

# Quality parameters for passive image interferometry tested at the Groningen network

E.B. Fokker<sup>1,2</sup> and E.N. Ruigrok<sup>1,2</sup>

<sup>1</sup>Department of Earth Sciences, Utrecht University, Princetonlaan 8a, 3584 CB, Utrecht, the Netherlands. E-mail: [e.b.fokker@uu.nl](mailto:e.b.fokker@uu.nl)

<sup>2</sup>R&D Seismology and Acoustics, Royal Netherlands Meteorological Institute, Utrechtseweg 297, 3731 GA, De Bilt, the Netherlands

Accepted 2019 May 15. Received 2019 April 8; in original form 2018 December 3

## SUMMARY

Passive image interferometry is widely used to retrieve velocity variations as function of time. The cross-coherence (spectrally normalized cross-correlation) of ambient noise recorded at two receivers is an estimate of the Green's function that due to velocity changes will be stretched or compressed in time. The relative velocity change ( $\Delta v/v$ ) is determined by the time stretching ( $-\Delta t/t$ ) that yields highest correlation between the reference cross-coherence and stretched lapse cross-coherence. The estimation of  $\Delta v/v$  could be used, for example, to warn for hazardous situations developing in a volcano setting or due to degradation of a civil engineering structure. Before a warning would be issued, however, one would like to have a handle on the quality of the medium-change estimate.

In the Groningen area in the Netherlands a well-sampled network of seismometers exists. This allows direct assessment of the quality of the  $\Delta v/v$  estimation by many different receiver paths sampling the same medium. We use this quality assessment of  $\Delta v/v$  to test other possible quality parameters that could also be used in settings with only a sparsely sampled seismic network. The quality of the  $\Delta v/v$  determination appears to be well described by the correlation coefficient which is used to determine the velocity change. Consequently, in order to measure medium changes with a high certainty, it is important to have a high consistency between the lapse and the reference cross-coherence. A quality estimate based on this correlation coefficient can also be applied when there is only one receiver pair available. If the quality is insufficient, it can be improved at the cost of temporal resolution.

This study also investigates possible causes of the measured medium changes: variations in temperature, soil moisture, air pressure, water table, gas production rate and subsidence. We find a weak anticorrelation with temperature, and a weak correlation with the gas production rate and subsidence. The observed medium change is likely a complicated combination of different processes taking place.

**Key words:** Coda waves; Seismic interferometry; Seismic noise.

## 1 INTRODUCTION

The possibility to detect seismic velocity changes is key to monitor natural phenomena with possibly hazardous implications. Before a volcanic eruption, when temperatures inside a volcano increase, and before an earthquake, when stresses increase, the seismic velocity changes. Monitoring these velocities may provide crucial information for predictions or better understanding of volcanic eruptions and earthquakes (e.g. Brenguier *et al.* 2008; Wegler *et al.* 2009; Obermann *et al.* 2014), just as the understanding of other tectonic deformations inducing stress changes (e.g. Brenguier *et al.* 2014; Takano *et al.* 2017). Also industrial activities in the subsurface lead to velocity perturbations. For example, the pressure front due to wastewater injection may be tracked by measuring velocity changes

or the deformation development at an enhanced geothermal site (Hillers *et al.* 2015). Monitoring these velocity changes benefits both the economics and safety of these engineering projects.

Seismic velocities can be monitored actively or passively. Active sources such as explosives or seismic vibrators, however, are hard to use continuously. Furthermore, they are expensive and may be dangerous to use. Passive techniques provide viable alternatives. Passive image interferometry (PII; Sens-Schönfelder & Wegler 2006) is such a passive technique. It uses ambient noise to measure relative changes in the seismic velocity ( $\Delta v/v$ ). This procedure employs two major steps. First, the cross-coherence of ambient noise recorded at two receivers is calculated to estimate the Green's function (Wapenaar *et al.* 2010). The Green's function is the response at one receiver, caused by an impulse at the other receiver, generating waves

for all possible paths to the first receiver. This Green's function estimate is constructed from one or a few days of data. We refer to it as the lapse Green's function. Second, the stretching or compressing of this lapse Green's function in time can be measured with respect to a reference Green's function and used to calculate  $\Delta v/v$  (Lobkis & Weaver 2003).

As outlined by Weaver *et al.* (2011) the estimated  $\Delta v/v$  contains a physical part, actual changes in the medium, but also there is a spurious contribution, for example, caused by changes in noise field. Both the estimate of  $\Delta v/v$  and its uncertainty vary with time. In order to use velocity monitoring as hazard or engineering tool, it is crucial to have a good estimate of the quality of the determination. This quality can directly be assessed when many station pairs sample the same medium with different azimuths. However, it is not obvious which quality parameter could be used when only one or a few station pairs are available.

To find a suitable generally applicable quality parameter, we apply PII to 1 yr of seismic noise recorded at the Groningen network (Dost *et al.* 2017). The Groningen network has a sufficient azimuthal distribution of station pairs to estimate the quality of the  $\Delta v/v$  determination directly, using the consistency of independent measurements sampling the same medium from different azimuths. This quality is compared with a suit of candidate quality parameters that could also be applied in networks with less optimal azimuthal coverage. The first candidate quality parameter is constructed from the correlation coefficients that are found after optimally stretching a lapse Green's function with the reference Green's function. The second is based on a quantification of spurious contributions and describes how well the lapse Green's function is estimated. The third is based on how isotropically the medium is illuminated from ambient seismic noise.

In the following we test the different quality parameters and we investigate possible causes of the medium perturbations as observed over Groningen. The next section outlines the setting of our field test. The section thereafter describes the implementation of PII and how the different quality factors are defined. This is followed by a description of the results of applying PII and the various quality parameters on actual data. Possible underlying causes of the observed medium changes are investigated in Section 5. We close the paper with a discussion and conclusions.

## 2 SETTING

Groningen is a province in the northeast of the Netherlands. It hosts the largest gas field in Western Europe. Production started in the early sixties after an accidental discovery in 1959 (de Jager & Visser 2017). In 1991 the first earthquake was recorded which could be ascribed to pressure depletion of the reservoir. For better monitoring the induced seismicity, a regional seismic network was constructed in 1995 (Dost *et al.* 2017). Since 2015 a new network is operational dedicated to monitoring the Groningen field: the G-network. This network contains of 70 stations placed in the near surface. At each station there is an accelerometer at the Earth's surface and there are geophones at 50, 100, 150 and 200 m depth (Hofman *et al.* 2017).

All data from the G-network is openly available via KNMI (1993). Throughout this work we use the continuous recordings at depth of 50 m. These instruments well record surface waves from various noise sources and are less affected by local anthropogenic activity than the surface accelerometers. The sensors are 4.5 Hz three-component geophones which noise recordings well capture the Peterson New Noise Model down till about 0.3 Hz (Spica *et al.* 2018).

Below this frequency the instrumental noise starts to dominate over seismic noise.

From the 70 stations, we select a subset of stations in the middle of the Groningen field (Fig. 1a). This subset contains of 14 stations that sample an area of roughly  $14 \times 14$  km<sup>2</sup>. An area has been chosen that may be considered 1-D for the upper 1700 m (Duin *et al.* 2006); It has unconsolidated Cenozoic sediments for the first 800 m, overlaying a thick layer of chalk (900 m). Below there is a thin blanket of Jurassic and Triassic sediments in the west (100 m) overlying a thick layer of Permian evaporites (1 km). In the east the evaporites are only half the thickness, but there are about 600 m of Jurassic and Triassic sediments. Below the evaporites, at about 3 km depth, the gas reservoir resides in a Permian sandstone formation. 91 unique station pairs sample the area with a rich distribution in azimuths (Fig. 1b).

Below 1 Hz, the noise field is dominated by the microseism (Longuet-Higgins 1950). The double-frequency microseism peak falls out of the sensitive range of the instruments. However, Groningen is close to the shallow the North Sea, which yields ocean-wave solid-earth interaction that can be recorded with frequencies beyond 1 Hz. The dominant noise direction is northwest, which corresponds to the location of the North Sea and North Atlantic with respect to the array (Kimman *et al.* 2012; Spica *et al.* 2018). From 1 Hz onwards anthropogenic sources start to dominate, like cars, trains and industrial activity.

