). At the same time, the Groningen setting is ideal for testing the robustness of damping-estimation methods due to the large quantity of measurements available. The compaction of the Groningen gas field leads to induced seismicity, which has been recorded since 1991. A small network of VAs had been established in 1995. By the end of 2014 and beginning of 2015, a large seismic network was installed covering the entire gas field with 69 VAs (Fig. 1). At each VA there is an accelerometer at the Earth's surface and there are 4.5 Hz geophones at 50, 100, 150 and 200 m depth. A large number of induced events have been

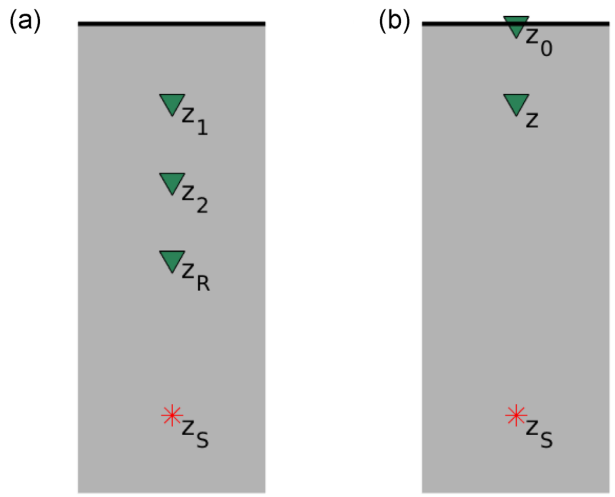

Figure 2. Configurations for the application of borehole seismic interferometry. (a) Using actual sources at depth z_S, responses are retrieved between the lowermost receiver z_R and receivers at other depth levels: z_1 and z_2. Damping is estimated between z_1 and z_2. (b) Using actual sources at depth z_S, responses are retrieved between a receiver at the Earth's surface z_0 and receivers below at depth level z. For estimating damping, the retrieved response at z is used.

recorded over these VAs. In a previous work (Hofman *et al.* 2017) these records were used to estimate near-surface interval velocities. Detailed *S*-wave profiles are available at 62 VA locations over the top 30 m from seismic cone penetration tests. The unconsolidated sediments reach a depth of approximately 800 m. A model for the entire field has been published in Kruiver *et al.* (2017). In addition, damping values have been measured in the lab for soil samples obtained from Groningen (Zwanenburg *et al.* 2020).

In the following we first describe the two different methods that we test to obtain damping values over the top 200 m in Groningen. We discuss estimating transfer functions (TFs) with seismic interferometry and we make a choice for the specific approach of deriving Q from these TFs. Next, the damping profiles are estimated for the entire G-network. Also, an estimate of damping below 200 m depth is obtained from surface waves. Finally, the damping profiles at individual stations are used to build a damping model at each grid point above the Groningen field.

2 METHOD

We will use seismic interferometry by deconvolution to retrieve empirical TFs between different borehole sensors. The direct waves within these functions are used for estimating damping values. We discuss two different methods for estimating Q. One method uses upgoing waves only, the other method uses amplitude differences of up- and downgoing waves.

2.1 Upgoing method

The rationale of the upgoing method is presented using a 1-D analytic example. The 1-D configuration is sketched in Fig. 2(a). There is a vertical array of sensors, which is illuminated from below from a source at depth level z_S. In the following analysis we assume that medium properties vary only with depth z. Q is a dimensionless factor expressing anelastic attenuation per cycle. For simplicity, it is assumed there is a (piecewise) homogeneous Q field over the VA set-up. In case of a homogeneous velocity field or a velocity gradient, there is only an upgoing wave and a free-surface reflected downgoing wave that is recorded at sensor depth z_R. The upgoing wave can be written as

$$U^-(z_R, z_S) = S(z_R, z_S)\mathrm{e}^{-j\omega t}\mathrm{e}^{-\omega t/(2Q)}, \tag{1}$$

where j is the imaginary unit, ω is the angular frequency and t is the traveltime from source position z_S to (reference) receiver position z_R. For U, the superscript $^-$ denotes that just the upgoing wavefield is taken. $S(z_R, z_S)$ is the resulting amplitude due to source and propagation effects. U and S are in capitals to denote that they are quantities expressed in the frequency domain.

The response at z, which is any receiver position above position z_R is written as

$$U^-(z, z_S) = U^-(z_R, z_S)F(z, z_R)\mathrm{e}^{-j\omega\tau}\mathrm{e}^{-\omega\tau/(2Q)}, \tag{2}$$

where τ is the traveltime between the two receiver levels. $F(z, z_R)$ describes amplitude modulation by elastic propagation effects between the receiver levels (e.g. impedance reduction) and $e^{-\omega\tau/(2Q)}$ describes the anelastic attenuation.

By deconvolving the upgoing response at z by the upgoing response at z_R we obtain the upgoing direct wave between these two depth levels:

$$D^-(z, z_R) = F(z, z_R)\mathrm{e}^{-j\omega\tau}\mathrm{e}^{-\omega\tau/(2Q)}. \tag{3}$$

That is, the deconvolution result D^- is equal to upgoing impulse response between the two depth levels.

In practice, the wavefield recorded at depth level z_R contains both upgoing and downgoing waves: $U = U^- + U^+$. The downgoing (free-surface reflected) wavefield at z_R can be written as

$$U^+(z_R, z_S) = U^-(z_R, z_S)F(z_R, z_R)\mathrm{e}^{-j\omega\tau_R}\mathrm{e}^{-\omega\tau_R/(2Q)}, \tag{4}$$

where τ_R is the traveltime from z_R to the free surface and back again and $F(z_R, z_R)$ is the elastic amplitude modulation over this same trajectory.

By deconvolving the upgoing wave at z (eq. 2) by the complete response at z_R (eqs 1 and 4) the following expression is obtained

$$D_T^-(z, z_R) = \frac{F(z, z_R)\mathrm{e}^{-j\omega\tau}\mathrm{e}^{-\omega\tau/(2Q)}}{1 + F(z_R, z_R)\mathrm{e}^{-j\omega\tau_R}\mathrm{e}^{-\omega\tau_R/(2Q)}}. \tag{5}$$

D_T^- denotes the (upgoing) TF. Eq. (5) can be written as a series expansion in which eq. (3) is the first term (Snieder & Safak 2006). The other terms are later arriving reverberations between the actual free surface and an imposed fully reflecting surface at z_R with a reflection coefficient of -1. In the following we use only the retrieved upgoing wave at earlier times (eq. 3).

Eq. (3) cannot directly be used to estimate Q. In practice, the recordings contain noise. The absolute amplitude of the deconvolution is affected by the noise in the denominator and the implemented stabilization term. By multiplication with a noise function and by labelling the upper receiver as z_1 we obtain

$$D^-(z_1, z_R) = N_1(\omega)F(z_1, z_R)\mathrm{e}^{-j\omega\tau_1}\mathrm{e}^{-\omega\tau_1/(2Q)}. \tag{6}$$

The deconvolution can be repeated between the reference receiver at z_R and another receiver at z_2 which is below z_1 (Fig. 2a):

$$D^-(z_2, z_R) = N_2(\omega)F(z_2, z_R)\mathrm{e}^{-j\omega\tau_2}\mathrm{e}^{-\omega\tau_2/(2Q)}. \tag{7}$$

Eqs (6) and (7) are both obtained with the same denominator and the same stabilization. Hence, it may be assumed that the noise functions are similar. With the assumption that they are identical, by deconvolution of eqs (6) and (7), the direct wave between both

depth levels is obtained:

$$D^{-}(z_1, z_2) = F(z_1, z_2)\mathrm{e}^{-j\omega\tau_{12}}\mathrm{e}^{-\omega\tau_{12}/(2Q)}, \tag{8}$$

where τ_{12} is the interval traveltime between depth levels 1 and 2.

From eq. (8), interval Q and timing can be estimated when the profile is known well enough to first correct for the elastic amplitude term $F(z_1, z_2)$ (Appendices A and B). Note, however, that also this second deconvolution may suffer from noise and (small) amplitude modulation due to stabilization. Therefore, in practice, maximum amplitudes are picked from the retrieved direct waves (eqs 7 and 8) and, from the ratio of these amplitudes, Q is estimated (Section 2.3).

2.2 Up-down method

In this section, we consider a configuration with a receiver at the Earth's surface and one or more receivers at depth. This is akin the configuration chosen in Fukushima *et al.* (2016) and Haendel *et al.* (2019).

As before, we consider a homogeneous field or a vertical gradient in elastic parameters, a (piecewise) homogeneous Q field over the VA setup and illumination from below. In this configuration we choose the reference level z_0 at the free surface (Fig. 2b). The response at this level can be written as

$$U(z_0, z_S) = S(z, z_S)F(z_0, z)\mathrm{e}^{-j\omega t}\mathrm{e}^{-\omega t/(2Q)}, \tag{9}$$

where t denotes the traveltime from the source to the receiver and $S(z, z_S)$ contains source and propagation amplitude terms up till depth level z, which is some depth below the free surface. The remaining elastic amplitude modulation, including the free-surface effect, is in $F(z_0, z)$.

The upgoing wave at depth level z can be written as

$$U^{-}(z, z_S) = S(z, z_S)\mathrm{e}^{-j\omega t}\mathrm{e}^{j\omega\tau}\mathrm{e}^{-\omega t/(2Q)}\mathrm{e}^{\omega\tau/(2Q)}, \tag{10}$$

where τ denotes the traveltime from the free surface to z. Similarly, the free-surface reflected downgoing wave at depth level z can be written as

$$\begin{aligned} U^{+}(z, z_S) = {} & S(z, z_S)F(z_0, z) \\ & \times F(z, z_0)\mathrm{e}^{-j\omega t}\mathrm{e}^{-j\omega\tau}\mathrm{e}^{-\omega t/(2Q)}\mathrm{e}^{-\omega\tau/(2Q)}, \end{aligned} \tag{11}$$

where $F(z, z_0)$ is the elastic amplitude modulation for the downgoing trajectory.

By deconvolving the upgoing wave at z by the response at z_0, we obtain

$$D^{*-}(z, z_0) = 1/F(z_0, z)\mathrm{e}^{j\omega\tau}\mathrm{e}^{\omega\tau/(2Q)}. \tag{12}$$

which is a time-reversed version of an upgoing wave, which is denoted with the complex conjugate sign * in the frequency domain. Also the attenuation is flipped. As a consequence, the retrieved direct wave is amplified from z_0 to z.

Similarly, by deconvolving the downgoing wave at z by the response at z_0, we obtain

$$D^{+}(z, z_0) = F(z, z_0)\mathrm{e}^{-j\omega\tau}\mathrm{e}^{-\omega\tau/(2Q)}. \tag{13}$$

This deconvolution result corresponds to a physical direct wave between z_0 and z with positive propagation time and attenuation due to loss factor Q.

Finally, by deconvolving the downgoing (eq. 13) with the time-reversed upgoing wave (eq. 12) we obtain

$$D(z, z) = F(z, z_0)F(z_0, z)\mathrm{e}^{-j\omega 2\tau}\mathrm{e}^{-\omega\tau/Q}, \tag{14}$$

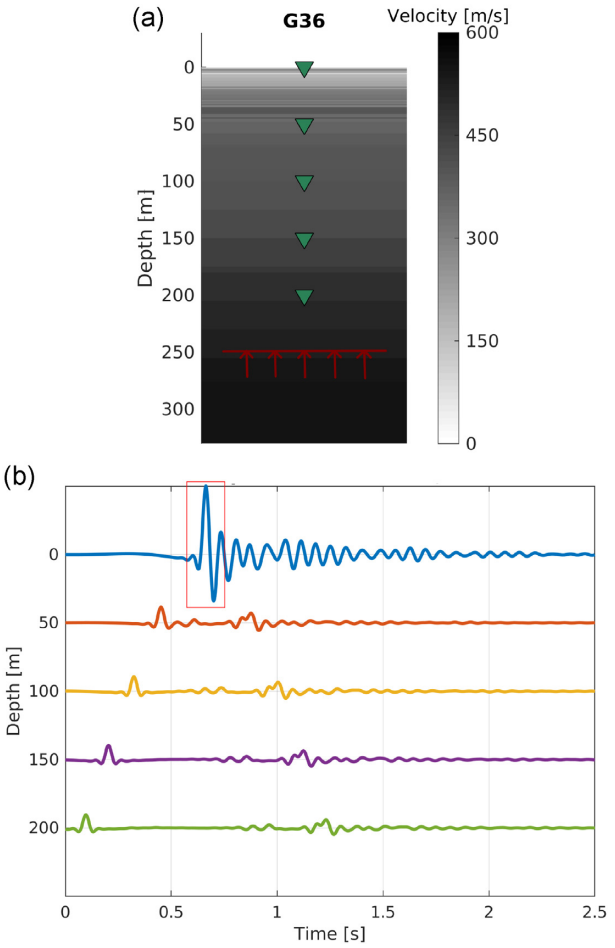

Figure 3. (a) *S*-wave model at borehole station G36, with five receiver levels (green triangles) and a plane-wave source at 250 m depth. (b) Elastic *S*-wave response at the five different receiver levels. The forward modelling is performed with a reflectivity code. A band-limited delta pulse ([2 20] Hz) is used as a source-time function. The red box shows the time window that is used to approximate the direct wave at the Earth's surface.

$D(z, z)$ has the timing, elastic amplitude modulation and anelastic attenuation for a direct wave propagating from z to the free surface and back again.

In case of a velocity gradient $1/F(z_0, z) = F(z, z_0)$ and the elastic propagation effects are removed from eq. (14). For a homogeneous medium, $F(z_0, z)$ only describes the free-surface effect, which is counteracted by $F(z, z_0)$. In both a homogeneous medium and a velocity-gradient medium, no scattering takes place. As soon as there is scattering, the direct wave undergoes transmission losses on its trajectory between z to the free surface and back again and thus $F(z_0, z)F(z, z_0) < 1$. Thus, before Q can be estimated from D (eq. 14), first a correction for transmission loss would need to be implemented. Note, however, that the need for this correction follows from only using direct waves in the above derivation. In the numeric example below it is shown that when the complete response is used in the deconvolution, also in the 1-D scattering case, no corrections are needed.

Fig. 3 shows the response over a borehole for a plane-wave source at 250 m depth. The forward modelling is performed without anelastic attenuation. The model is a velocity gradient between 200 and 50 m depth (Fig. 3a), resulting in an isolated upward going pulse at the lower depth levels. In the upper 50 m there are large impedance

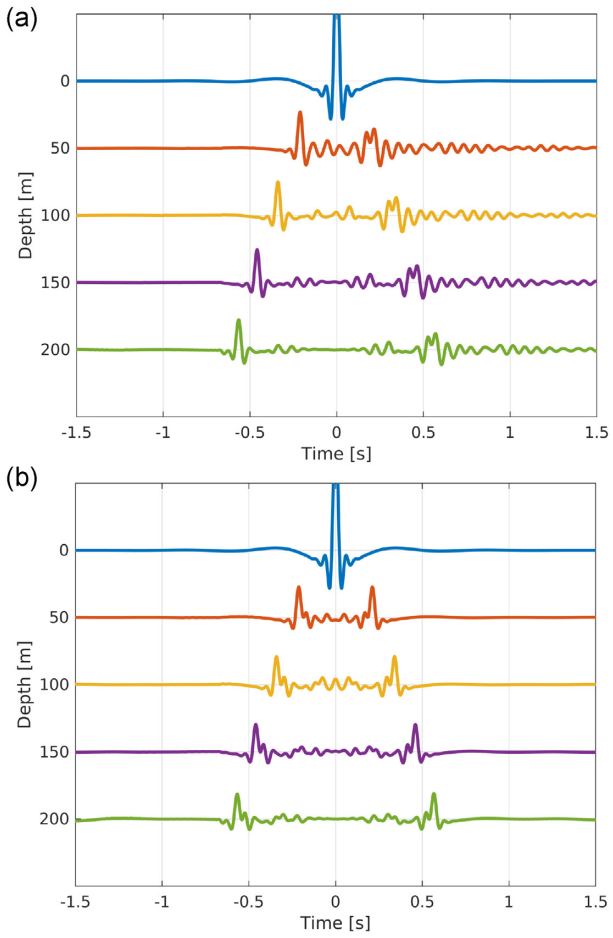

Figure 4. Responses retrieved with deconvolution (a) when an approximation of the direct wave recorded at the Earth's surface is used (as indicated with the red box in Fig. 3b) and (b) when the total response at the Earth's surface is used. In both cases, the complete forward modelled response, as shown in Fig. 3(b), is taken for the 50–200 m depth levels.

contrasts, resulting in a long scattering sequence recorded at the Earth's surface and a complicated downgoing wave sequence at the lower depth levels. When the direct wave is isolated from the z_0 response, as in eqs (12) and (13), the deconvolution results are shown in Fig. 4(a). It can be seen that the retrieved downgoing wave indeed has lower amplitudes than the upgoing wave at the same recording level (due to transmission loss).

From seismic-interferometry theory (e.g. Wapenaar *et al.* 2010b) it follows that in the elastic case, a response R is retrieved that is symmetric or antisymmetic in time. That is $[R(\,-t)]^t = R(t)$ or $[R(\,-t)]^t = -R(t)$, where ^{t} denotes time reversal. If instead of only the direct wave, the complete responses (Fig. 3b) are taken and deconvolved, a time-symmetric result is obtained (Fig. 4b). The direct-wave elastic amplitude loss that was observed in Fig. 4(a) is compensated by scattering contributions that also travel between z_0 and z. Thus, when there is scattering, all this scattering needs to be included in the deconvolution process. When this is done, deconvolution of the upgoing and downgoing retrieved waves yields

$$H(z,z) = \mathrm{e}^{-j\omega 2\tau}\mathrm{e}^{-\omega\tau/Q}, \tag{15}$$

in which transmission loss is not present and only the anelastic attenuation remains. Though H does not describe a physically correct wave, it is a convenient expression to use as a base for estimating Q.

From the above analysis one can learn that by using both the up- and downgoing waves as in Fukushima *et al.* (2016) one can estimate Q without the need to know elastic parameters. However, to remove the imprint of transmission losses, long time windows are to be used in the deconvolution process, which windows include the complex near-surface reverberations. Using short time windows will lead to inclusion of transmission loss in H and therewith to overestimation of damping.

In the ideal case without noise, the deconvolution result is zero prior to the retrieval of the upgoing wave (Fig. 4). This characteristic is used in Appendix C for estimating the signal-to-noise ratio (SNR) of the interferometric process and for propagating this SNR to a damping uncertainty.

2.3 Q-estimation methods

There is a suite of Q estimation techniques that takes advantage of the precise amplitude versus frequency development of the direct wave over multiple depth levels. Notably, the spectral-ratio method and the linear-approximation method based on the analytic trace (Tonn 1991). The large appeal of these methods is that, in ideal cases, transmission effects and geometric spreading are left out of the equation. Damping is estimated from the spectral slope, which slope is not affected by processes that are frequency independent. The main disadvantage of these methods is that they are susceptible to small distortions in the spectra, e.g. caused by interference of the direct wave with subtle scattering (e.g. Matsushima *et al.* 2016). On the other side of the spectrum there are methods that only use the decay of the maximum amplitude of the direct wave. These methods are more robust, but may need a correction for transmission loss and geometric spreading.

For the Groningen application we may assume that the retrieved responses have negligible geometric spreading over the near-surface VAs. For the upgoing method we make amplification corrections (Appendix A). For the up-down method, there is no elastic amplitude loss or gain that needs correction (eq. 15).

A straightforward approach would to extract Q from a ratio of amplitudes. Taking the ratio of the amplitude of the downgoing and upgoing wave equates to taking the absolute value of eq. (15). Isolating Q from this, we find

$$Q(f) = \frac{-2\pi\tau f}{\ln(|H(f)|)}, \tag{16}$$

where ln is the natural logarithm and τ is the one-way traveltime between the receiver level at depth and the surface. This equation is used in Fukushima *et al.* (2016) to find a frequency-dependent Q factor. Note that in Fukushima *et al.* (2016) the factor of 2 is not present in the numerator. This is because they use a τ' denoting a two-way traveltime.

Alternatively, Q can be estimated from time-domain amplitudes. A robust implementation is to use a ratio of maximum amplitudes:

$$Q = \frac{-2\pi\tau f^d}{\ln\frac{A^+}{A^-}}, \tag{17}$$

where A^+ and A^- are the maximum amplitudes of the (direct) down- and upgoing waves, respectively and f^d is the dominant frequency.

Tonn (1991) compared Q estimation methods and he found the maximum method to be most robust. With the maximum method, the analytic signal (e.g. Schimmel & Paulssen 1997) is taken from

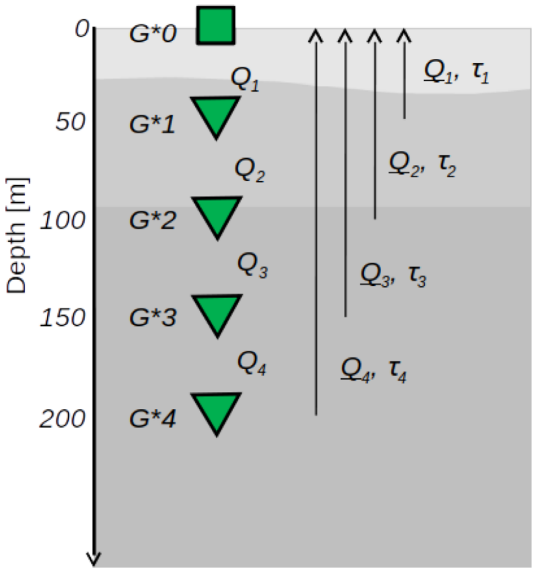

Figure 5. Q estimation setup for a G-network station. G∗ is short for any station (G01–G70; however G15 does not exist). At each station there is a vertical array of receivers with an accelerometer (green square) at the surface and geophones (green triangles) at depth. The receiver levels are numbered from 0 at the Earth's surface to 4 at 200 m depth. τ_1 to τ_4 are the one-way *S*-wave traveltimes from receivers 1 to 4, to the surface. Using seismic interferometry, $\underline{Q}_1$ to $\underline{Q}_4$ can be estimated. By subsequently harmonic de-averaging (eq. 21) interval Q values are found: Q_1 to Q_4.

the retrieved seismogram. Hence, instantaneous amplitudes (envelopes) and instantaneous frequencies are used. In particular, the instantaneous frequency F is more stable than the dominant frequency f^d in the presence of noise. Using the maximum method and neglecting geometric spreading, we can estimate Q between the free surface and depth level z with

$$Q = \frac{-\pi\tau(F^- + F^+)}{\ln\frac{E^+}{E^-}}, \tag{18}$$

where $E^{\pm}$ and $F^{\pm}$ are the envelope and the instantaneous frequency of the direct upgoing and downgoing waves. The upgoing method equivalence of eq. (18) can be derived from eq. (8), yielding (after elastic amplitude correction)

$$Q = \frac{-\pi\tau_{12}(F^1 + F^2)}{2\ln\frac{E^1}{E^2}}, \tag{19}$$

where $E^{1,2}$ and $F^{1,2}$ are the envelope and the instantaneous frequency of the upgoing wave at two different receiver levels and τ_{12} is the traveltime between these levels.

With the above approach, an average loss factor $\underline{Q}_i$ is estimated, which is the average attenuation over the depth interval from the free surface to depth level z_i. For a Groningen setup with 4 receivers at depth (Fig. 5) $\underline{Q}_i$ can be estimated for 4 partly overlapping depth intervals. $\underline{Q}_i$ can be subdivided into the loss factor from the free surface to depth level z_{i-1} and an additional loss factor from level z_{i-1} to z_i :

$$e^{-\omega\tau_i/2\underline{Q}_i} = e^{-\omega\tau_{i-1}/2\underline{Q}_{i-1}} e^{-\omega(\tau_i-\tau_{i-1})/2Q_i} \tag{20}$$

From the above expression the interval Q_i can be isolated:

$$Q_i = \frac{\tau_i - \tau_{i-1}}{\tau_i/\underline{Q}_i - \tau_{i-1}/\underline{Q}_{i-1}}. \tag{21}$$

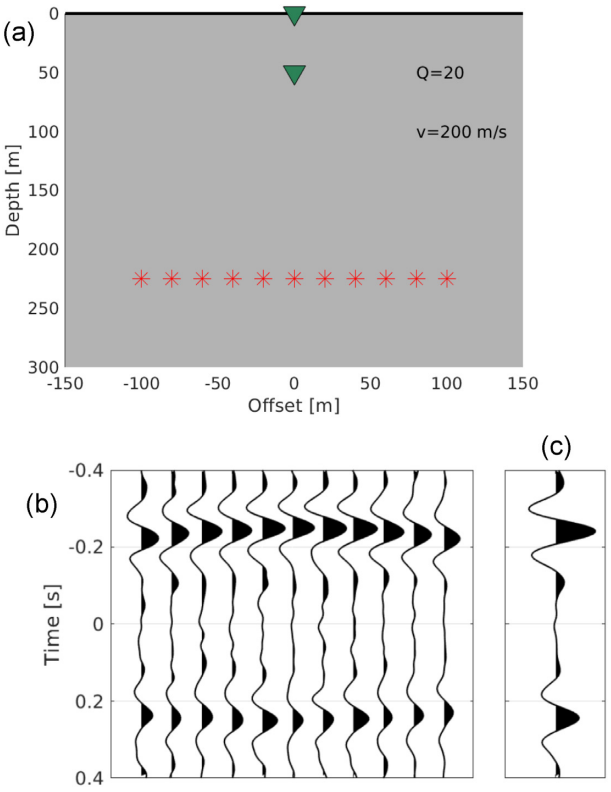

Figure 6. Numerical implementation of up-down method. (a) Configuration with an homogeneous subsurface, an array of sources (red stars) and two receivers (green triangles). (b) Deconvolution gather: the result of deconvolving the responses at the lower receiver with the ones at the upper receiver. (c) Result of stacking the deconvolutions over the sources.

By using the above de-harmonic averaging equation to a five-sensor VA, four interval Q values are estimated. For example Q_2 is the factor for the 50–100 m depth interval (Fig. 5).

2.4 Integration

In Section 2, we assumed that there is one source, illuminating the VA from below with an angle of incidence equal to zero. In reality, the seismicity provides a range of illumination. In Groningen, angles of illumination are observed between about -30° and +30° at the 200 m depth level (Hofman *et al.* 2017). Each source provides a different illumination and therewith a different deconvolution result. Within the context of seismic interferometry, deconvolutions need to be stacked (integrated) before subsurface parameters can be estimated. In the current implementations of the up-down method, however (Fukushima *et al.* 2016; Haendel *et al.* 2019) Q is estimated from each deconvolution and afterwards Q is averaged over all estimates. Below we numerically test both approaches.

Seismic interferometry (e.g. Wapenaar *et al.* 2010a) is the application of an integral equation with correlations of waveforms, or deconvolutions of waveforms, in the integrand. When this integral equation is used to obtain a virtual source at a receiver location, the integration is over sources, or plane-wave illumination angles (Ruigrok *et al.* 2010). For a configuration as in Fig. 6(a) the subsurface sources can be remapped to retrieve the response as if there were a source at the upper receiver which response were measured

by the lower receiver:

$$T(\mathbf{x}, \mathbf{x}_0) = \int_{\mathbf{x}_S} \frac{U(\mathbf{x}, \mathbf{x}_S)}{U(\mathbf{x}_0, \mathbf{x}_S)}, \tag{22}$$

where $\mathbf{x}_S$ are source positions at depth, $\mathbf{x}$ is a receiver location at depth and $\mathbf{x_0}$ is a receiver at the free surface. The configuration and deconvolution in the above equation is the same as in the up-down method. However, the integration is added to isolate from all illumination angles only the vertical propagation between both receivers as can be shown with a stationary-phase approximation of this integral (Snieder 2004).