## 3 METHODS AND MATERIAL

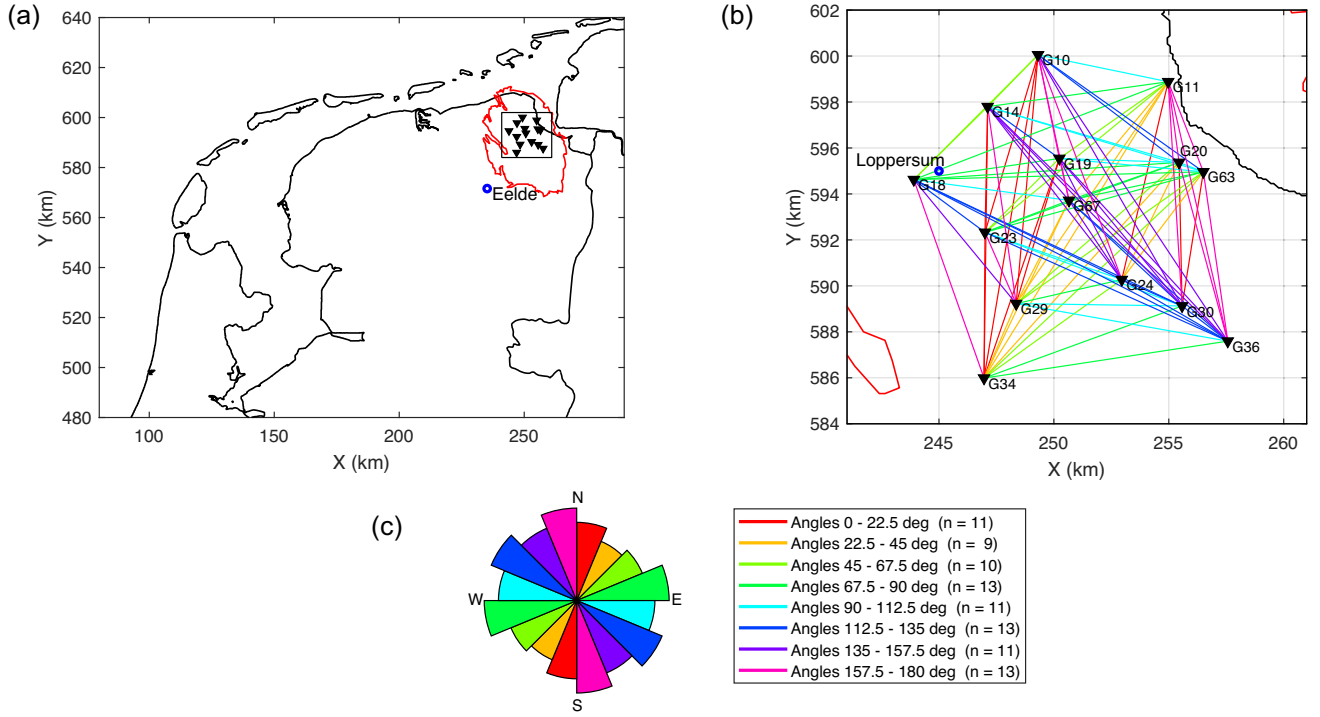
### 3.1 Passive image interferometry

Wapenaar *et al.* (2010) derived that the cross-coherence  $H$  of the displacements at two receivers provides a fine estimate of the Green's function if the illumination is good. The cross-coherence is a cross-correlation of the waveforms with additionally a spectral normalization; the cross-correlation is normalized with the noise amplitude spectra at the two receivers. This normalization works well to whiten the spectrally unbalanced noise cross-correlation. In this cross-coherence  $H(x_A, x_B, t)$  positive times represent the causal Green's function (i.e. the response at  $x_B$  due to a pulse at  $x_A$ ) and negative times represent the anticausal Green's function (i.e. the time inverted response at  $x_A$  due to a pulse at  $x_B$ ).

Typically a reference Green's function ( $H_{\text{ref}}$ ) is estimated by stacking cross-coherences over 1 yr of data. Also for time-periods of a single day, cross-coherences are stacked to obtain an estimate of the state of the medium on that exact day: the lapse Green's function  $H_{\text{lapse}}$ . Relative changes in propagation velocity would stretch or compress the lapse Green's function with respect to the reference Green's function. The time lapse seismic imaging technique (Lobkis & Weaver 2003; Weemstra *et al.* 2016) determines a correlation coefficient  $CC$  between  $H_{\text{lapse}}$  stretched in time with factor  $(1 - \epsilon)$ , and  $H_{\text{ref}}$ :

$$CC(\epsilon) = \frac{\int_{t_1}^{t_2} H_{\text{lapse}}[t(1 - \epsilon)] H_{\text{ref}}[t] dt}{\sqrt{\int_{t_1}^{t_2} (H_{\text{lapse}})^2 [t(1 - \epsilon)] dt} \sqrt{\int_{t_1}^{t_2} (H_{\text{ref}})^2 [t] dt}}. \quad (1)$$

The stretching factor  $\epsilon$  for which the correlation coefficient is highest serves as an estimate of the relative velocity change  $\Delta v/v = \epsilon$ .  $t_1$  and  $t_2$  control the time window that is taken from the cross-coherence to estimate  $CC$  and  $\epsilon$ .  $\Delta v/v$  is determined more accurately when this time window is chosen such that in the cross-coherence



**Figure 1.** Maps of locations of seismic stations in the northeast of the Netherlands that are used in this study. The coordinates are in the Cartesian Dutch Triangular System (in Dutch: Rijksdriehoekskoördinaten).  $X$ : Easting (km),  $Y$ : Northing (km). (a) Overview map with circumference of the Netherlands northern coastline (black), circumference of the Groningen gas field (red), the location of the seismic stations (black triangles), the location of the weather station in Eelde (blue point) and a small box which indicates the location of (b). (b) The subset of G-network stations (Dost *et al.* 2017) that are used in this study and the location of Loppersum where the InSAR subsidence data originated from. (c) Colour coding of the connecting lines in (b) indicating the angle with the north in accordance with the colour coding used in Fig. 5.

the part of direct waves is disregarded (Poupinet *et al.* 1984; Sens-Schönfelder & Wegler 2006; Weemstra *et al.* 2016), since these direct waves are most sensitive to the (daily varying) illumination of the medium.

The time window including the direct wave is chosen as  $t_1 = -t_{\max}$  and  $t_2 = t_{\max}$ , where  $t_{\max}$  is the maximum lag time for which the cross-coherence is calculated. For the time window excluding the direct wave integrals  $\int_{t_1}^{t_2} \dots dt$  are replaced by  $\int_{-t_{\max}}^{-t_{\text{direct}}} \dots dt + \int_{t_{\text{direct}}}^{t_{\max}} \dots dt$ , where  $t_{\text{direct}}$  is the traveltime of a wave travelling a direct path between the receivers, depending on the phase velocity and the distance between the receivers. Traveltime  $t_{\text{direct}}$  is calculated with a margin to ensure that all energy of the direct waves is disregarded.

The average  $\Delta v/v$  over multiple receiver pairs is calculated as an azimuthally weighted average, together with an azimuthally weighted standard deviation. The weights are proportional to the range of azimuths that a receiver pair represents (i.e. the smallest angle between the two receiver pairs divided by two).

For the application of PII one could use the freely available Python package created by Lecocq *et al.* (2014) which is suitable to monitor velocity changes from ambient noise.

### 3.2 Quality assessment parameters

Under the assumption of velocity changes being isotropic, the different receiver pairs sampling the same medium from different azimuths should detect the same  $\Delta v/v$ . Detecting differences in the velocity changes would then indicate that the velocity changes were not determined accurately. Thus, the standard deviation of the velocity change over the different receiver pairs sampling the medium

from different azimuths is used as a base to derive a consistency quality parameter  $Q_{\text{PII}}(t)$ :

$$Q_{\text{PII}}(t) = 1 - \frac{\sigma_{\Delta v/v}^{\text{lapse}}(t)}{\sigma_{\Delta v/v}^{\text{year}}}, \quad (2)$$

where the daily varying standard deviation  $\sigma_{\Delta v/v}^{\text{lapse}}(t)$  quantifies the differences between the measurements, and the standard deviation of the yearly variations  $\sigma_{\Delta v/v}^{\text{year}}$  describes the yearly perturbation of the medium and acts as a normalization parameter. These standard deviations are defined as

$$\sigma_{\Delta v/v}^{\text{lapse}}(t) = \sqrt{\frac{\sum_{i=1}^n (\Delta v/v_i(t) - \overline{\Delta v/v}(t))^2}{n-1}},$$

$$\sigma_{\Delta v/v}^{\text{year}} = \sqrt{\frac{\sum_{t=1}^{365} (\overline{\Delta v/v}(t) - \overline{\overline{\Delta v/v}})^2}{365-1}},$$

where  $\Delta v/v_i(t)$  is the velocity change at time  $t$  obtained with receiver combination  $i$ ,  $\overline{\Delta v/v}(t)$  is the velocity change averaged over all  $n$  receiver combinations and  $\overline{\overline{\Delta v/v}}$  is  $\overline{\Delta v/v}(t)$  averaged over a whole year.

The quality factor approaches 1 when the measurements of all receivers pairs give very similar  $\Delta v/v_i(t)$  ( $\sigma_{\Delta v/v}^{\text{lapse}}(t) \ll \sigma_{\Delta v/v}^{\text{year}}$ ). A quality factor much smaller than 1 will be reached when there is a large inconsistency within the network of the estimated daily  $\Delta v/v_i(t)$ . Therefore, consistency quality factor  $Q_{\text{PII}}(t)$  is a good measure for the reliability of the daily estimations of velocity change  $\Delta v/v$ .

The second quality factor  $Q_{CCF}(t)$  is constructed from the highest correlation coefficient calculated using eq. (1) and averaging over all receiver combinations. The coefficient quality is given by

$$Q_{CCF}(t) = \sum_{i=1}^n \frac{CC_i(t)}{n}. \quad (3)$$

Here  $CC_i(t)$  represents the largest correlation coefficient  $CC(\epsilon)$  (eq. 1) used to calculate  $\Delta v/v(t)$  between receiver combination  $i$ , and  $n$  represents the amount of receiver combinations. This quality factor approaches 1 when the optimally stretched lapse cross-coherence is a perfect match of the reference cross-coherence. The quality factor will be lower than one when the lapse cross-coherence can not well be written as a stretched copy of the reference cross-coherence. Strong correlation between consistency quality  $Q_{PII}$  and coefficient quality  $Q_{CCF}$  would indicate that a high similarity between the lapse and reference Green's function is needed to estimate  $\Delta v/v$  well.

The third quality factor uses the power distribution of noise sources around the Groningen setting. We use the method of Ruigrok *et al.* (2017) to calculate the power of waves coming from different directions. Their method of cross-correlation beamforming requires as input the cross-coherence that we also used in the previous subsection. The beamforming results, however, are aliased due to the relatively large distances between the receivers. We limit the effect of the aliasing by only focussing on the beampower in the slowness range in which surface waves exist. To determine sensible slowness bounds we first determine the phase velocities of these surface waves.

For this purpose we apply multichannel analysis of surface waves (MASW) as introduced by Park *et al.* (1998). MASW uses the move-out of the surface waves over an array of receivers to obtain average phase velocities as function of frequency. As a first step we transform the estimated reference Green's functions to the wavenumber frequency domain. The amplitudes squared represent the power as function of frequency and wavenumber. The wavenumber axis is converted to a phase velocity axis using  $v_{ph}(\omega) = \omega/k(\omega)$ , where  $\omega$  is the angular frequency and  $k$  is the wavenumber. The connections between maxima of the power as function of phase velocity and frequency represent the dispersion curves of the surface waves. We smooth the dispersion curves and plot them on top of the power plot.

We use the slowness bounds of the dispersion curve with the highest power to improve the beamforming results by stacking the beamforming results of tight frequency and slowness ranges. With this stacked power  $P(\theta)$  as function of direction and its maximum  $P_{max}$ , we define the illumination quality  $Q_{ILL}(t)$  as the ratio between the average power of all directions and the maximum power:

$$Q_{ILL}(t) = \frac{1}{360 \cdot P_{max}(t)} \int_0^{360} P(\theta, t) d\theta. \quad (4)$$

This quality factor will be 1 when all directions have the same power and approaches zero when all power is concentrated at one direction. Strong correlation between consistency quality  $Q_{PII}$  and illumination quality  $Q_{ILL}$  would indicate that an azimuthally balanced illumination is required to determine  $\Delta v/v$  accurately.

In a fourth quality assessment we compare, within the cross-coherences  $H(t)$ , the power before the first arrival with the power after the first arrival. First, we use the highest velocity calculated with MASW, together with the distance between two receivers, to calculate the arrival time of the first surface wave, and we use the lowest velocity to calculate the arrival time of the last surface wave arriving through a direct path between the receivers. To allow the

complete waveform to come in we take for  $\tau_1$  the first arrival time, minus a small time of 2 s, and for  $\tau_2$  the last arrival time plus an additional time of 10 s. Then we compare the average power before (causal and acausal) the first arrival  $P_0 = \frac{1}{2\tau_1} \int_{-\tau_1}^{\tau_1} H(t)^2 dt$  with the average power of the surface waves  $P_1 = \frac{1}{2(\tau_2 - \tau_1)} \int_{\tau_1}^{\tau_2} H(t)^2 dt + \frac{1}{2(\tau_2 - \tau_1)} \int_{\tau_1}^{\tau_2} H(t)^2 dt$  for every receiver combination for which  $\tau_1 > 0$  and calculate the average ratio  $\langle P_0/P_1 \rangle$ . Now we define the cross-coherence quality as

$$Q_H(t) = 1 - \left\langle \frac{P_0(t)}{P_1(t)} \right\rangle. \quad (5)$$

The quality factor approaches 1 when the power of the response before the first arrival is much smaller than the power of the response of the actual waves, and will be negative when the power before the first arrival is larger than the power of the ballistic surface waves. Strong correlation between consistency quality  $Q_{PII}$  and cross-coherence quality  $Q_H$  would indicate that  $\Delta v/v$  can only be determined accurately if the Green's function is estimated well.

## 4 RESULTS

### 4.1 Green's function estimation and dispersion analysis

The estimation of the velocity changes and the quality there-off is done using lapse and reference Green's functions. Estimations of these Green's functions are shown in this section. 1 yr of continuous data is used in the analysis, starting on 2016 March 30. The reference Green's functions are further used to estimate geometric dispersion relations for surface waves in Groningen. These are needed for two of the quality parameters that are computed in the next section.

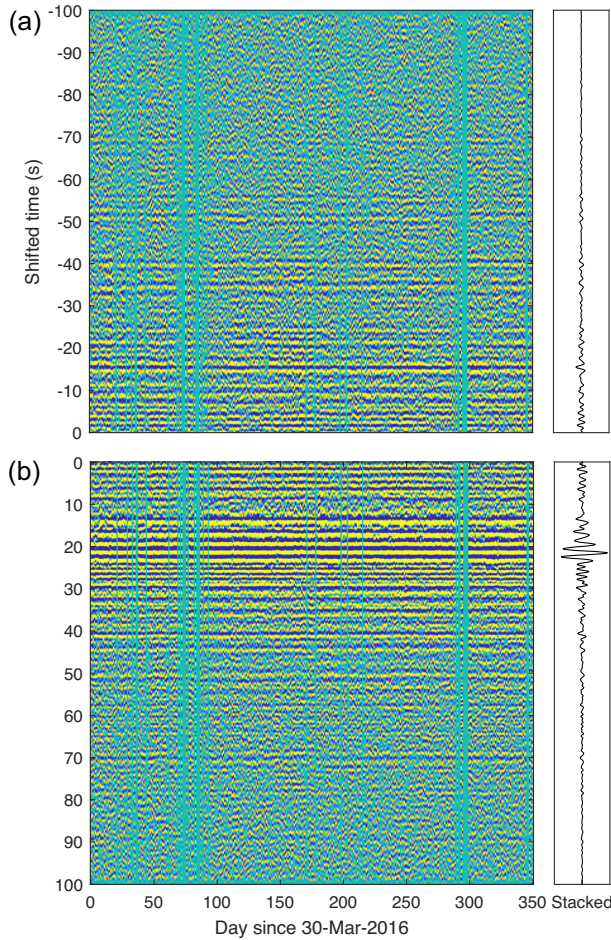
Fig. 2 shows an example of lapse cross-coherences, together with the reference cross-coherence. For every day's calculation we stacked the cross-coherences calculated from 50 percent overlapping time windows of 20 min duration, where the first time window ranges from 0:00 to 0:20 UTC, the second from 0:10 to 0:30 UTC, etc.

Fig. 3 shows the cross-coherences of all receiver combinations shown in Fig. 1, stacked over 1 yr. These reference cross-coherences are plotted as function of distance between the receivers for the radial (a), the transverse (b) and the vertical (c) component. The arrival times increase with distance, as expected. The transverse component shows arrivals with a larger moveout than on the radial component. This would indicate that Love waves are slower over the unconsolidated Neogene sediments than Rayleigh waves. Unlike the radial component, the vertical component shows clearly two arrivals, which implies two different modes of the Rayleigh wave.

Using the reference cross-coherences (Fig. 3), the phase velocities are calculated for the shallow subsurface of Groningen. Fig. 4 shows the power spectrum of the retrieved surface waves as function of frequency and phase velocity for Rayleigh waves (a) and Love waves (b). The power of the Rayleigh wave is constructed by summing the powers of the radial and vertical components. Several modes are found both for the Rayleigh and the Love wave, but different modes dominate for different frequency ranges. Only modes that exist consistently over the patch of receivers (Fig. 1) can well be picked. The velocities we find for the Rayleigh waves are comparable to the velocities found by Spica *et al.* (2018), but the Love-wave velocities we find are much lower.

From Fig. 4 we find that most energy propagates in the first higher mode with phase velocities between 850 and 1000 m/s, corresponding to frequencies of 0.65 and 0.45 Hz. Assuming that the





**Figure 2.** Lapse cross-coherences (colourplot) for 1 yr of seismic data of the radial components of receivers G11 and G29 (Fig. 1) in the frequency range [0.30 1.0] Hz, together with the reference cross-coherence (i.e. all lapse cross-coherences of 1 yr stacked). The cross-coherences are splitted in an anticausal (a) and a causal (b) part.

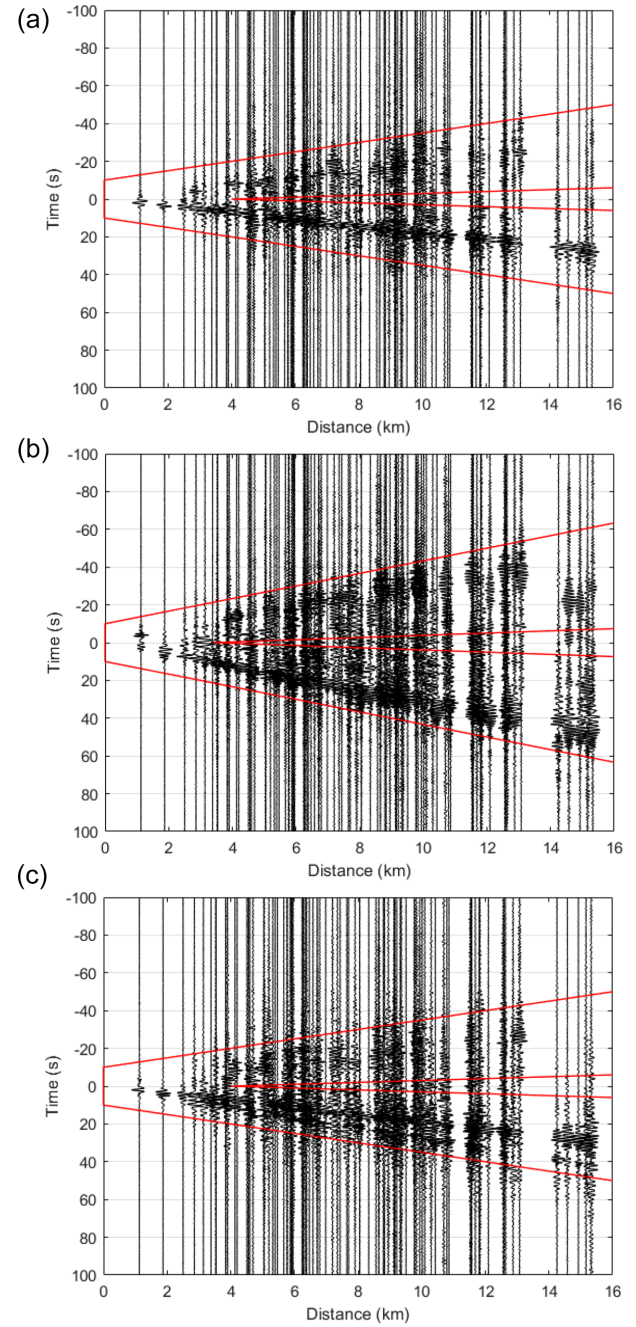
maximum sensitivity of a wave is at one-third of the wavelength, the depths of maximum sensitivity range from 440 to 740 m. Therefore, most energy propagates within the 800 m deep package of Cenozoic sediments.