In practice, the above integration is approximated by a summation over different sources. Additionally, the deconvolution needs to be stabilized:

$$T(\mathbf{x}, \mathbf{x}_0) \approx \sum_{\mathbf{x}_S} \frac{U(\mathbf{x}, \mathbf{x}_S) U^*(\mathbf{x}_0, \mathbf{x}_S)}{|U(\mathbf{x}_0, \mathbf{x}_S)|^2 + \epsilon}, \tag{23}$$

where * denotes the complex conjugate, $|U|^2$ is the power spectrum of U and ϵ is a stabilization constant.

Fig. 6 shows a numerical example. A homogeneous model is used with a velocity of 200 m s^{-1} and $Q = 20$. A distribution of sources is placed to mimic illumination from -30° to +30°. There are two receivers, one at 50 m depth, the other at the free surface. The responses are forward modelled using a Ricker wavelet with a dominant frequency of 8 Hz. Gaussian white noise is added with amplitudes of up to 20 per cent of the direct wave. Fig. 6(b) shows the result of applying the deconvolution for each individual source and Fig. 6(c) is the result after summing the deconvolutions (eq. 23). Subsequently, Q is estimated with the maximum method, from the amplitude difference of the wave found at negative and positive time lags.

Averaging over the parameters estimated for the individual deconvolutions yields $Q = 21.3 \pm 5.6$. When using the stacked deconvolutions, $Q = 21.8$ is obtained. Both estimates are somewhat higher than the actual Q, which is 20. With the used illumination and level of noise, both approaches give a similar result. With higher noise levels, however, a part of individual deconvolutions have an insufficient SNR for Q estimation, whereas estimating Q from the stack remains possible due to noise suppression in the stacking process. For the robustness of implementation we therefore opt for the interferometric approach in the following. The uncertainty in the Q estimation is obtained from the noise that remains in the interferometric result (Appendix C).

3 RESULTS

In this section, results are shown of applying Q-estimation to G-network stations in Groningen, the Netherlands. First local TFs are estimated with seismic interferometry and, subsequently, damping factors are derived with either the upgoing method (Section 3.1) or the up-down method (Section 3.2). In Section 3.3, the microseism is used to find damping values in the unconsolidated sediments below 200 m depth.

For both interferometric methods, the first steps are similar. Eq. (23) is applied for receiver combinations at the VAs. Fig. 7 shows the results for one receiver combination at station G36. From the earthquake catalogue, all events are selected within 20 km epicentral distance and with a minimum local magnitude of 1.5 (Fig. 1). Most events are induced at reservoir level, which is at about 3 km depth (e.g. Spetzler & Dost 2017). The events are bandpass filtered between 2 and 20 Hz and are shown as function of source-receiver

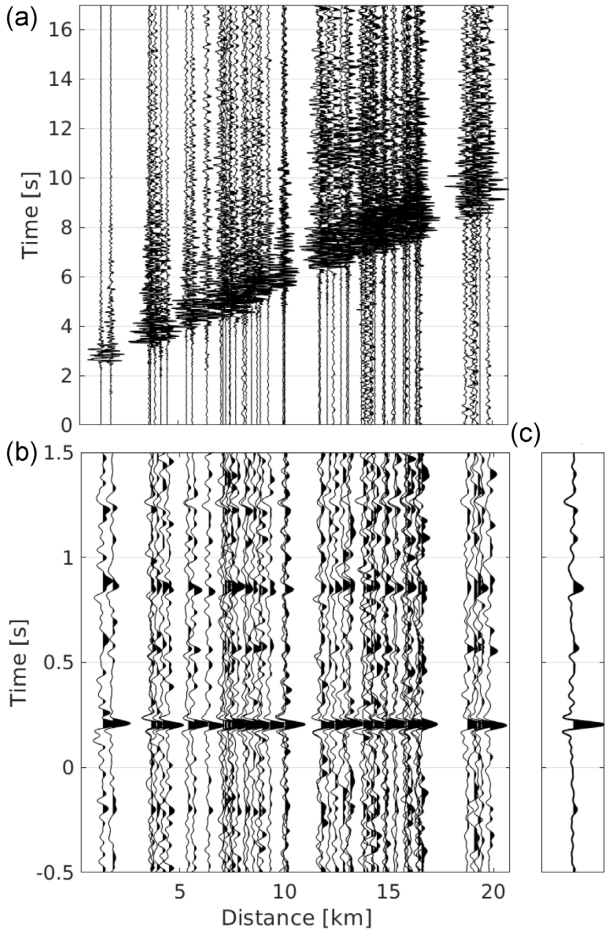

Figure 7. Interferometric processing at G36. (a) The transverse-component receiver gather at one depth level, showing the recording of 55 events as function of distance, bandpass filtered between 2 and 20 Hz. (b) Deconvolution gather: result of deconvolving the recordings at the 200 m depth level from the recordings at 100 m depth. As a stabilization constant, 10 per cent of the median of the denominator in eq. (23) is taken. (c) Result of stacking all the deconvolutions to obtain an estimate of a TF between 200 and 100 m depth.

distance (Fig. 7a). In this band, all events have a sufficient SNR. Below 2 Hz only the larger events still have signal that stands out from the noise. Above 20 Hz the wavefield undergoes strong scattering and a local 1-D assumption diminishes to hold. The transverse component event recordings (Fig. 7a) are a good approximation of the *S*-wave response. Nearly no amplitudes can be noted prior to the first *S*-wave arrival. The recordings, however, show a large variation from event to event. With the deconvolution step, the source and path effects are largely removed. What remains are mainly the local TFs. In Fig. 7(b), it can be seen that the estimation of these TFs is similar, irrespective of the event (distance) used. Each deconvolution result by itself is not yet a perfect estimate of the TF because of imperfect illumination, noise and the stabilization constant. By stacking over the deconvolution gather, a response is obtained that better approximates the local TF in amplitude and phase (Fig. 7c).

3.1 Upgoing method

Fig. 8 shows the result of retrieving *SH* TFs between the lowest receiver level at G36 and all other receivers. At the 200 m depth level, the TF corresponds to a band-limited delta pulse. This pulse can be

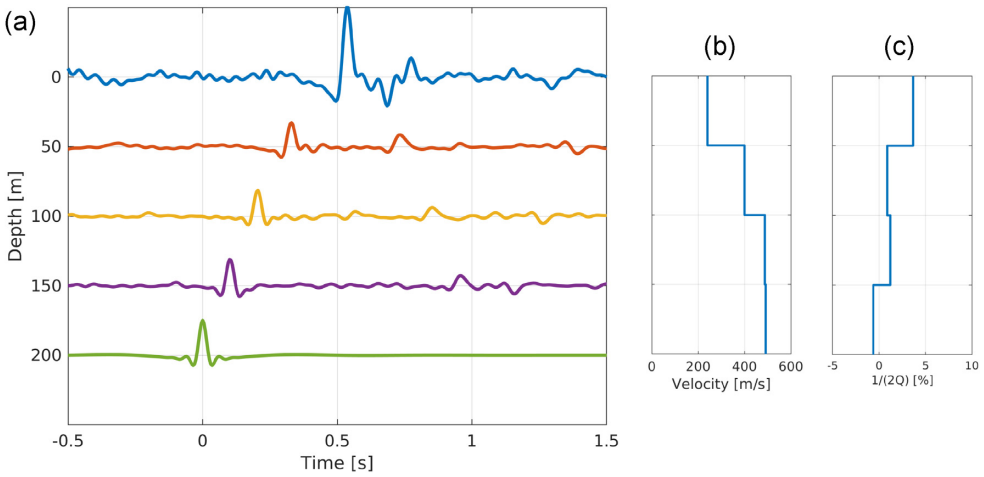

Figure 8. (a) The results of applying seismic interferometry by deconvolution at station G36 with a virtual source constructed at the 200 m depth level. The transverse-component result is shown in the band [2 20] Hz. (b) *S*-wave interval velocity model estimated from the timing of the upgoing wave in (a). (c) Damping interval model estimated from the amplitude decay of the upgoing wave in (a).

followed upwards to the other receiver levels. At the free surface, a strong amplitude gain can be noted, which is a combination of the free surface effect (factor of 2) and additional amplification in the shallow subsurface. It can be seen that part of the energy remains trapped in the near surface, leading to reverberations following the direct pulse. It can also be seen that a part of the energy does propagate downwards and shows a pulse succession that is almost symmetrical (in time) to the upgoing wave. Interestingly, the TF has a -1 reflection coefficient at the deconvolution level (200 m in this case, see eq. 5 and Snieder & Safak 2006). As a consequence, at delay times larger than 1 s, an upward travelling wave can be noted with a flipped polarity.

In Fig. 8(a), the upgoing wave has the highest SNR and is least affected with interference. From the delay times of the pulse arriving at the different depth levels, an interval velocity model is estimated as shown in Fig. 8(b). This procedure is done for all 69 VAs and yields an updated interval velocity model with respect to the one published in Hofman *et al.* (2017). Since 2017 many more events are available which further improves the estimate. Also, instead of applying cross-correlations, now deconvolutions have been implemented, which result in TFs which are more amenable to traveltime picking than the (impulse) responses that are estimated with cross-correlations (Nakata & Snieder 2012). For most sites and depth intervals, the difference between the Hofman model and the updated model is small. The newly picked delay times are also used for the Q estimation (τ_{12} in eq. (19)).

The estimated TFs are corrected for elastic propagation effects (Appendix A) including the free-surface effect. Subsequently, the attenuation is estimated from the upgoing wave using the maximum method (Section 2.3). This results in an estimate of the attenuation parameter Q for the upper three depth intervals. For obtaining an estimate for the deepest depth interva, a TF is retrieved with a virtual source constructed at the Earth's surface. Fig. 8(c) shows the damping estimated for station G36 (blue line).

At all other G stations, the Q-estimation procedure is repeated. Stations are removed where there is a known malfunction of one of the horizontal components, at one of the depth levels. Also, all stations without optimized velocity model (Appendix B) are removed. For an accurate elastic propagation correction, a detailed velocity model of the near surface is deemed vital. For the remaining 36 stations, Fig. 9 shows the estimated Q values.

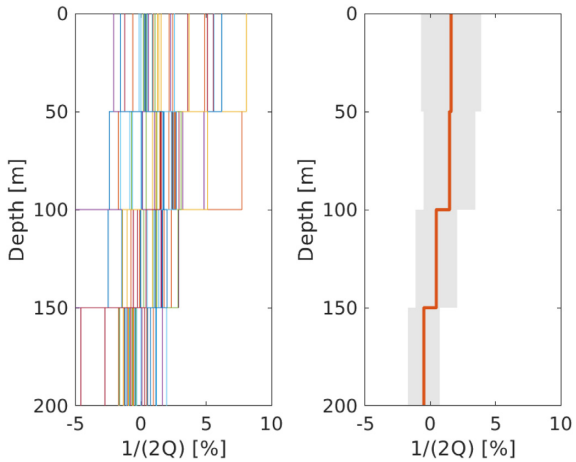

Figure 9. Damping profiles estimated with the upgoing method at 36 sites of the G-network. Left: distribution of the 36 profiles. Right: mean (orange line) and spread (-sigma to sigma) of damping values over the network. The mean Q values for the four depth intervals are—from top to bottom—31, 34, 107 and −103.

The mean estimated values look reasonable for the upper three depth intervals. However, the average level for lowest depth interval (200 to 150 m) is negative. This means that waves are amplified. At some locations there could be focusing of the wavefield that is misattributed to Q. It is physically very unlikely, however, to obtain a mean negative value over 36 sites, Also at other depth levels there are both negative and very large values in the distribution (Fig. 9) that seem unrealistic of the actual damping conditions.

For a few sites, we investigated in detail what could cause the estimation of un-physical damping values. First the orientation of horizontal components, and its uncertainty, was assessed in detail (Ruigrok *et al.* 2019). As the uncertainty was found to be less than a few degrees at most geophones, this could be ruled out as a cause of the observed amplitude differences. Secondly, it was checked whether there could be flaws in the intererometric processing. This hypothesis could be rejected by a modelling exercise and by plotting amplitude development of raw waveform data and comparing this with expected amplification based on the detailed subsurface model. Fig. 10 shows this comparison for stations G36 and G39. On this

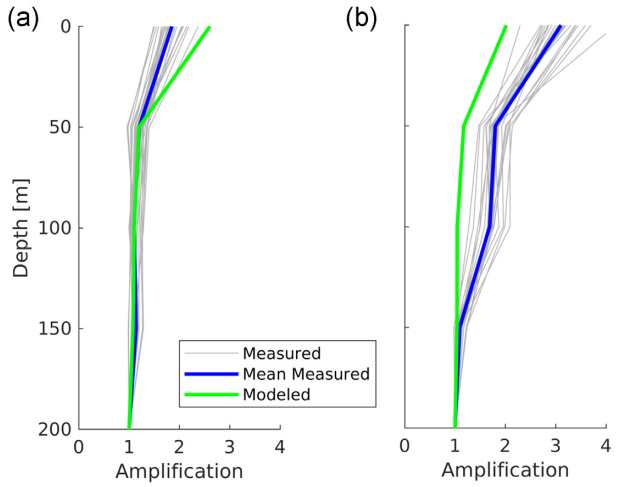

Figure 10. Comparison of measured and modelled amplification for (a) station G36 and (b) station G39. For a distribution of events, the grey lines show the measured transverse-component peak ground velocity as function of depth. The maximum amplitude that is picked would generally be the direct upgoing *S*-wave. The mean peak-ground-velocity development (blue line) is compared with the elastically modelled amplification (green line) using the optimized model. Since in the actual subsurface, damping should be positive, the green line should depict larger values than the blue one. For G36 this is not the case at 150 m depth, which results in a negative Q estimate at the lowest depth interval (Fig. 8c). At G39 there is a large unexplained amplitude jump from 150 to 100 m depth, without a large difference in impedance between these two depth levels.

raw data, amplitude jumps can be noted that cannot reasonably be explained by site effects.

The most likely remaining hypothesis is a mismatch, at some sites, between the *in situ* response of the geophones and the assumed nominal factory response. At four locations, there are nearly co-located geophones and broad-band stations. At these sites, it is confirmed that there are frequency-dependent amplitude differences between the geophones and broad-band sensors (appendix IV in Bommer *et al.* 2019).