#### 4.2 Estimating medium changes and the quality there-off

In this section we show the results of the application of PII and the different quality parameters as introduced in Section 3.

Fig. 5 shows  $\Delta v/v$  obtained using the radial, transverse and vertical component, when in the stretching technique (eq. 1) the full cross-coherences are used (a, c, e), or the late arrivals only (b, d, f). Each curve represents an average of  $\Delta v/v$  of receiver pairs binned with orientation bins of  $22.5^\circ$ . In each computation, both the anticausal and causal parts of the cross-coherences are used. When the direct waves are omitted (panels on the right-hand side of Fig. 5) larger velocity variations are estimated. In this case the estimated velocity variations are more consistent for receiver pairs with different azimuths. The calculation parameters used to obtain Fig. 5 are given in Table 1.

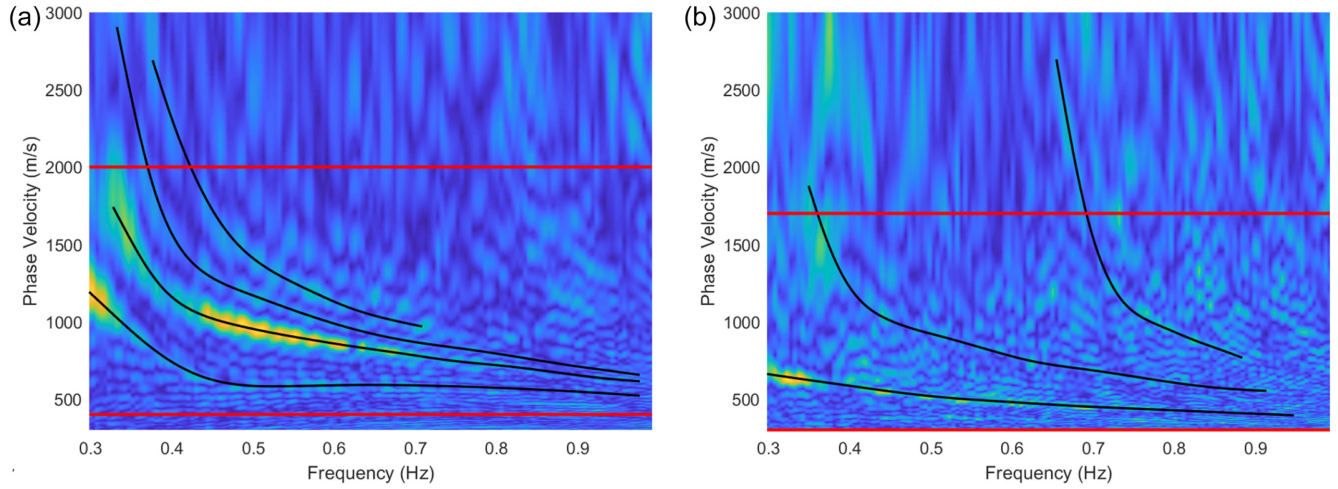
The averages of  $\Delta v/v$  over all station-pair azimuths are shown in Fig. 6. These are azimuthally averaged means that are calculated



**Figure 3.** Reference Green's functions plotted as function of the distance between receiver pairs in the frequency range [0.30 1.0] Hz for the radial (a), transverse (b) and vertical (c) components. For 1 yr of continuous data, cross-coherences were calculated and stacked. The red lines represent the times  $\tau_1$  and  $\tau_2$  between which all arrivals are expected of ballistic surface waves travelling a direct path between the receivers.

with the late arrivals only. Also the azimuthally weighted standard deviations  $\sigma_{\Delta v/v}^{\text{lapse}}(t)$  are indicated.

We now have arrived at a point where we can assess whether the quality of the  $\Delta v/v$  determination is correlated with the other quality factors. From the data underlying Fig. 6 we compute consistency quality  $Q_{\text{PII}}(t)$  in accordance with eq. (2) and coefficient quality  $Q_{\text{CCF}}(t)$  in accordance with eq. (3). The comparisons between these quality factors are shown in Fig. 7 for the three different components, using the procedure excluding the arrivals of the direct



**Figure 4.** Dispersion curves of the phase velocity obtained by applying MASW (Park *et al.* 1998) to the cross-coherences shown in Fig. 3 for the Rayleigh waves (a) and for the Love waves (b). The power in (a) is obtained by stacking the power of the radial and vertical component, while (b) just represents the power of the transverse component. For both types of surface waves multiple modes are found as indicated by the black lines. The red line indicates the velocity bounds between which most of the energy propagates.

surface waves. A first inspection reveals a strong correlation between  $Q_{\text{PII}}(t)$  and  $Q_{\text{CCF}}(t)$  for the radial and transverse components. For the vertical component the sizes of the qualities do not match. The correlation coefficients between the quality factors, however, show a strong correlation for all components. The correlation coefficients are shown in Table 2 for both the procedures excluding ( $Q_{\text{PII, ex}}$ ) and including ( $Q_{\text{PII, in}}$ ) the direct wave in the determination of  $\Delta v/v$ . The correlations are lower for the Qualities determined including the direct wave.

In order to compute illumination quality  $Q_{\text{ILL}}(t)$  (eq. 4) we first need to compute the azimuthally varying beampower  $P(\theta)$  and maximum beampower  $P_{\text{max}}$ . The beampower is calculated for frequencies for which the main phase velocity is determined well (Fig. 4). For every frequency bandwidth of 0.02 Hz, slowness bounds are calculated by inverting the phase velocity at the minimum and maximum frequency. Using these frequency bandwidths and corresponding slowness bounds, we calculate the power as function of time and angle with the north (clockwise). Fig. 8 shows this power of the waves recorded at the radial and vertical components, stacked over all frequency bandwidths between 0.34 and 0.73 Hz ([0.34 0.36] Hz, [0.35 0.37] Hz, etc.). At most frequencies the dominant source is in the direction of angles between  $280^\circ$  and  $350^\circ$ . This can be due to microseisms on the North Sea, which is located in the northwest, as seen from Groningen. A large change in illumination can be seen between days 13 and 98, which is a period with a large number of not working stations. Beampowers from this period, therefore, have not been used in further calculations.

From the data underlying Fig. 8 we derive the quality of the power distribution in accordance with eq. (4):  $Q_{\text{ILL}}$ . We compare this quality to the consistency quality of the determination of  $\Delta v/v$ :  $Q_{\text{PII}}$ . Fig. 9(a) shows the time-series of these quality parameters. The correlation coefficients between these qualities are shown in Table 2. These coefficients are below 0.1. Hence it can be concluded that a correlation is absent. Over the course of the deployment, different stations were inactive and therewith the array response function changed. This resulted in a varying amount of aliasing, effecting the  $Q_{\text{ILL}}$  parameter, without actual change in illumination. A large effect can especially be seen between days 13 and 98 in Fig. 8. For this reason, we did not include the values of the quality  $Q_{\text{ILL}}$  of these days in the calculations of the correlation between  $Q_{\text{PII}}$  and  $Q_{\text{ILL}}$ .

Due to an increase in spatial aliasing, there is some inconsistency between the frequency bandwidth used for computing the beam-power ([0.34 0.73] Hz) and in the determination of the velocity variations ([0.30 1.0] Hz). We have assumed that outside the [0.34 0.73] Hz range the energy field was comparable to the field inside the range. We do however not expect a large effect on the correlation between the two quality factors if there is some deviation from this distribution outside the range mentioned.