3.2 Up-down method

With the up-down method, TFs are retrieved between a ground-level receiver and a receiver at depth (eqs 12 and 13). Subsequently, the amplitude difference between the upgoing and downgoing wave is used to estimate Q (eq. 15). The downgoing wave in general has poorer SNR than the upgoing wave (e.g. Fig. 8a). Therefore, the TF estimation is repeated with both the radial and the transverse component. With the approximation that the local structure is horizontally layered, the radial-component TF would be the same as the transverse-component TF. Both are added to further enhance the SNR. Fig. 11 shows the resulting TFs for one VA. With respect to the modelled deconvolution result (Fig. 4) the actual case shows broader pulses and up- and downgoing waves with smaller amplitudes. This is to be expected as no losses were implemented in the modelling case.

From the retrieved TFs at each VA, four Q values are estimated with the maximum method (eq. 18) and interval Q values are isolated with eq. (19). The TF retrieval and Q estimation are repeated for all G-network stations without known sensor issues on any of the horizontal components. Next, all stations are omitted where the downgoing wave has insufficient SNR. After these station rejection steps, 44 are remaining. Their damping distribution is shown in Fig. 12. The damping distribution over the 44 sites appears log-normal for the upper 50 m. For the top 50 m in Fig. 12 the values range from 0.25 to 5.09 per cent damping. Most of the values, however, are near 1.7 per cent. For this reason, we list here the median values, which are—from top to bottom—1.67, 0.77, 0.34 and 0.37 per cent, respectively.

For the hazard model, damping estimates over the top 200 m suffice ($\underline{Q}_4$ in Fig. 5). These estimates can be made with only the surface sensor and the 200 m depth sensor. From the 69 VAs, four are removed where issues are known to exist with one of the horizontal channels (G050, G454, G494 and G634). Furthermore, G35 and G33 are removed because of suspicious amplitude behaviour. G354 has a downgoing wave that is larger than the upgoing wave. G334 has a weak upgoing wave. The damping at the remaining 63 sites is shown in Fig. 13, as function of the average velocity over the top 10 meters V_{S10}. This upper 10 m is descriptive since this is the depth range in which most of the lateral heterogeneity occurs. The error bars are obtained with the error propagation described in Appendix C. In general, there is higher uncertainty for higher damping levels. Moreover, Fig. 13 shows that all sites with high velocities in the top 10 m have low damping. Low velocities, however, are not necessarily connected to high damping, as can be judged from the large spread in damping values; sites with velocities lower than 180 m/s show damping ranging from 0.5 to 2.7 per cent.

Fig. 14 shows the spatial distribution of the estimated damping values over the top 200 m. There is no obvious spatial pattern, but for a few higher values in the east and generally lower damping values in the south. In the southern part of the region, also the top part of the soil column consists of sands (van Ginkel *et al.* 2019). All the low damping values are found where the top column shows very low velocities (Fig. 13) corresponding to a large presence of clay and peat.

In engineering seismology κ_0 is often used as a site term (Appendix D). κ_0 is related to Q as (e.g. Van Houtte *et al.* 2018)

$$\kappa_0 = \int_z \frac{1}{QV} \mathrm{d}z. \tag{24}$$

Hence, κ_0 can be seen as a site-response term that describes the accrued attenuation over z, the depth interval of interest. Another site term that is often used is the average *S*-wave velocity over the top 30 m: V_{S30}. We compute κ_0 over the top 200 m, using the interferometrically derived Q and V values, and plot it against V_{S30} taken from the optimized model (Appendix B). The resulting distribution (Fig. 15) shows that there is an approximate log-linear relationship between the two site terms.

To determine a possible frequency dependence of Q we apply eq. (16). We use the same 44 sites as for generating Fig. 12 and we keep the same frequency band ([2 20] Hz). Fig. 16 shows the resulting values converted to damping. For most sites, the estimated damping shows an apparent frequency dependence. Given the wide spread in $\partial d/\partial f$ derivatives—including positive slopes—the estimated frequency dependence per site is likely spurious. This would be due to the inability to select pure upgoing and downgoing waveforms for the deconvolution process. In the Groningen high-scattering environment, there is always some level of interference with other waves. With the assumption that interference is random, the mean over 44 sites does give a robust measure of field-wide frequency dependence. After taking the mean over all sites, a damping estimate remains that shows a modest drop of damping (increase in Q) with frequency over the used frequency range (blue lines in Fig. 12).

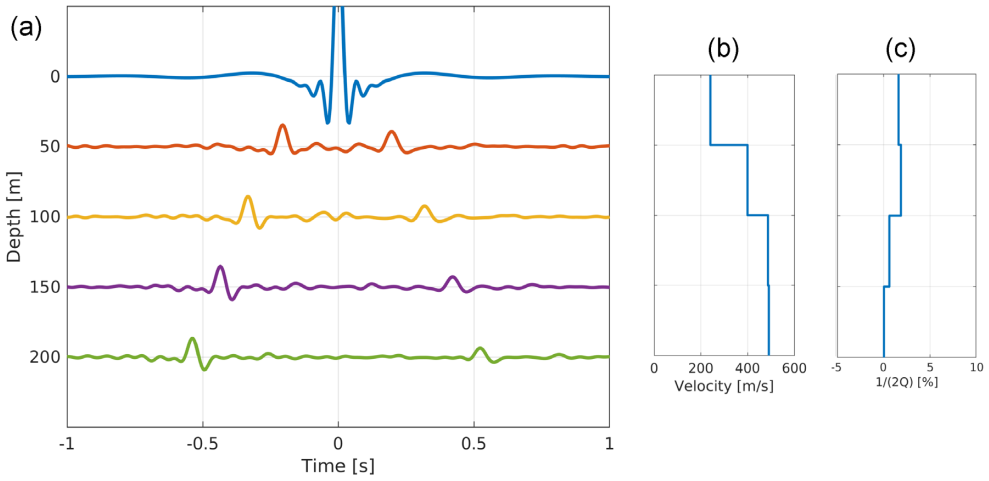

Figure 11. (a) The results of applying seismic interferometry by deconvolution at station G36 with a virtual source constructed at 0 m depth. The transverse-component result is shown in the band [2 20] Hz. (b) *S*-wave interval velocity model estimated from the timing of the upgoing wave in (a). (c) Damping interval model estimated from the amplitude differences between the up- and downgoing waves in (a).

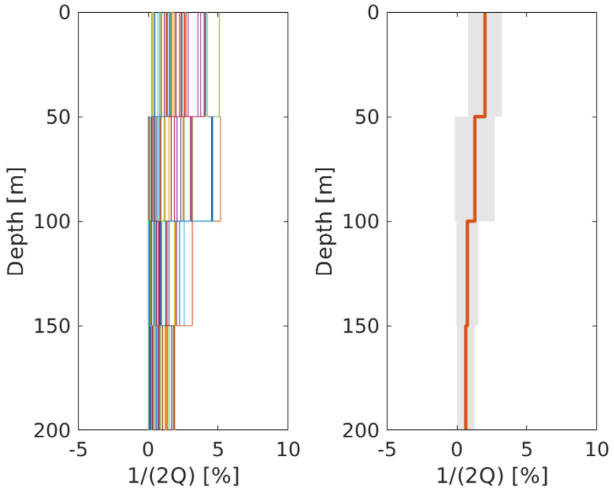

Figure 12. Damping profiles estimated with the up-down method at 44 sites of the G-network. Left: distribution of the 44 profiles. Right: mean (orange line) and spread (-sigma to sigma) of damping values over the network. The mean Q values for the four depth intervals are—from top to bottom—25, 38, 76 and 87.

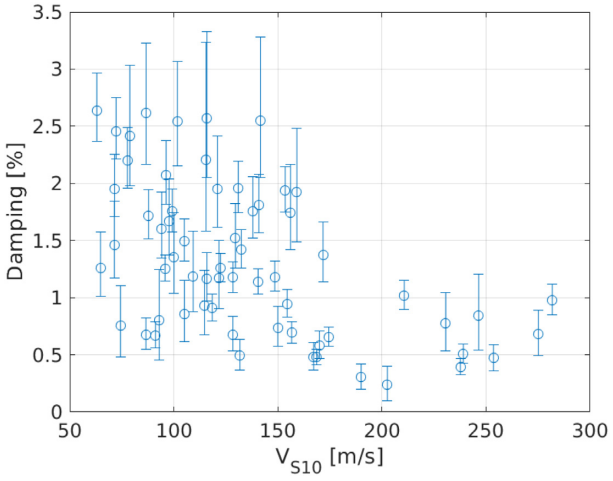

Figure 13. Estimated damping over the top 200 m plotted as function of *S*-wave velocity over the top 10 m. The circles depict the mean damping values for 63 sites. The error bars show the 68 per cent confidence interval (Appendix C).

As a sanity check, the frequency-domain results are similar to the ones obtained with the maximum method, after averaging over stations and over frequencies. For example, for the top 50 m, an average damping is estimated of 1.91 per cent or $Q = 26$. This is very close to the average value found with the up-down method (Fig. 12) which is $Q = 25$.

In Appendix D, the results of the up-down method are compared with results obtained with modelling aided spectral-ratio approach. It is shown that, averaged over the sites, the results of the up-down method are similar to the ones of the spectral-ratio method. However, for individual sites, large differences exist in the estimated κ_0. The spectral-ratio results are sensitive to the exact frequency-band used and the assumed TF.

3.3 Damping at depths below 200 m

The unconsolidated sediments from the North Sea Group are on average ~800 m deep below Groningen (Kruiver *et al.* 2017). Using VAs from the G-network, we obtained damping estimates in the top 200 m. To obtain also an estimate of damping at larger depths, surface waves are used. Over the G-network, continuously microseism surface waves are recorded which have their origin in ocean gravity waves (Kimman *et al.* 2012). In the so-called double-frequency (DF) peak, the microseism is dominated by wave-wave interactions in the Pacific and Atlantic Ocean and has a peak near 0.2 Hz, as in the new high noise model (NHNM; Peterson (1993)). In Groningen, the DF microseism is a-typical. Fig. 17 shows a 1 yr distribution of power-spectrum density (PSD) levels at a 100 m deep borehole. The peak frequency is with ~0.4 Hz considerably higher than in the NHNM. Moreover, between 0.3 and 2 Hz the power levels are mostly higher than in the NHNM. This high-frequency microseism has an origin in the North Sea, which is separated from mainland Groningen only by the shallow Wadden Sea.

We select a day with a high sea state at the North Sea: 2017 June 16, and compute for each sensor the PSD on the vertical component. The source of the microseism is much further north than the coast.

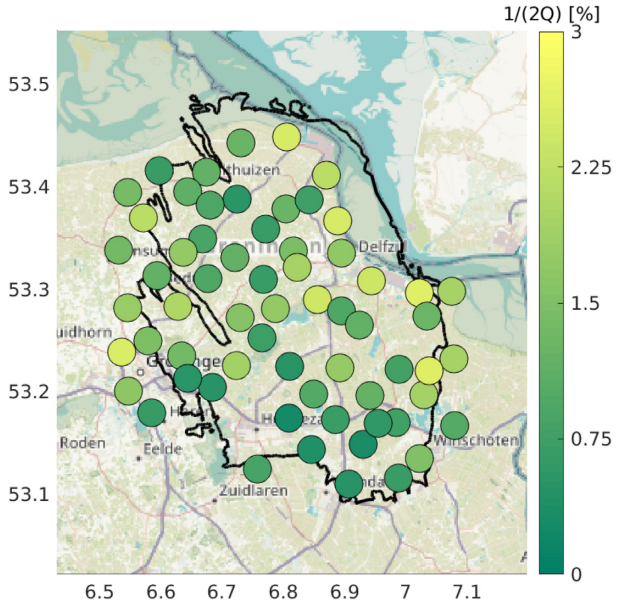

Figure 14. Estimated mean damping ($1/(2Q)$) over the top 200 m at 63 sites over Groningen. The black line denotes the outline of the Groningen gas field. The background map is from openstreetmap.org. Coordinates are in degrees latitude and longitude (WGS 84).

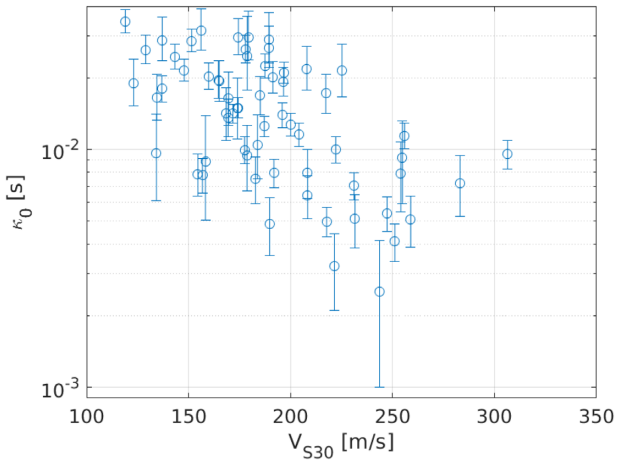

Figure 15. Estimated κ_0 over the top 200 m plotted as function of *S*-wave velocity over the top 30 m. The circles depict the mean κ_0 values for 63 sites. The error bars show the 68 per cent confidence interval (Appendix C).

As a consequence, incoming waves are nearly planar over the G-network. Thus, geometrical spreading is assumed negligible and the power decay is modelled with an exponential term only:

$$P(x) = P_0 e^{\frac{-\omega x}{Qv}}, \tag{25}$$

where ω is the angular frequency, x the distance from the active coast and v the propagation velocity. P_0 is the power of the microseism at the position of the active coast.

P_0 and Q are estimated from the data. The 50 m depth level is used as it records high levels of microseism, but is not much affected by, e.g., tilting noise, like some of the surface installations. The recorded PSDs are fitted with eq. (25) for frequencies from 0.4 to 0.8 Hz, with steps of 0.1 Hz. The Rayleigh-wave phase velocity is frequency dependent. The values are taken from the dispersion curve estimated in Fokker & Ruigrok (2019). Fig. 18 shows the measured power levels and the fitted models for the five different frequency bins.