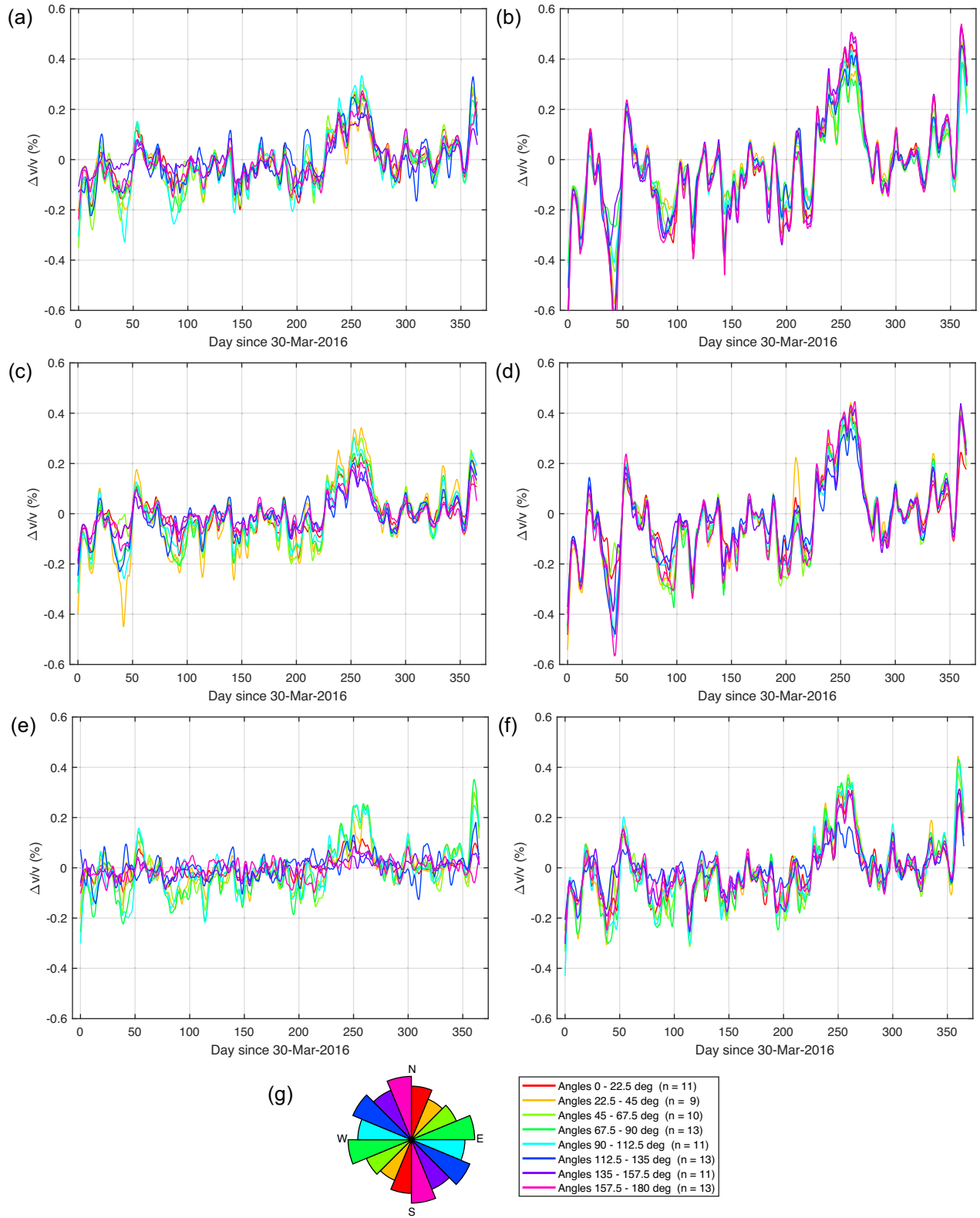
From the lapse cross-coherences we compute the quality parameter that is indicative of the quality of the Green's function estimate. We compute this cross-coherence quality  $Q_{\text{H}}$  in accordance with eq. (5) and compare this quality to consistency quality  $Q_{\text{PII}}$ . The correlation coefficients between these qualities are shown in Table 2. These coefficients are below 0.2. Hence it can be concluded that a correlation is absent.

## 5 INTERPRETATION

In this section we search for underlying causes of the observed medium changes (Fig. 6). In the frequency band we used ([0.30 1.0] Hz) for detecting the medium changes, the surface waves sample primarily the top 800 m of unconsolidated Cenozoic sediments. These sediments might be affected by gas production from, and compaction of, the underlying reservoir. From above, these sediments undergo influence from the atmospheric state. Here we investigate a possible correlation both with gas production and meteorological parameters.

Seismic velocities could be changed by meteorological conditions through changes in water saturation, pressure and temperature (e.g. Wang *et al.* 2017). Fig. 10 shows the comparison of  $\Delta v/v$  with meteorological parameters from a nearby weather station (Eelde; Fig. 1a, [lon,lat,alt]=[6.585°, 53.125°, 5.20 m]). A weak anticorrelation ( $R = -0.32$ ) can be seen between temperature and  $\Delta v/v$ . When temperature goes up, velocity goes down. Air pressure does not show any substantial correlation with  $\Delta v/v$ . The third meteorological parameter is soil moisture. Soil moisture is difficult to measure *in situ*. Instead a proxy is used: the measured precipitation minus a modelled evaporation. The soil moisture goes down during the growing season of plants, when much evaporation takes place





**Figure 5.** Velocity variations between the receiver combinations shown in Fig. 1, calculated within frequency bandwidth [0.30 1.0] Hz using the full cross-coherence (a,c,e) and the late arrivals only (b,d,f) for the radial (a,b), the transverse (c,d) and the vertical (e,f) components. The receiver combinations were separated by orientation calculating the average of all receiver combinations that have an orientation within a 22.5 degree bin as indicated in the Rose diagram and the legend (g). The velocity variations are smoothed over a smoothing time of 5 d using local regression with weighted linear least squares and a linear polynomial model. Detailed calculation parameters are given in Table 1. The colours from this figure match the colours of the lines between receivers in Fig. 1. The used reference cross-coherences are shown in Fig. 3.

**Table 1.** Detailed calculation parameters used to obtain velocity variations as shown in Figs 5 and 6.

Frequency bandwidth	[0.30 1.0] Hz
Time window for stretching (direct wave included)	[-100 100] s
Time window for stretching (direct wave excluded)	$[-100 -t_{\text{direct}}]$ $+ [t_{\text{direct}} 100]$ s
$t_{\text{direct}}$	$5\text{ s} + d/300\text{ m}$ $\text{s}^{-1}$
Sample frequency	10 Hz
Smoothing length of $\Delta v/v$ curves	5 d

through the leaves of vegetation. It starts to go up again around October, when plants loose there leaves (around day 190 in Fig. 10c). This yearly cycle of soil moisture depletion and replenishment cannot be traced in  $\Delta v/v$ .

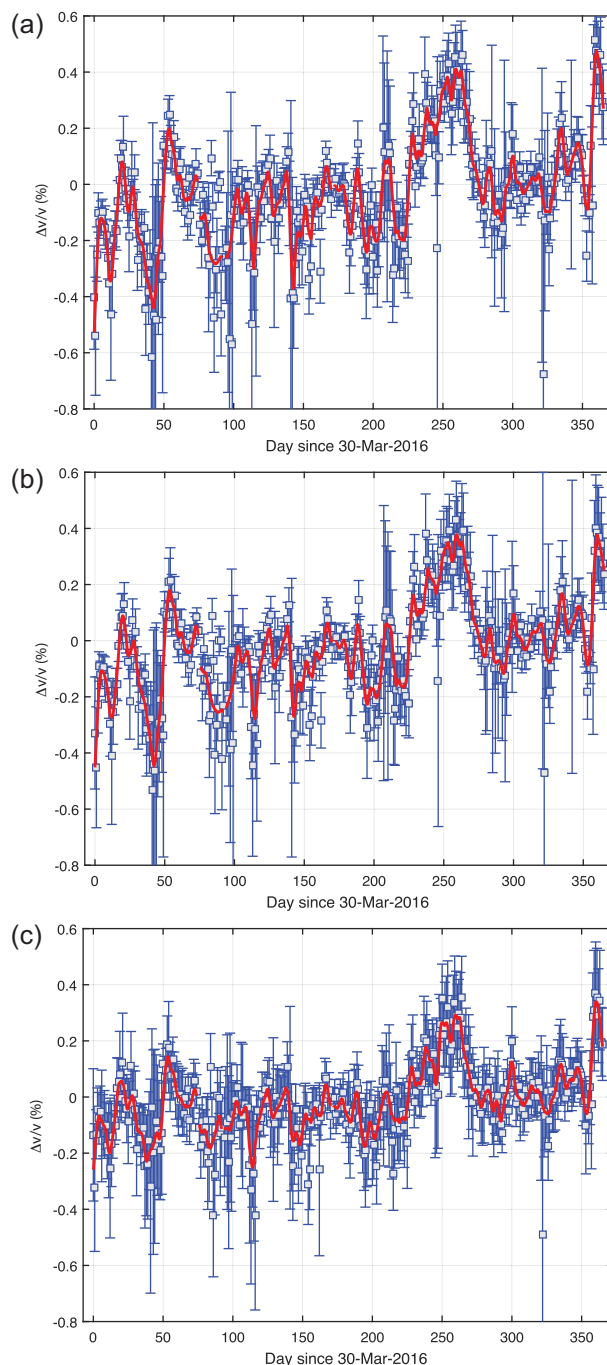
We also test a possible correlation of  $\Delta v/v$  with the ground water table. Daily water levels in the area are obtained from <http://www.dinoloket.nl>. Ground water variations vary from site to site. However, the yearly trend at all sites follows the same pattern as the soil moisture proxy. The water table fluctuations are less than 1 m over a year. As with soil moisture, no correlation can be seen between  $\Delta v/v$  and the ground water table.