The Rayleigh wave at 0.8 Hz has a maximum sensitivity at approximately $1/3\lambda \approx 500/0.8/3 = 200$ m depth. The Rayleigh wave at 0.4 Hz has a maximum sensitivity at approximately $1/3\lambda \approx 700/0.4/3 = 580$ m depth. Therewith, a large part of the NSG sediments is sampled. Fig. 18 shows no indications of a frequency or depth dependence of the NSG damping below 200 m depth. From the five frequencies, Q is 99 ± 2. Q standard deviations of less than 2 are obtained when bootstrap resampling the damping estimates per frequency.

4 FIELD-WIDE DAMPING MODEL

The prediction of site response for hazard analyses at the Groningen field requires a damping model that can be applied throughout the entire field (Rodriguez-Marek *et al.* 2017); not only at the sites that are sampled with the G-network. For this purpose, we use the up-down estimates of Q (Section 3.2). These estimates correspond to average Q (or equivalently, damping) values from the surface to the depth of the instrument used in the estimation of Q. To extrapolate these values to different locations across the field, we first use the GeoTOP model (Stafleu *et al.* 2011) to obtain a detailed soil profile for any location across the field. We then build initial low-strain damping (d) profiles using published models for low-strain damping linked to soil properties. Finally, these d profiles are scaled to match the average damping measured for the stations of the G-network. This approach results in low-strain damping profiles that are consistent with the d dependency on soil type and on confining stress that is well established through laboratory testing, and allows for a rational extrapolation of the measured damping at the G-network stations to any location in the Groningen field. The key element in this approach is the scaling factor (D_{fact}) that is needed to scale the laboratory-based low-strain damping profiles. The scaling factor is obtained using the following steps:

(i) For each of the G-network stations where Q was estimated, the average effect of damping from the deepest instrument (200 m depth) to the surface is estimated using the κ_0 parameter (eq. 24 and Fig. 15).

(ii) Laboratory-based low-strain damping (d_{lab}) profiles for each station are constructed using the laboratory-based models of Menq (2003) for sands, Darendeli (2001) for clays, and a model for peats developed from tests of Groningen peats (Zwanenburg *et al.* 2020). The soil index parameters that are input parameters to these models are obtained either from correlations to the Cone Penetration Tests (CPTs) tip resistances measured at the stations, or from parameters associated with each soil type.

(iii) The d_{lab} profiles are integrated to obtain the equivalent effect of damping over the entire column using:

$$\kappa_{0,\text{lab}} = \int_{z=0}^{z=200} \frac{2d_{\text{lab}}}{V} \mathrm{d}z, \tag{26}$$

(iv) For each station, the value of D_{fact} is obtained by equating the values from eq. (26) with the values obtained from eq. (24): $\kappa_0 = D_{\text{fact}}\kappa_{0,\text{lab}}$. The resulting D_{fact} values are plotted in Fig. 19 as a function of V_{S30}.

Observe that the D_{fact} is strongly correlated with V_{S30}. Therefore, a linear model [$\ln(D_{\text{fact}})$ versus $\ln(V_{S30})$] is fitted to the data in Fig. 19. Constraints are placed to limit the maximum values of D_{fact} to not exceed the value predicted for the lowest V_{S30} for all

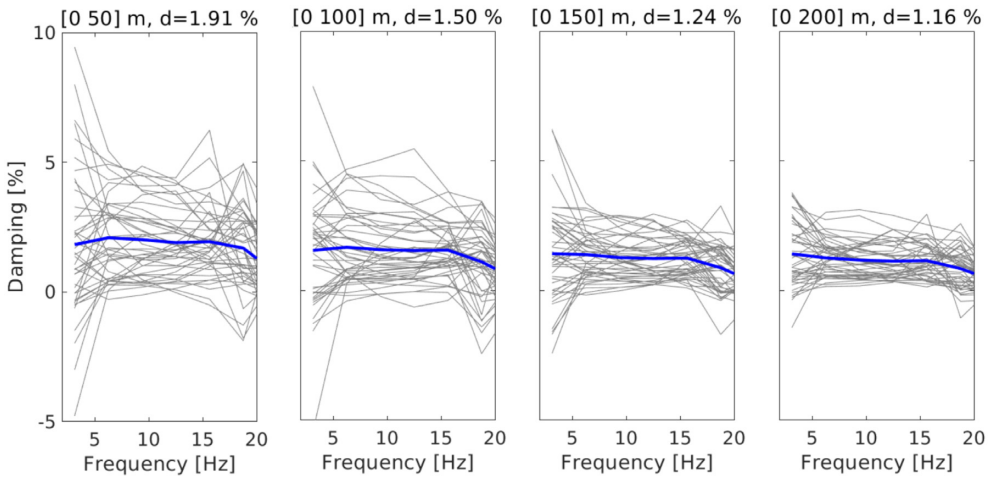

Figure 16. Estimated frequency-dependent damping (grey lines) for 44 sites over Groningen and the average over these sites (thick blue line). From left to right, the estimates are shown for the top 50, 100, 150 and 200 m, respectively.

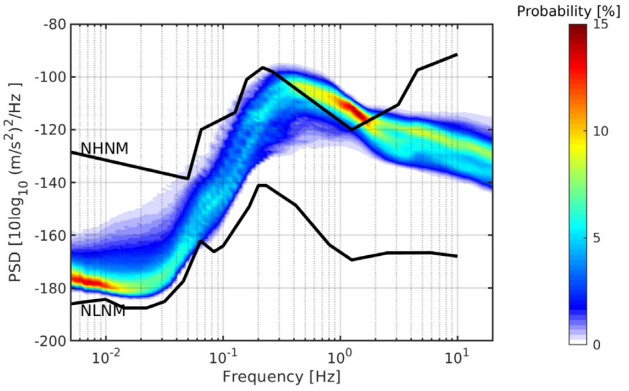

Figure 17. The vertical-component power-spectrum density distribution in 2020 at station G81B. For computing this distribution, the recipe in McNamara & Buland (2004) was followed. The measured distribution is compared with the new low noise model (NLNM) and new high noise model (NHNM) from Peterson (1993).

the stations considered, and a lower bound of 1.0 is placed to avoid reducing the laboratory estimates of damping. The lower limit is a conservative choice based on the fact that field estimates of damping have been often observed to exceed laboratory estimates due to effects not captured in a laboratory sample (i.e. scattering due to small-scale heterogeneities in the field (e.g. Cabas *et al.* 2017)). On the other hand, there is scant evidence that field damping should be lower than laboratory estimates. The central damping model in Fig. 19 corresponds to a best-estimated model for the field. To obtain uncertainty bounds on this model, two uncertainties were considered: the uncertainty in Q estimates (Appendix C), which were propagated to uncertainties in the estimated D_{fact} and uncertainties in the regression of the linear models described above. The latter dominate the total uncertainty. An upper and lower model for the D_{fact} is then obtained using the same approach as for the central value, but using the upper- and lower-bound of the 95 per cent confidence intervals from the regression fit of the κ_0 versus V_{S30} relationship. The upper and lower limits on D_{fact} (seen as the horizontal bounds to the D_{fact} models in Fig. 19) were developed based on the judgment of the authors. The models in Fig. 19 can be used to develop low-strain damping profiles for any location in the field where a stratigraphic column can be determined and V_{S30} is measured or estimated. In Groningen, these columns are available for each 100 x 100 m grid-cell.

5 DISCUSSION

In the preceding sections, we considered two interferometric methods for estimating Q: the upgoing and the up-down method. The upgoing method was initially seen as favourable. The upgoing retrieved waves have higher SNR than downgoing waves and exhibit less interference with complex arrivals. However, during the course of this research we found clear disadvantages for the upgoing method. A detailed local velocity model is needed to apply corrections for elastic propagation effects. Moreover, the *in situ* response of the geophones needs to be known in great detail.

In Groningen, the nominal factory settings are used for instrument deconvolution. The down-hole geophones have been designed to have a 4.5 Hz resonance frequency, a coil resistance of 3500 Ω, a sensitivity of 78.9 V m^{-1} s^{-1} and an open-circuit damping of 0.58. The damping is increased to 0.7 by addition of a shunt resistor. The *in situ* responses may have too large an offset from the nominal responses to allow the extraction of a refined amplitude attribute such as damping. There are many possible reasons for the *in situ* response to be different to that of the nominal:

(i) Geophones at depth operate at different pressures and temperatures than in the reference condition (1×10^5 Pa and 20 °C).
(ii) Geophones are at the end of long cables, up to 205 m, of which some might be damaged.
(iii) Geophones are sensitive to tilt.
(iv) The coupling affects the response.
(v) Different electronic components could be susceptible to aging and fatigue. For instance, the permanent magnet in the coil could reduce in strength and total resistance could be increased due to oxidation at connection points.

Downhole geophones are also deployed at Hi-net (Japan). These are 1.0 Hz geophones with a similar damping as used in Groningen. The geophones are deployed in a pressurized vessel that has been lowered in a cased borehole. The horizontal components are tilt sensitive. However, a system has been built that adjusts for tilt (Obara *et al.* 2005). The *in situ* resonance frequency and damping

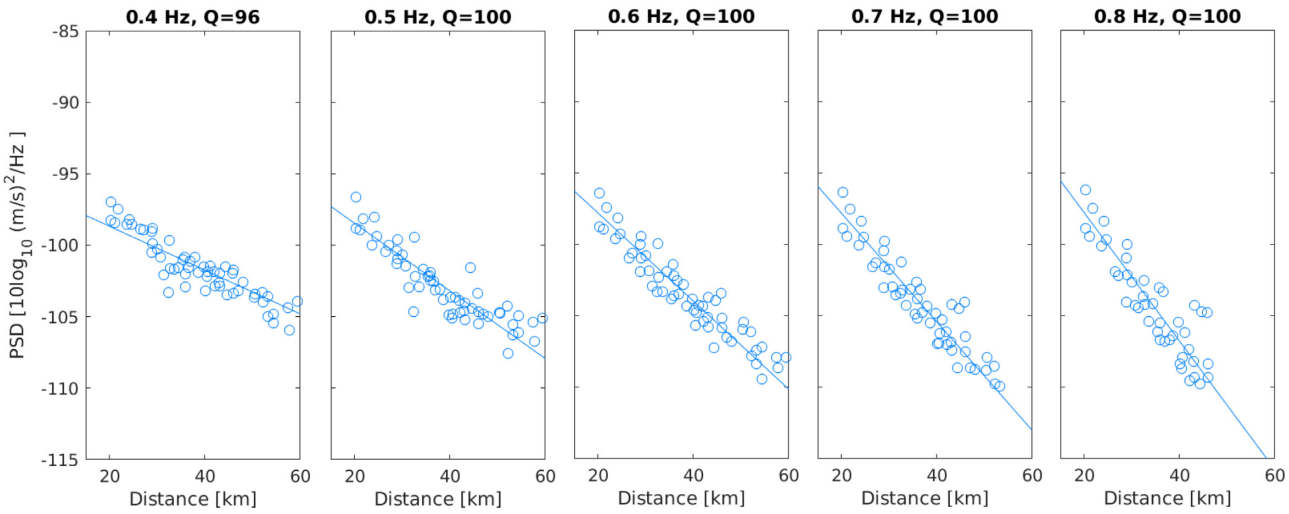

Figure 18. Recorded PSDs (circles) and fitted attenuation curves (lines) for five different frequencies. The recordings at 50 m depth are used. The Q values as estimated in the fitting procedure are shown above the curves. The PSDs are shown as function of distance from the active North-Sea coast, which is defined at the north side of the Wadden-Sea barrier islands, as an approximate west-east line 20 km north of the G-network.

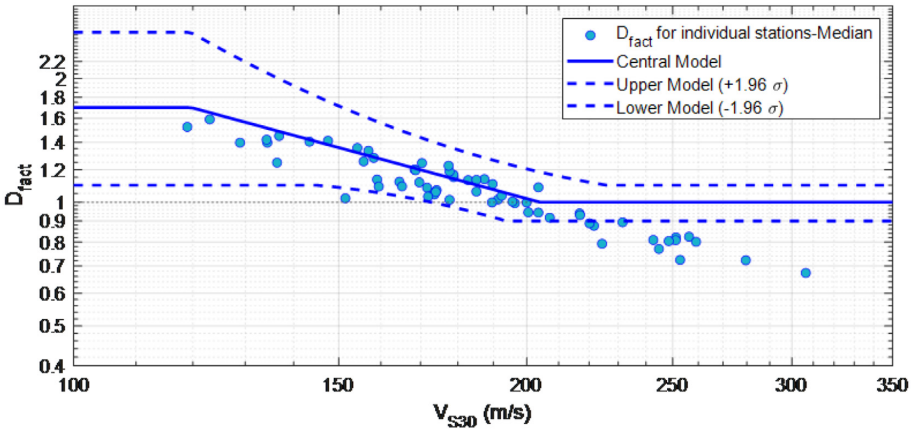

Figure 19. Estimates of D_{fact} for stations in the G-Network plotted against the V_{S30} for the station. The upper and lower models correspond to a 95 per cent uncertainty bound that consider uncertainty in each of the values of D_{fact} and uncertainty on the model fit.

are measured every day, using a test-coil signal. Ueno *et al.* (2015) describe the results over a 10-yr period. Most of the sensors showed fluctuations of less than 5 per cent over that time frame, thus showing little ageing effects. Over the entire Hi-net more than 95 per cent of the sensors were found to have natural frequencies in the range of [1.0 1.2] Hz and damping constants in the range of [0.6 0.8]. In Groningen, the response variations over the network are likely larger than 10 per cent. A VA is constructed by flushing a temporary hole, lowering a geophone cable and filling up the hole again with indigenous material. This results in a less controlled environment than at Hi-net. The exact impact on the *in situ* response is material for a follow-up study.

The main advantage of the up-down method is that propagation corrections are not needed and that most sensor imperfections are left out of the equation. An imperfectly known instrument response would similarly affect the up- and the downgoing wave measured at the same geophone. The main drawback of the method is that at sites with very strong damping and/or scattering, the downgoing wave has amplitude levels close to the amplitude level of scattered waves. This leads to a large uncertainty in the damping value. However, this uncertainty can be well quantified (Appendix C).

With the spectral-ratio method (e.g. Rodriguez-Marek *et al.* 2017) usually the assumption is made that elastic propagation effects are frequency independent. This assumption breaks down in Groningen where resonances exist over near-surface layers. We applied an adapted spectral-ratio method in which we corrected for the elastic TF (eq. D3). Though the Groningen near-surface is quite well known (e.g. Fig. 3) for many sites the model is not good enough to exactly mimic the actual resonances. As a consequence, the spectral-ratio results remain susceptible to these resonances, especially below 10 Hz where they are strongly present in the spectra.

The largest damping value measured in the field was 5.1 per cent, which was obtained over the top 50 m at site G43 (Fig. 12). At this site, a large fraction of the top 50 m consists of clay, along with a few meters of Holocene peat. The measured damping is lower than the largest damping value measured in the lab when sampling local peaty soils (Zwanenburg *et al.* 2020). They measured the strongest damping for the upper peat deposits (Hollandveen) which have a shear-wave velocity of about 35 m/s and small-strain damping of about 6 per cent.

He *et al.* (2021) applied a spectral-ratio method to induced events measured over the Groningen VAs. They did not apply a propagation correction, but utilized a multitime window approach to mitigate instability of the spectral ratio. They found a strong frequency dependence of Q between 2 and 10 Hz, with very strong damping at 2 Hz (values between 4.5 and 9 per cent). With a frequency-domain implementation of the up-down method (Fig. 16) we found only a weak frequency dependence. Fukushima *et al.* (2016) also concluded that between 2 and 10 Hz little frequency dependence could be observed for sediments at KiK-net stations in Japan.

He *et al.* (2021) show spatial plots of estimated Q^{-1} values at 2, 4, 6 and 8 Hz. Their results at 8 Hz are best compared with ours, considering that we estimated damping over a band between 2 and 20 Hz. For the depth interval 0 to 200 m, they found Q^{-1} values ranging between about 0.02 and 0.05, which equates to damping values between 1.0 and 2.5 per cent. For the same depth range, we find values between 0.24 and 2.6 per cent (Fig. 13).

In Fig. 19, it was shown that the estimated damping values, on average, come close to laboratory results; on average D_{fact} is about 1.0. However, for sites with a $V_{S30} < 200\,\text{m s}^{-1}$ the measurements show higher damping values. For sites with $V_{S30} > 200\,\text{m s}^{-1}$, on the other hand, the measurements show smaller values than the laboratory results. Several reasons can explain the differences between laboratory and field estimates of damping, including differences in the frequency range used in laboratory tests and that used for field estimates, differences in the soil micro-structure that result from sampling, and potential differences in confining stress. In addition, laboratory-based models for low-strain damping are developed using soils that have different geological origins but have common geotechnical-based index parameters, thus are bound to have a large scatter.

Often, Q is subdivided into a scattering Q and an intrinsic Q. The former describes losses due to small-scale scattering of the wavefield, the latter describes losses due to anelastic attenuation. In the frequency band in which we applied the Q estimation ([2 20] Hz) 3-D scattering is thought to be small. This notion comes from the measured TFs, which were constructed by using incoming waves with steep angles incidence (Section 2.4). These TFs could be explained well with measured 1-D profiles (e.g. Fig. B1). 1-D transmission loss was taken into account for the upgoing method (Appendix A) and 1-D scattering was shown not to impact the relative amplitudes in the up-down method (Fig. 4). Hence, the estimated damping values are likely a fair estimate of intrinsic Q. In this research we did not study the implications of applying the methods in settings where 3-D effects are important.

6 CONCLUSION

In this analysis we considered two methods for estimating damping from a vertical-array recording of local seismicity. With both methods, the first step is the application of seismic interferometry by deconvolution to obtain an estimate of the local TF. With the upgoing method, the amplitude decay of the retrieved upgoing waves is used to estimate damping. With the up-down method, the amplitude difference between the retrieved upgoing and free-surface reflected downgoing wave is used. We found the up-down method to be more robust, because of the following reasons: (1) no error-prone elastic corrections need to be applied and (2) small differences in effective instrument response between different depth levels are irrelevant.

For the upgoing method, the amplitude of the upgoing direct wave is affected by both elastic and anelastic effects. In order to estimate the anelastic attenuation it was necessary to remove the elastic amplification first. Dependent on the knowledge of the subsurface between different recording depth levels, we discussed three possible approximations: (1) one-interface model, (2) gradient model and (3) multilayered model. For the Groningen near surface, Models 1 and 2 were too great a simplification. Model 3 was implemented using, among others, a measured S-wave profile over the top 30 m. Despite the fact that elastic compensation could be accurately determined, non-physical damping values were estimated for a number of boreholes. We consider it likely that the underlying cause was small differences in the effective response functions of the different geophones. These differences are not reflected in the nominal factory response, which was assumed for the processing.

In the up-down method, the damping is estimated from the deconvolution result H (eq. 15). H is not a physical TF, since it does not contain 1-D elastic propagation effects between the up- and downgoing waves, although it does contain the anelastic effects. We showed with 1-D numerical modelling (Fig. 4) that H needs to be estimated with long time windows of local events. Only if the complex reverberations are included in the interferometric process, the direct wave in H is corrected for elastic transmission loss and Q estimation can be made without knowing the 1-D soil profile in detail. We did not study the influence of 3-D effects as we deemed a local 1-D approximation valid at most Groningen sites, for the frequency band used.

With the up-down method, the damping could be estimated as function of frequency. For individual sites, we found these estimates to be highly oscillatory. Averaged over all sites, a stable curve was found with weak frequency dependence between 2 and 20 Hz.

We also applied a variant of the spectral-ratio method to deconvolved waveforms (Appendix D). Deconvolutions were taken of local events with magnitudes higher than $M_{\text{L}} = 2.5$. This allowed utilizing a larger frequency band than was used in the up-down method. We found the spectral-ratio results to be highly susceptible to the exact frequency band that was used.

For estimating damping we used interferometric theory. That is, deconvolutions of recordings at various depth levels where stacked over local events. This yields a best estimate of the local TF and a best estimate of damping. In other implementations (e.g. Fukushima *et al.* 2016) parameters are derived from each event, to yield a distribution in parameters from which the uncertainty is derived. For our best estimate we derived the uncertainty from the SNR of the used data attributes, as shown in Appendix C. For a homogeneous model we tested both approaches and found little difference in the estimated mean value.

Damping of unconsolidated layers below 200 m depth was found from the attenuation of the microseism over Groningen. For a depth range between about 200 and 580 m, a Q of 99 was found, with no notable variation over this depth range.

The estimated Q profiles were used to derive a damping model for the top 200 m of the entire Groningen field. A scaling relation was derived by comparing estimated Q profiles with low-strain damping profiles that were built using published models for low-strain damping linked to soil properties. This scaling relation, together with the soil-properties based damping model, allowed upscaling the model to each 100 x 100 m grid-cell in the Groningen field. The resulting model is consistent with the estimated Q values over the G-network and with the stratigraphic profiles of the Groningen field.

ACKNOWLEDGMENTS

The work of ER has partly been funded by TURNkey (Towards more Earthquake-resilient Urban Societies through a Multi-sensor-based Information System enabling Earthquake Forecasting, Early Warning and Rapid Response actions), which is a project funded by the EU within the H2020 – SC5 – 2018 – 2 call. We would like to thank Jun Matsushima and an anonymous reviewer for a detailed and critical assessment.

DATA AVAILABILITY

Waveform data from the Netherlands Seismic and Acoustic Network (KNMI 1993) is openly available. Waveform data, station metadata and event catalogues can be obtained through http://rdsa.knmi.nl/. The 2021 update of the Groningen ground motion model, to which the damping model is one of the ingredients, has been published as Bommer *et al.* (2021).

REFERENCES

Anderson, J.G. & Hough, S.E., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. seism. Soc. Am.,* **74**(5), 1969–1993.

Aoi, S. *et al.*, 2020. MOWLAS: NIED observation network for earthquake, tsunami and volcano, *Earth Planets Space,* **72**(1), 1–31.

Bommer, J.J. *et al.*, 2017. Framework for a ground-motion model for induced seismic hazard and risk analysis in the Groningen gas field, the Netherlands, *Earthq. Spectra,* doi:10.1193/082916EQS138M.

Bommer, J.J. *et al.*, 2019. V6 ground-motion model for induced seismicity in the Groningen gas field, NAM Study Report.

Bommer, J.J. *et al.*, 2021. V7 ground-motion model for induced seismicity in the Groningen gas field, NAM Study Report.

Cabas, A., Rodriguez-Marek, A. & Bonilla, L.F., 2017. Estimation of site-specific kappa (κ_0)-consistent damping values at KiK-net sites to assess the discrepancy between laboratory-based damping models and observed attenuation (of seismic waves) in the field, *Bull. seism. Soc. Am.,* **107**(5), 2258–2271.

Darendeli, M.B., 2001. *Development of a New Family of Normalized Modulus Reduction and Material Damping Curves,* The University of Texas at Austin.

Dost, B., Ruigrok, E. & Spetzler, J., 2017. Development of probabilistic seismic hazard assessment for the Groningen gas field, *Neth. J. Geosci.,* **96**(5), s235–s245.

Edwards, B., Ktenidou, O.-J., Cotton, F., Abrahamson, N., Van Houtte, C. & Fäh, D., 2015. Epistemic uncertainty and limitations of the κ_0 model for near-surface attenuation at hard rock sites, *Geophys. J. Int.,* **202**(3), 1627–1645.

Fokker, E. & Ruigrok, E., 2019. Quality parameters for passive image interferometry tested at the Groningen network, *Geophys. J. Int.,* **218**(2), 1367–1378.

Fukushima, R., Nakahara, H. & Nishimura, T., 2016. Estimating S-wave attenuation in sediments by deconvolution analysis of KiK-net borehole seismograms, *Bull. seism. Soc. Am.,* **106**(2), 552–559.

Haendel, A., Ohrnberger, M. & Krüger, F., 2019. Frequency-dependent quality factors from the deconvolution of ambient noise recordings in a borehole in West Bohemia/Vogtland, *Geophys. J. Int.,* **216**(1), 251–260.

He, Y., Li, J., Tian, W. & Li, Z., 2021. Characterization of the near-surface shear wave attenuation in the Groningen gas field using borehole recording, *Geophys. J. Int.,* **226**(3), 2057–2072.

Hofman, L., Ruigrok, E., Dost, B. & Paulssen, H., 2017. A shallow velocity model for the Groningen area in the Netherlands, *J. geophys. Res.,* **122,** 8035–8050.

Kimman, W., Campman, X. & Trampert, J., 2012. Characteristics of seismic noise: fundamental and higher mode energy observed in the northeast of the Netherlands, *Bull. seism. Soc. Am.,* **102**(4), 1388–1399.

KNMI, 1993. *Netherlands Seismic and Acoustic Network,* Royal Netherlands Meteorological Institute (KNMI), Other/Seismic Network , doi:10.21944/e970fd34-23b9-3411-b366-e4f72877d2c5.

Kruiver, P.P. *et al.*, 2017. An integrated shear-wave velocity model for the Groningen gas field, the Netherlands, *Bull. Earthq. Eng.,* doi:10.1007/s10518-017-0105-y.

Matsushima, J., Ali, M.Y. & Bouchaala, F., 2016. Seismic attenuation estimation from zero-offset VSP data using seismic interferometry, *Geophys. J. Int.,* **204**(2), 1288–1307.

McNamara, D.E. & Buland, R.P., 2004. Ambient noise levels in the continental United States, *Bull. seism. Soc. Am.,* **94**(4), 1517–1527.

Menq, F.-Y., 2003. *Dynamic Properties of Sandy and Gravelly Soils,* The University of Texas.

Nakata, N. & Snieder, R., 2012. Estimating near-surface shear wave velocities in Japan by applying seismic interferometry to KiK-net data, *J. geophys. Res.,* **117**(B1).

Noorlandt, R. *et al.*, 2018. Characterisation of ground motion recording stations in the Groningen gas field, *J. Seismol.,* **22,** 605–624.

Ntinalexis, M. *et al.*, 2019. Ground-motion networks in the Groningen field: usability and consistency of surface recordings, *J. Seismol.,* **23,** 1233–1253.

Obara, K., Kasahara, K., Hori, S. & Okada, Y., 2005. A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention, *Rev. Sci. Instrum.,* **76**(2), 021301.

Parolai, S., Bindi, D., Ansal, A., Kurtulus, A., Strollo, A. & Zschau, J., 2010. Determination of shallow S-wave attenuation by down-hole waveform deconvolution: a case study in Istanbul (Turkey), *Geophys. J. Int.,* **181**(2), 1147–1158.

Peterson, J.R., 1993. *Observations and modeling of seismic background noise,* U.S. Geological Survey, **93–322,** 94 p..

Rodriguez-Marek, A., Kruiver, P.P., Meijers, P., Bommer, J.J., Dost, B., van Elk, J. & Doornhof, D., 2017. A regional site-response model for the Groningen gas field, *Bull. seism. Soc. Am.,* **107**(5), 2067–2077.

Ruigrok, E., Campman, X., Draganov, D. & Wapenaar, K., 2010. High-resolution lithospheric imaging with seismic interferometry, *Geophys. J. Int.,* **183**(1), 339–357.

Ruigrok, E., Domingo-Ballesta, J., van den Hazel, G., Dost, B. & Evers, L., 2019. Groningen explosion database, *First Break,* **37**(8), 37–41.

Schimmel, M. & Paulssen, H., 1997. Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.,* **130**(2), 497–505.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E,* **69**(4), 046610.

Snieder, R. & Safak, E., 2006. Extracting the building response using seismic interferometry: Theory and application to the Millikan Library in Pasadena, California, *Bull. seism. Soc. Am.,* **96**(2), 586–598.

Spetzler, J. & Dost, B., 2017. Hypocentre estimation of induced earthquakes in groningen, *Geophys. J. Int.,* **209**(1), 453–465.