Gas from the Groningen gas field is produced from 20 production clusters (<http://www.namplatform.nl>). Three of these production sites largely stopped production in 2014 (Muntendam-Bos *et al.* 2017). From the 17 remaining clusters, we select the ones that are well within the area that is sampled with PII (Fig. 1). We leave out production from clusters at the southern edge of this area. Fig. 11(a) shows both the total monthly production over these sites and  $\Delta v/v$ . Some correlation can be noted. Especially the production peak between days 200 and 300 coincides with a peak in  $\Delta v/v$  during the same period.

Over the year, overall there is an increase in  $\Delta v/v$  (Fig. 6). This could potentially be attributed to compaction. The production of gas from the reservoir leads to compaction of the reservoir rock, which is largely accommodated by subsidence of the Earth's surface. This subsidence is routinely monitored with spirit levelling, InSAR and GPS (van Eijs & van der Wal 2017). The current total subsidence is about 30 cm in the middle of the subsidence bowl, which is expected to grow to about 47 cm in 2080 (van Thienen-Visser & Fokker 2017). Potentially, the compaction at reservoir level also results in compaction of overlying Cenozoic sediments, especially near the middle of the subsidence bowl. The surface waves that we used in the application of PII sampled a significant portion of this 800 m package of unconsolidated sediments. Compaction would lead to a smaller distance of the grains and thus to an increase in seismic velocity. Fig. 11(b) shows the subsidence as derived from InSAR. The upgoing trend corresponds with the  $\Delta v/v$  signal, but the oscillations over the year show little correlation.

## 6 DISCUSSION

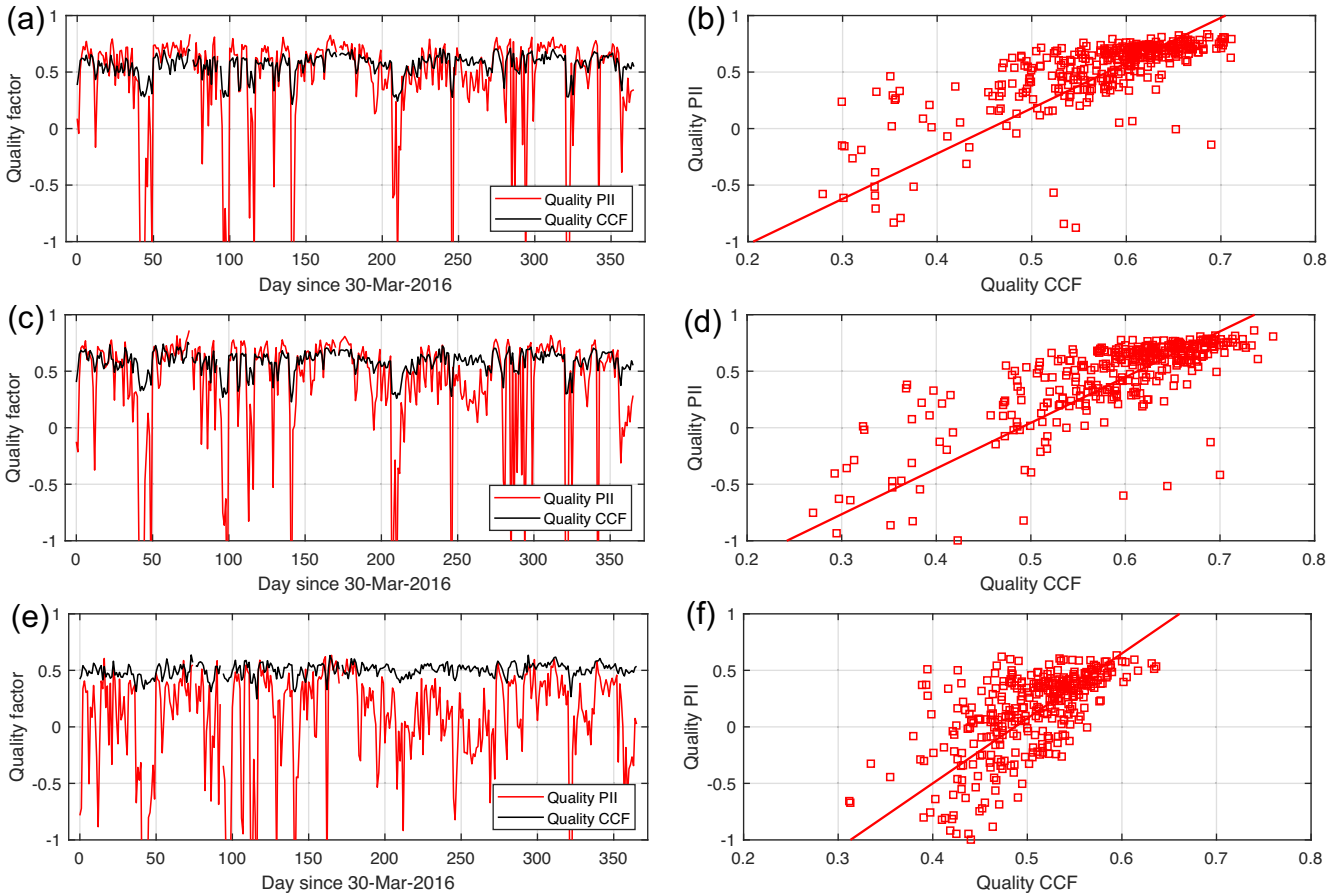
In Section 4 we extracted velocity variations from noise cross-coherences over radial, transverse and vertical components. For all components a consistent  $\Delta v/v$  was found. The  $\Delta v/v$  functions, however, were most pronounced on the horizontal components (Fig. 6). This was also reflected in overall higher consistency qualities  $Q_{\text{PII}}$  for the horizontal components (Fig. 7).



**Figure 6.** Azimuth-weighted average of the velocity variation plotted per day with its standard deviation (blue squares with error bars) and smoothed with 5 d (red line), estimated with late arrivals only (a) for the radial, (b) for the transverse and (c) for the vertical components. Detailed calculation parameters are given in Table 1.

From the three quality factors we compared with  $Q_{\text{PII}}$  in Section 4 we found that only coefficient quality  $Q_{\text{CCF}}$  is descriptive of the quality of the velocity determination. This notion corresponds with findings of Hadziioannou *et al.* (2009) who found in a laboratory experiment that retrieving a high quality Green's function is not important for estimating velocity changes. Instead, what is important is that the coda of the Green's function is retrieved in a stable





**Figure 7.** Consistency quality factor  $Q_{PII}$  in accordance with eq. (2) (red; direct wave excluded) and coefficient quality factor  $Q_{CCF}$  in accordance with eq. (3) (black) for the radial (a,b), the transverse (c,d) and the vertical (e,f) components. (a,c,e) Qualities as function of time. (b,d,f) Quality  $Q_{PII}$  plotted against quality  $Q_{CCF}$  with the best linear fit. The correlation coefficients between the quality factors are shown in Table 2.

**Table 2.** Correlation coefficients (at zero lag) between quality factor  $Q_{PII}(t)$  and quality factors  $Q_{CCF}(t)$ ,  $Q_{ILL}(t)$  and  $Q_H(t)$ . Quality factors  $Q_{PII,ex}$  and  $Q_{PII,in}$  (eq. 2) give the consistency qualities of the velocity measurement when the first arrivals are excluded and included, respectively. Quality factors  $Q_{CCF}$  (eq. 3),  $Q_{ILL}$  (eq. 4) and  $Q_H$  (eq. 5) are the correlation-coefficient-based, illumination-based and the cross-coherence-based qualities, respectively.

	$Q_{CCF}$	$Q_{ILL}$	$Q_H$
R $Q_{PII,ex}$	0.780	-0.081	-0.132
R $Q_{PII,in}$	0.558	-0.183	0.068
T $Q_{PII,ex}$	0.762		0.076
T $Q_{PII,in}$	0.632		0.153
Z $Q_{PII,ex}$	0.710	0.061	-0.134
Z $Q_{PII,in}$	0.442	0.001	-0.017

fashion. Also Hobiger *et al.* (2012) show, for a field study, a direct link between the quality of the  $\Delta v/v$  estimate and the correlation coefficient.