Stafleu, J., Maljers, D., Gunnink, J., Menkovic, A. & Busschers, F., 2011. 3D modelling of the shallow subsurface of Zeeland, the Netherlands, *Neth. J. Geosci.,* **90**(4), 293–310.

Tonn, R., 1991. The determination of the seismic quality factor Q from VSP data: a comparison of different computational methods, *Geophys. Prospect.,* **39**(1), 1–27.

Trampert, J., Cara, M. & Frogneux, M., 1993. SH propagator matrix and *qs* estimates from borehole- and surface-recorded earthquake data, *Geophys. J. Int.,* **112**(2), 290–299.

Ueno, T., Saito, T., Shiomi, K. & Haryu, Y., 2015. Monitoring the instrument response of the high-sensitivity seismograph network in Japan (Hi-net): effects of response changes on seismic interferometry analysis, *Earth Planets Space,* **67**(1), 1–10.

van Elk, J., Bourne, S.J., Oates, S.J., Bommer, J.J., Pinho, R. & Crowley, H., 2019. A probabilistic model to evaluate options for mitigating induced seismic risk, *Earthq. Spectra,* **35**(2), 537–564.

van Ginkel, J., Ruigrok, E. & Herber, R., 2019. Assessing soil amplifications in Groningen, the Netherlands, *First Break,* **37**(10), 33–38.

Van Houtte, C., Ktenidou, O.-J., Larkin, T. & Holden, C., 2018. A continuous map of near-surface s-wave attenuation in new zealand, *Geophys. J. Int.*, **213**(1), 408–425.

Wapenaar, K., 2006. Green's function retrieval by cross-correlation in case of one-sided illumination, *Geophys. Res. Lett.*, **33**(19), doi:10.1029/2006GL027747.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X. & Verdel, A., 2010a. Tutorial on seismic interferometry: Part 1 — basic principles and applications, *Geophysics*, **75**(5), 75A195–75A209.

Wapenaar, K., Slob, E., Snieder, R. & Curtis, A., 2010b. Tutorial on seismic interferometry: Part 2 — underlying theory and new advances, *Geophysics*, **75**(5), 75A211–75A227.

Zwanenburg, C. *et al.*, 2020. Assessment of the dynamic properties of Holocene peat, *J. Geotech. Geoenviron. Eng.*, **146**(7), 04020049.

APPENDIX A: CORRECTION FOR ELASTIC AMPLIFICATION

For estimating Q it is ideal to have multiple receivers in the same layer. This circumvents the need to make impedance and reverberation corrections. In Groningen, we do not have this ideal case. Especially between the free surface and the 50 m recording level, strong impedance contrasts exist. For most methods, propagation effect through the layering need to be taken into account before anelastic losses can be estimated.

During the upward trajectory of the wave front, waves generally amplify due to the reduction of impedance. At the same time, resonances may occur. These are frequency-dependent amplitude modulations due to waves reverberating between impedance contrasts.

Often the precise near-surface buildup is not known. However, one may have a good estimate of the impedance (density times velocity, ρv) at receiver levels z_2 and z_1. One possible assumption is that the impedance difference occurs at a single interface. In that case one can use the Zoeppritz equation to find the amplification term:

$$F(z_1, z_2) = \frac{2\rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2}. \tag{A1}$$

Another possible assumption is that there is a smooth impedance gradient. In that case, the amplitude at depth level z_1 with respect to the amplitude at z_2 is written as

$$F(z_1, z_2) = \sqrt{\frac{\rho_2 v_2}{\rho_1 v_1}}, \tag{A2}$$

Using eq. (A2) or (A1) makes a significant difference. Assume that both the density and velocity are halved going from z_2 to z_1. For this case, Fig. A1 shows the transmission amplitude as function of the amount of layers between the two depth levels. If there are a few layers only, the one-interface assumption is better. If there are many layers, the smooth gradient approximation is valid, which results in this example in a 25 per cent higher transmitted amplitude than with the one-interface assumption. For a more complicated subsurface structure, and when taking finite frequency effects into account, none of the two approximations is valid.

In Groningen, a detailed near-surface model is available. Hence, no end-member impedance-correction assumption (eq. A1 or A2) needs to be made, but the actual elastic effect can be modelled and corrected for:

$$F(z_1, z_2) = \left| \frac{\Gamma(z_1, z_R)}{\Gamma(z_2, z_R)} \right|, \tag{A3}$$

where $\Gamma(z_1, z_R)$ is the modelled TF from a reference depth z_R to receiver level z_1, in the frequency domain. For the direct wave and

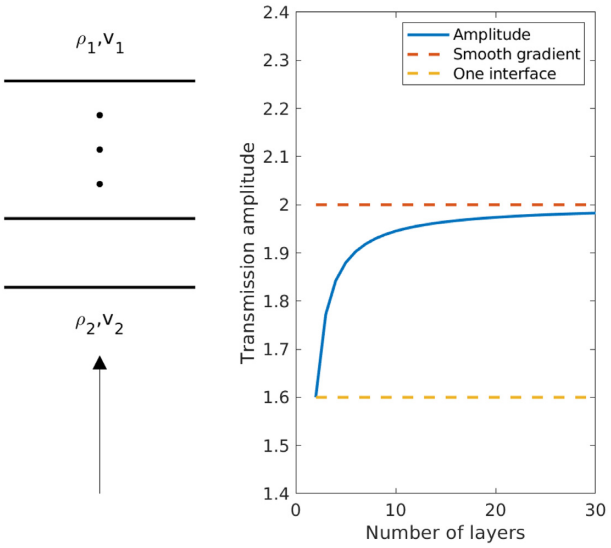

Figure A1. Left: a plane wave front with amplitude 1 impinges from below on a stack of layers and grows in amplitude due to reduction in impedance from depth level 2 to 1. Right: the transmitted amplitude above the stack of layers, as function of the amount of layers. It is assumed that the both the velocity and density are halved from depth level 2 to 1 and that the reduction in impedance occurs in equal steps from layer to layer.

when Q estimation is done in the time domain, the above function can be approximated as a frequency independent factor:

$$F(z_1, z_2) \approx \frac{\max(\gamma(z_1, z_R))}{\max(\gamma(z_2, z_R))}. \tag{A4}$$

Eq. (A4) is used, for example, to remove elastic propagation effects from eq. (8). The forward modelling of the TFs is done in the same frequency band that is used for the field-data processing. Modelling is performed using the reflectivity code of Wapenaar (2006), rewritten from acoustic to *SH* propagation.

APPENDIX B: VELOCITY-MODEL UPDATING

The previous appendix shows importance of a well-calibrated velocity model. For the upgoing method, only a good estimate of Q can be made when elastic amplitude modulation is first removed. Therefore we seek to update an existing model using new measurements to derive an optimized velocity model.

For the forward modelling we use the *S*-wave model from Kruiver *et al.* (2017) as a base. The top 50 m in that model was obtained by assigning seismic velocities to lithostratigraphic layers in a detailed 3-D geologic model which is called GeoTOP (Stafleu *et al.* 2011). The depth range between 50 and 120 m depth was obtained through a surface wave inversion over a dense grid of active sources and receivers.

In 2019 and 2020, Cone Penetration Tests (CPTs) and Shear wave Cone Penetration Tests (SCPTs) were made at almost all G-network stations. These measurements of *S*-wave velocity V_S over the top 30 m are used to locally update the Kruiver model. Fig. B1(a) shows, for one borehole station, the Kruiver model in grey and the locally measured *S*-wave model in black. The density model is taken entirely from Kruiver *et al.* (2017). The quality of the resulting model can be checked by comparing modelled and measured TFs. The measured TF is obtained with deconvolution seismic interferometry

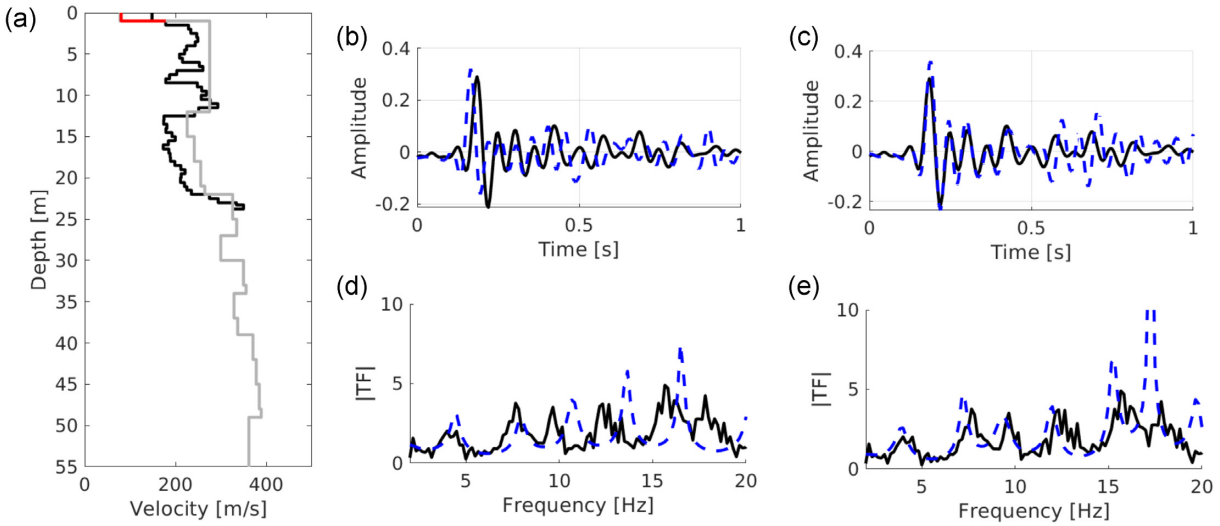

Figure B1. Example of using the SCPT and locally estimated TF for updating the V_S profile. (a) The GeoTOP-based velocity profile at G56 (grey line) with the upper part replaced with the SCPT measurement (black line) from which the top 1 m, is replaced by a model (red line) consistent with the measured TF between 50 and 0 m depth. Panels (b) and (c) show the measured (black line) and modelled (blue dashed line) TF (b) before and (c) after updating the model. Panels (d) and (e) show the corresponding amplitude spectra (d) before and (e) after updating the model. The *SH*-wave TF is modelled without losses. Hence, the modelled amplitudes are somewhat higher than the observed amplitudes, especially for the higher frequencies.

using local seismicity (eq. 23). The elastic modelling is performed for the stack of layers. The TF for the Kruiver model has a direct wave (first pulse on the blue line in Fig. B1b) that is faster than in the measured TF (black line in same figure).

The SCPTs and the measured TFs between the 50 and 0 m depth level are used to update the *S*-wave model. In the top metres ($\leq$5 m) the SCPT interpretation is not reliable (Noorlandt *et al.* 2018) due to overlapping *P* and *S* waves, short traveltimes and noisy records. Although it makes up only a small portion of the 50 m column, the top 5 m of the model have a large effect on the TFs, both in respect of correct timing of the direct wave as in explaining free-surface reverberations. The upper part of the soil column is adjusted with velocity values that match the timing of the TF and that are consistent with the lithology description that comes from the CPT. For a few stations, the model is also updated between 30 and 50 m depth. If the lowest SCPT *S*-wave recording shows values that are largely different from the GeoTOP derived model, a taper is applied to go from SCPT values at 30 m depth to GeoTOP values at 50 m depth. The updated model (Kruiver + SCPT + TF-based update) is used in the following and referred to as the optimized model. Fig. B1 illustrates the updating process for station G56. SCPT's are available at all 69 G-network borehole stations, with the exception of G05, G06, G07, G29, G43, G52, G53, G54, G58, G59, G60, G63, G68 and G69.

APPENDIX C: ERROR MODEL

In the main text, a damping model is derived from the amplitude difference of the up- and downgoing waves obtained at borehole sensors. Both the up- and downgoing waves are affected by noise, which limits the ability to find the true damping. This noise could have different origins. Either the noise comes from waves that are not vertically propagating, but which nevertheless have not been stacked out completely in the interferometric process (Section 2.4). Or, it are waves that are part of the TF which concur with the up- and downgoing direct wave. Here we assess the impact on the noise on the damping estimate.

In eq. (18), there are three (sets of) parameters that are derived from the data: Δt_i, E^+/E^- and $(F^- + F^+)$. In the following, we assume that a (noise-induced) error on the amplitude (envelope) quotient E^+/E^- has the strongest impact on the damping estimate.

Fig. 4(b) shows a response as retrieved in the ideal case. In this lossless case, the acausal and causal parts are symmetric. Moreover, the non-zero response is bounded within $-T_r(z)$ and $+T_r(z)$, where $T_r(z)$ is one-way traveltime from depth level z to the Earth's surface plus an additional margin of $1/f^d$, where f^d is the dominant frequency. In an actual case, the illumination is not perfect and noise is present, leading to imperfect cancellation of spurious terms, which we coin SInoise. For example, this SInoise can be identified in Fig. 11 prior to $-T_r(z)$ and after $+T_r(z)$. Hence, the SNR of the up- or downgoing wave we can be expressed in dB as

$$\mathrm{SNR} = 10 \log_{10} \frac{P_S}{P_N}, \tag{C1}$$

where P_S is the average power in a 0.1 s time window around the up- or downgoing wave and P_N is the average power in a 0.3 s time window prior to $-T_r(z)$.