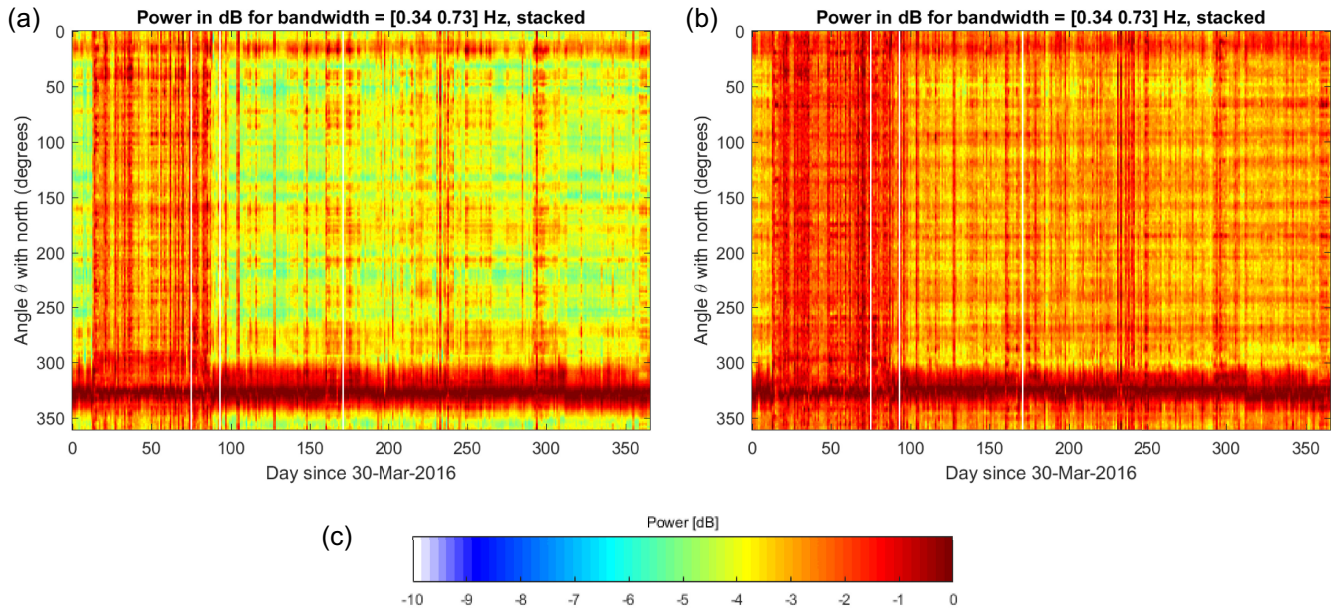
Fig. 2 shows for the Groningen case that the lapse cross-coherence does not have a high quality. It is not symmetric in time and has many spurious arrivals at acausal times, before the arrival of the surface wave. Hence, it is not a good estimate of the actual Green's function. This bad quality lapse cross-coherence can be explained by the biased illumination from the northwest as found in Fig. 8. Nevertheless, the coda that is retrieved in Fig. 2 is very stable and is consistent with the reference cross-coherence. This yields the

overall high coefficient quality values ( $Q_{CCF}$ ) when excluding the direct wave (Fig. 7). The match of  $Q_{CCF}$  with  $Q_{PII}$  confirms that it is the similarity between the lapse cross-coherence and the reference cross-coherence that determines the quality of the estimated medium change. This further supports the PII uncertainty parameter derived by Weaver *et al.* (2011) that is based on the CC (eq. 1).

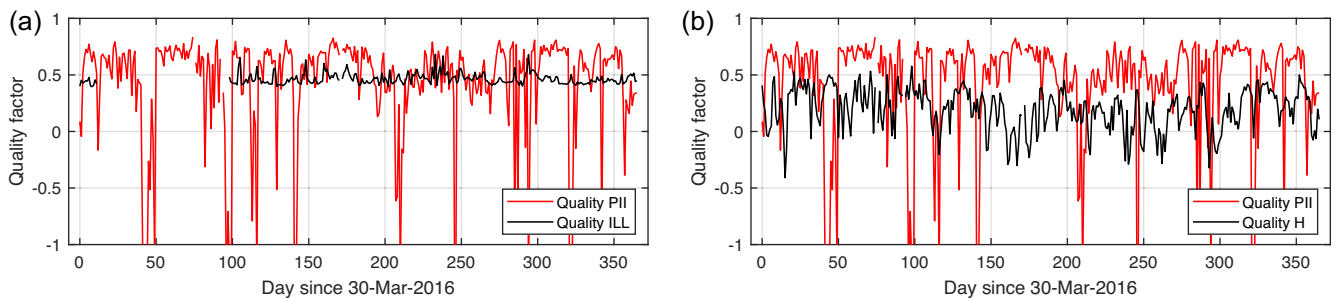
The direct wave is often excluded in PII studies (e.g. Hobiger *et al.* 2012) as it is more sensitive to changes in illumination. One might expect therefore that the PII signal becomes less oscillatory when the direct wave is left out. In Fig. 5 it can be seen that the variance over the receiver pairs decreases when the direct wave is left out. This is consistent with the reduced sensitivity to illumination. However, the amplitude of the PII signal increases when the direct wave is left out. This unexpected observation might be a specific feature of applying PII in a soft-sediment setting where most energy of the ballistic Rayleigh wave is in the first overtone (Fig. 4). The coda waves might be dominated by fundamental-mode reverberations which are more sensitive to shallower structure where changes are more likely to occur.

Similarly remarkable in Fig. 5 is the higher amplitude on the horizontal components. This might indicate that the changes are especially accrued in the surface-wave coda and that the contribution of  $P$ -wave coda on the vertical results in a damping of the  $\Delta v/v$  signal.

We did not find a strong correlation between meteorological variables and  $\Delta v/v$ . Only a weak anticorrelation of  $\Delta v/v$  with tem-



**Figure 8.** Power as function of backazimuth and day since 2016 March 30 of waves recorded on the radial (a) and on the vertical (b) components. The powers are stacked over all frequency bandwidths between 0.34 and 0.73 Hz ([0.34 0.36] Hz, [0.35 0.37] Hz, etc.). The scaling of the power is shown in (c).



**Figure 9.** Time-series of consistency quality  $Q_{PII}$  in accordance with eq. (2) (red; direct wave excluded), illumination quality  $Q_{ILL}$  in accordance with eq. (4) (a; black) and cross-coherence quality  $Q_H$  in accordance with eq. (5) (b; black) for the radial component. The correlation coefficients between the quality factors are shown in Table 2.

perature could be noted (Fig. 10). The absence of a strong correlation with the meteorological state might be attributed to the low frequencies used ( $<1.0$  Hz) which have a limited sensitivity to the top soil that is changed most by meteorological conditions. Moreover, changes in the water table are quite small over the year in Groningen (less than 1 m). This is much less than the approximately 40 m variation that was found on Merapi volcano (Sens-Schönfelder & Wegler 2006), where these ground water fluctuations were found to be the major contributor to the observed  $\Delta v/v$ .

We did find some correlation between  $\Delta v/v$  and the monthly gas production within the area sampled with PII. The large peak in the winter season could potentially be explained by a rise in production. Also we found an upgoing trend in  $\Delta v/v$  which might be linked to cumulative gas production through pressure reduction at reservoir level and compaction of both the reservoir and overlying sediments. For both observations, application of PII over a few years of data would be needed to exclude a coincidental match.

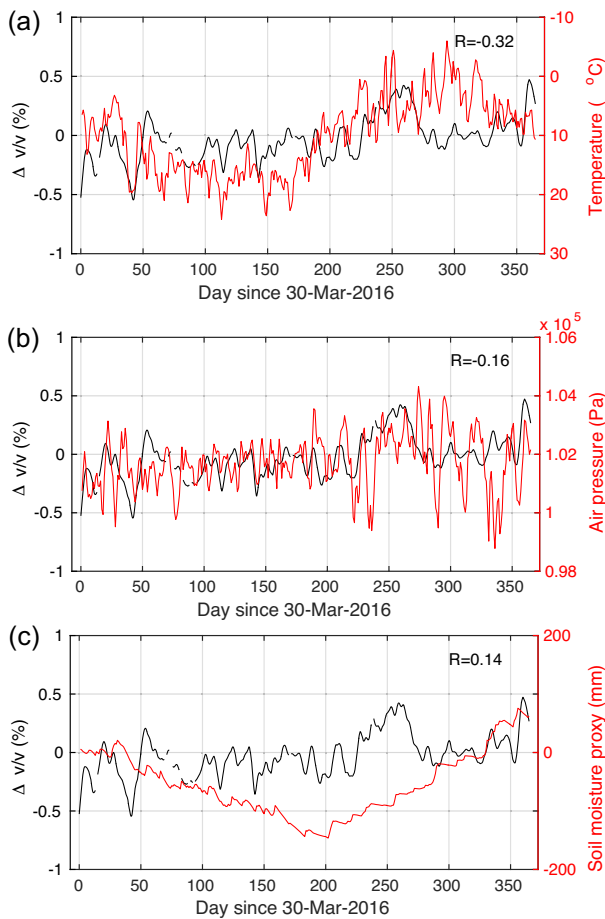
## 7 CONCLUSIONS

In the course of this research we determined dispersion relations of seismic waves for the shallow Groningen subsurface within fre-

quency bandwidth [0.30 0.95] Hz. Both for Rayleigh and Love waves we found multiple modes.

Variations of the seismic velocity ( $\Delta v/v$ ) in the shallow subsurface of Groningen were retrieved using PII. We employed two methods: one that included and the other that excluded the direct arrivals. Having station pairs sampling the same medium with many different azimuths allowed us to compute the quality (expressed as azimuthal consistency) of the determined  $\Delta v/v$ . The consistency quality of the results using the method excluding the direct arrivals were much better than those including them. We found consistent  $\Delta v/v$  for all components. However, the consistency of the estimates from the horizontal components were better than those from the vertical component.

We tested a quality factor based on the correlation coefficient between the lapse cross-coherence stretched in time and the reference cross-coherence. This quality factor did show high correlation with the azimuthal consistency of  $\Delta v/v$ . Therefore, also in a setting with a single receiver combination, a quality factor can be assigned to the estimated velocity change. The similarity of the lapse cross-coherence and the reference cross-coherence increases when the temporal resolution is decreased; one can improve estimation of medium changes at the cost of temporal resolution.