For mapping SNR to amplitude errors, we take a Monte Carlo approach, with the assumption that amplitude errors due to different noise realizations are normally distributed. In Hofman *et al.* (2017), a similar approach was taken to assess the uncertainty in velocity that was estimated with cross-correlation seismic interferometry. The blue seismogram in Fig. C1(a) shows a synthetic direct wave with a similar frequency content as the Groningen data. This synthetic wave is represented by a Ricker wavelet with a dominant frequency of 10 Hz, time shifted by 0.5 s. Different generations of noise with similar frequency content as the signal are added, yielding seismograms like the green one in Fig. C1(a). From the noisy direct wave, we take the envelope and compute the envelope error ϵ_E, normalized with the envelope of the clean signal. For each

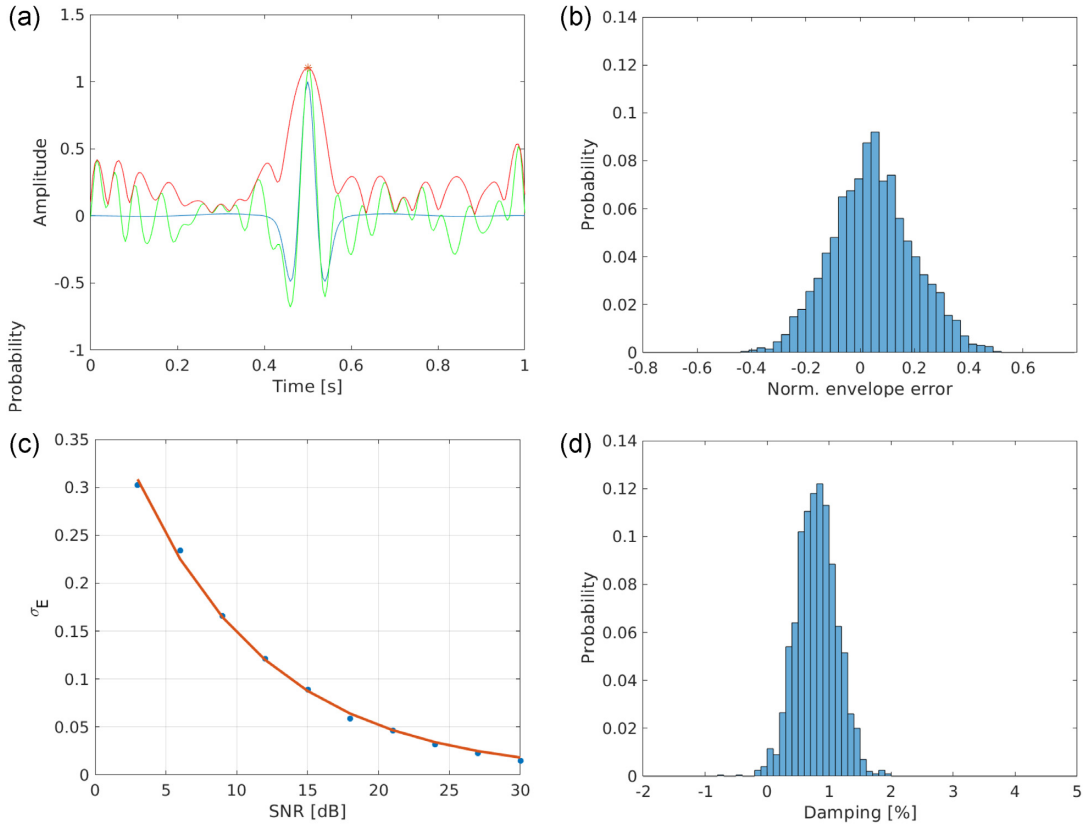

Figure C1. The panels show different steps of error propagation between amplitude (envelope) errors to a damping confidence region. (a) Clean wavelet (blue trace) distorted with noise with a SNR of 10 dB (green trace) and the envelope of signal+noise (red trace). The maximum envelope (red star) has an amplitude error of 0.1 or 10 per cent. (b) Amplitude (error) distribution for 2000 synthetic runs with different noise realizations with SNR=10 dB, yielding an amplitude standard deviation σ_E of 0.15. (c) SNR versus σ_E data points (blue dots) and a fitted function (red line). (d) Estimated damping distribution for Groningen scenario with a 200 m depth interval and an SNR of 10 dB.

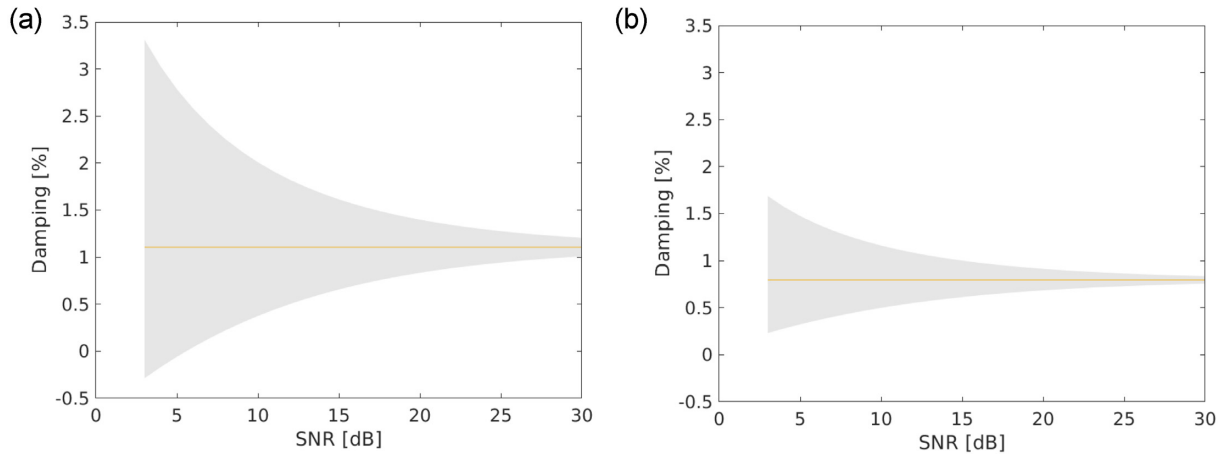

Figure C2. 68 per cent damping confidence regions (grey areas) as function of SNR, for (a) the top 50 m of soil scenario and (b) the top 200 m scenario.

SNR we compute 2000 noise realizations and corresponding envelope errors. This yields the ϵ_E probability distributions as shown in Fig. C1(b). From this distribution, the amplitude standard deviation σ_E is taken for this SNR. The process is repeated for a range of SNRs and the resulting SNR versus σ_E data points are fitted with a decaying exponential function (Fig. C1c) yielding

$$\sigma_E = 0.423\mathrm{e}^{-0.105SNR}. \quad \text{(C2)}$$

Q is estimated from a division of the maximum envelope of the down- and upgoing wave. Small-strain damping d and Q are related

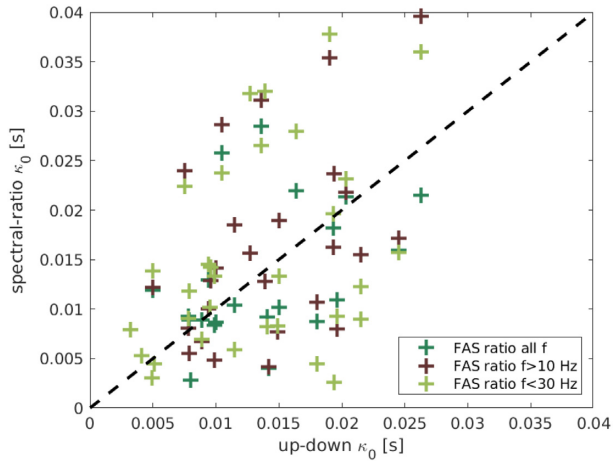

Figure D1. Cross-plot of κ_0 values obtained with the up-down method and the spectral-ratio method, using different frequency bands. All frequencies means the entire frequency band with SNR > 3 dB. If the outcomes of the different methods were equal, they would lie on the dashed black line.

as $d = 1/2Q$. Hence, by reworking eq. (18) the following expression is obtained for outer bounds of a confidence interval in terms of damping

$$\underline{d_i}^{\pm} = \frac{\ln\left(\frac{E^+}{E^-}\left(1 \pm \sqrt{\sigma^2_{E+} + \sigma^2_{E-}}\right)\right)}{-\pi\,\tau_i(F^- + F^+)}, \tag{C3}$$

where $\underline{d_i}$ is the damping from the free surface to depth level z_i and τ_i is the one-way traveltime over the same trajectory. σ_{E+} and σ_{E-} are the amplitude standard deviations of the down- and upgoing waves, respectively. The 68 per cent confidence interval is described by $[\underline{d_i}^- \; \underline{d_i}^+]$.

Instead of using eqs (C2) and (C3), also a damping distribution can be computed by running 2000 noise realizations and computing the damping for each realization using eq. (18). Fig. C1(d) shows the resulting damping probability density function (PDF) for a typical Groningen case. We use the mean velocity over the top 200 m in Groningen, which is 352 m s^{-1}. This corresponds to a mean one-way traveltime of τ_{200}=0.568 s. Furthermore we assume an instantaneous frequency of the up- and downgoing waves of 9.3 and 8.7 Hz, respectively. With a 40 per cent reduction in amplitude of the downgoing wave with respect to the upgoing wave, these parameters correspond to a damping of 0.8 per cent. The distribution shown in Fig. C1(d) for SNR =10 dB has a mean damping of 0.80 per cent and a standard deviation of 0.34 per cent. When instead the derived functions are used (eqs C2 and C3) a 68 per cent confidence interval is obtained between [0.5 1.16] per cent, which is in line with 0.80 ± 0.34. With both approaches a positive skew is obtained of about 0.15. In the main text, the damping confidence area is obtained by (1) determining the SNR of the up- and downgoing waves and (2) propagating to a damping confidence zone using eqs (C2) and (C3).

Fig. C2 shows the 68 per cent confidence regions for estimating damping over an average Groningen top 50 and 200 m soil profile, for a range of SNRs. For the 200 m profile, the same traveltime, (undisturbed) amplitude ratio and instantaneous frequency is used as above. For the 50 m scenario, a mean one-way traveltime of 0.230 s is taken and an amplitude ratio of 0.75. In Fig. C2, it can be seen again that the estimated damping distribution has a positive skew. Also, it can be seen that damping over the top 200 m can be estimated more accurately than over the top 50 m. The reason is that the size of the confidence region scales inversely with traveltime (eq. 21). Naturally, the confidence regions become smaller for larger SNRs.

APPENDIX D: SPECTRAL-RATIO APPROACH

An alternative approach to estimate damping in the upper 200 m can be made through analysis of the ratio of Fourier amplitude spectra (FAS) of downhole and surface records. This method is consistent with the commonly used model of Anderson & Hough (1984), which seeks to define the shape of high-frequency FAS of earthquake acceleration time-histories for use in engineering applications. In the Anderson & Hough (1984) model, the decay of high frequency spectral amplitudes is defined by a single parameter, κ, and assumed to be related to damping along both path and site, κ_r and κ_0, respectively (with $\kappa = \kappa_r + \kappa_0$). The log-linear decay of high-frequency acceleration FAS, $A(f)$, is given by

$$A(f) \propto \mathrm{e}^{-\pi f(\kappa_r + \kappa_0)}, \quad f_1 < f < f_2. \tag{D1}$$

The bounding frequencies f_1 and f_2 are defined by the earthquake's source corner frequency, f_c, with $f_1 \gg f_c$, and an upper bound, f_2, imposed by the noise floor.

Cabas *et al.* (2017) investigated the differences between κ determined at surface and downhole sensors of KiK-net in Japan and found a good correlation between their difference and the sites' V_{S30} values, although they observed only weak correlation with laboratory-derived damping measurements. A caveat of the Anderson & Hough (1984) model is that it does not consider complexities in source or path-effects, nor does it account for high-frequency site amplification or resonance effects. This can significantly bias individual estimates of κ (Edwards *et al.* 2015).

Similarly as in Section 2, the elastic effects are included here by writing the FAS as

$$A(f) = F(z, z_R)\mathrm{e}^{-\pi f(\kappa_r + \kappa_0)}, \tag{D2}$$

where $F(z, z_R)$ contains the elastic propagation amplitude effects between depth level z and a reference depth level z_R. Instead of individually determining in the borehole and surface sensors, we compute, and directly model the ratio of recorded FAS, which is vastly simplified as the source and path effects (to the base of the borehole) are identical for both surface FAS (A_0) and borehole (A). As a result, we no longer have a lower limit defined by the source, the path effects are removed, rather than simplistically modelled by κ_r, and we are left only with a ratio that describes the effect of the site:

$$\frac{A_0(f)}{A(f)} = F(z_0, z)\frac{\mathrm{e}^{-\pi f\kappa_0}}{1 + \mathrm{e}^{-2\pi f\kappa_0}}, \quad f < f_2. \tag{D3}$$

Note the above equation has large similarity with $|T^-(z_0, z, \omega)|$ (eq. 5). The elastic amplitude effects $F(z_0, z, f)$ are modelled with eq. (A3) as the ratio of the outcrop-motion and within-motion TFs at the surface and at depth, respectively. They are calculated using the 1D-SH TFs, leaving only the exponential decay as an unknown. The fitted term, κ_0, describes the attenuation due to the layers between the borehole sensor and the surface.

For moderate damping ($\kappa_0 \geq 0.02$ s) and high frequency (10 Hz $\leq f <$ 50 Hz) the spectral ratio decays within 92 per cent of κ_0. For such cases, a simple linear fit to the log of the spectral ratio (eq. D3) may therefore suffice, to obtain a slope of approximately $-\pi\kappa_0$. However, for lower frequency FAS, or for low damping,

we must consider the denominator in eq. (D3). Alternatively, the deconvolution result can be time windowed prior to computing the FAS. By time windowing the main pulse, only the numerator term in eq. (D3) remains. We fit eq. (D3) using a non-linear guided decent (Powell's) method with constraint that $\kappa_0 > 0$. The data are obtained from stations of the G-network for 6 recorded local events ($M_L >$ 2.5) (Ntinalexis *et al.* 2019). All data are assessed for SNR and only used in the bandwidth SNR > 3 for both surface and downhole records. This yields 32 sites with a spectral-ratio based estimate. The resulting κ_0 values are compared with the up-down method based estimates (Section 3.2) in Fig. D1. We repeat the spectral-ratio method for frequency bands restricted with $f > 10$ Hz (at 27 sites) and with $f < 30$ Hz (at 32 sites).

Averaged over the sites, the results of the spectral-ratio method are similar to the ones of the up-down method. However, for individual sites, large differences exist in the estimated κ_0. The spectral-ratio method results are very sensitive to the exact frequency-band used, without a clear pattern as function of frequency band. Also numerical tests (not included here) show that the spectral-ratio results are unstable when noise is added.