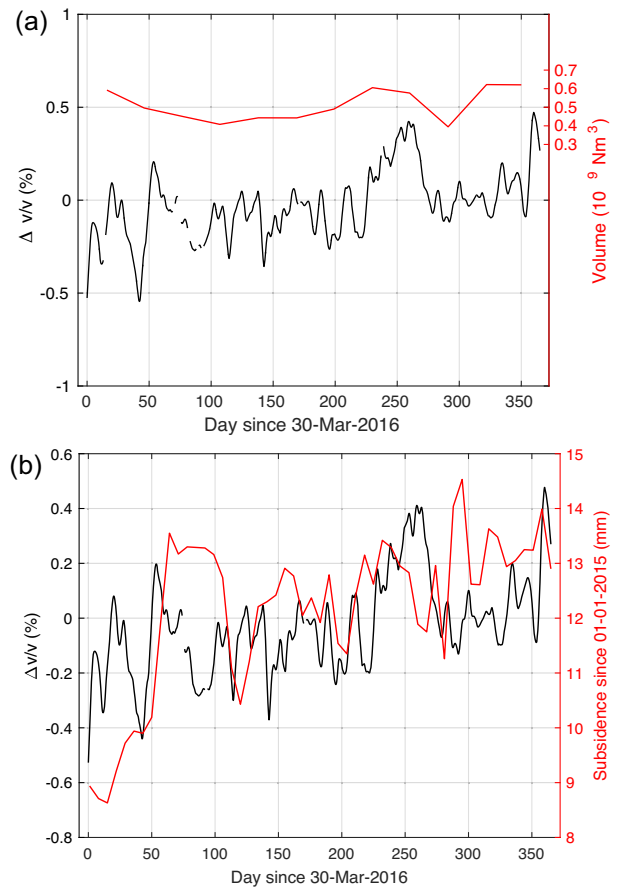
**Figure 10.** Comparison of estimated velocity change with meteorological variables detected in Eelde, NL (Fig. 1a). The black lines in the three panels are the mean medium change as estimated over the radial component (Fig. 6a). It is compared with the daily average temperature (a), the daily average air pressure (b) and a proxy for the daily average soil moisture (c). In the upper right-hand corner of the panels the normalized zero-lag cross-correlation coefficients ( $R$ ) are listed.

Variations within the power distribution of the noise were determined using cross-correlation beamforming between all shown receiver pairs. This also resulted in a quality factor. We investigated the correlation between the illumination quality and the consistency quality. It appeared that these qualities were not correlated.

We also evaluated the power within the cross-coherences from before and after the expected first arrival. This also resulted in a quality factor that did not correlate with the consistency quality.

As in previous research we found that a balanced illumination is not needed to retrieve velocity variations of good quality using PII. Nor is it necessary to have a good quality lapse Green's function estimate. It is important, however, to have a large resemblance between the lapse Green's function and the reference Green's function.

The observed medium change is likely a complicated combination of different processes taking place. We did not find one parameter that can largely explain the observations. We found a weak anticorrelation with temperature and we found a weak correlation with local gas production data and subsidence data.



**Figure 11.** Comparison of estimated velocity change with gas production data (a) and the subsidence in Loppersum (b), which is at the northwestern edge of the used subarray (Fig. 1b). The black line is the mean medium change as estimated over the radial component (Fig. 6a). The red line in (a) depicts the monthly production in billion Normal cubic metres ( $10^9 \text{ Nm}^3$ ). The production is a sum of six production clusters that are within the area that is spanned by the used network (Fig. 1); the names of these production clusters are 't Zandt, Leermens, Bierum, Ten Post, Overschild and Amsweer. The red line in (b) shows the subsidence obtained with InSAR (bodemdalingsskaart.nl).

## ACKNOWLEDGEMENTS

The authors kindly thank Hanneke Paulssen and Peter Fokker for their constructive suggestions and Gerhard Blab for his valuable advice. We thank Arie Verdel and Dirk Kraaijpoel for their insights and suggestions regarding possible explanations of the velocity variations. We much appreciate a thorough review by Qing-Yu Wang and an anonymous reviewer, who helped improving the manuscript. This work was supported by the Netherlands Organisation for Scientific Research (NWO), research programme DeepNL.

## REFERENCES

- Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercissian, A., 2008. Towards forecasting volcanic eruptions using seismic noise, *Nat. Geosci.*, **1**(2), 126–130.
- Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N., Briand, X., Emoto, K. & Miyake, H., 2014. Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science*, **345**(6192), 80–82.
- de Jager, J. & Visser, C., 2017. Geology of the Groningen field—an overview, *Neth. J. Geosci.*, **96**(5), s3–s15.



- Dost, B., Ruigrok, E.N. & Spetzler, J., 2017. Development of seismicity and probabilistic hazard assessment for the Groningen gas field, *Neth. J. Geosci.*, **96**(5), s235–s245.
- Duin, E.J.T., Doornenbal, J.C., Rijkers, R.H.B., Verbeek, J.W. & Wong, T.E., 2006. Subsurface structure of the Netherlands—results of recent onshore and offshore mapping, *Neth. J. Geosci.*, **85**(4), 245–276.
- Hadziioannou, C., Larose, E., Coutant, O., Roux, P. & Campillo, M., 2009. Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: laboratory experiments, *J. acoust. Soc. Am.*, **125**(6), 3688–3695.
- Hilliers, G., Husen, S., Obermann, A., Planès, T., Larose, E. & Campillo, M., 2015. Noise-based monitoring and imaging of aseismic transient deformation induced by the 2006 Basel reservoir stimulation, *Geophysics*, **80**(4), KS51–KS68.
- Hobiger, M., Wegler, U., Shiomi, K. & Nakahara, H., 2012. Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan, *J. geophys. Res.*, **117**(B9), doi:10.1029/2012JB009402.
- Hofman, L.J., Ruigrok, E.N., Dost, B. & Paulssen, H., 2017. A shallow seismic velocity model for the Groningen area in the Netherlands, *J. geophys. Res.*, **122**(10), 8035–8050.
- Kimman, W.P., 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.
- Lecocq, T., Caudron, C. & Brenguier, F., 2014. MSNoise, a python package for monitoring seismic velocity changes using ambient seismic noise, *Seismol. Res. Lett.*, **85**(3), 715–726.
- Lobkis, O.I. & Weaver, R.L., 2003. Coda-wave interferometry in finite solids: recovery of P-to-S conversion rates in an elastodynamic billiard, *Phys. Rev. Lett.*, **90**(25), 254302, doi:10.1103/PhysRevLett.90.254302.
- Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc. Lond., A*, **243**(857), 1–35.
- Muntendam-Bos, A.G., Roest, J.P.A. & de Waal, H.A., 2017. The effect of imposed production measures on gas extraction induced seismic risk, *Neth. J. Geosci.*, **96**(5), s271–s278.
- Obermann, A., Froment, B., Campillo, M., Larose, E., Planès, T., Valette, B., Chen, J. & Liu, Q., 2014. Seismic noise correlations to image structural and mechanical changes associated with the  $M_w$  7.9 2008 Wenchuan earthquake, *J. geophys. Res.*, **119**(4), 3155–3168.
- Park, C.B., Miller, R.D. & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel record, in *1998 SEG Annual Meeting*, Society of Exploration Geophysicists, doi:10.1190/1.1820161.
- Poupinet, G., Ellsworth, W.L. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. geophys. Res.*, **89**(B7), 5719–5731.
- Ruigrok, E.N., Gibbons, S. & Wapenaar, K., 2017. Cross-correlation beam-forming, *J. Seismol.*, **21**(3), 495–508.
- Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.*, **33**(21), doi:10.1029/2006GL027797.
- Spica, Z.J., Nakata, N., Liu, X., Campman, X., Tang, Z. & Beroza, G.C., 2018. The ambient seismic field at Groningen gas field: an overview from the surface to reservoir depth, *Seismol. Res. Lett.*, **89**, 1450–1466.
- Takano, T., Nishimura, T. & Nakahara, H., 2017. Seismic velocity changes concentrated at the shallow structure as inferred from correlation analyses of ambient noise during volcano deformation at Izu-Oshima, Japan, *J. geophys. Res.*, **122**(8), 6721–6736.
- van Eijs, R.M. & van der Wal, O., 2017. Field-wide reservoir compressibility estimation through inversion of subsidence data above the Groningen gas field, *Neth. J. Geosci.*, **96**(5), s117–s129.
- van Thienen-Visser, K. & Fokker, P.A., 2017. The future of subsidence modelling: compaction and subsidence due to gas depletion of the Groningen gas field in the Netherlands, *Neth. J. Geosci.*, **96**(5), s105–s116.
- Wang, Q.-Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T. & Aoki, Y., 2017. Seasonal crustal seismic velocity changes throughout Japan, *J. geophys. Res.*, **122**(10), 7987–8002.
- Wapenaar, K., Slob, E., Snieder, R. & Curtis, A., 2010. Tutorial on seismic interferometry: Part 2—underlying theory and new advances, *Geophysics*, **75**(5), 75A211–75A227.
- Weaver, R.L., Hadziioannou, C., Larose, E. & Campillo, M., 2011. On the precision of noise correlation interferometry, *J. geophys. Int.*, **185**(3), 1384–1392.
- Weemstra, C. *et al.*, 2016. Time-lapse seismic imaging of the Reykjanes geothermal reservoir, in *Proceedings of the European Geothermal Congress*, European Geothermal Energy Council (EGEC), Strassbourg.
- Wegler, U., Nakahara, H., Sens-Schönfelder, C., Korn, M. & Shiomi, K., 2009. Sudden drop of seismic velocity after the 2004  $M_w$  6.6 mid-Niigata earthquake, Japan, observed with passive image interferometry, *J. geophys. Res.*, **114**(B6), doi:10.1029/2008JB005869.

## SUPPORTING INFORMATION

Supplementary data are available at [GJI](https://doi.org/10.1002/gji.1367/5490351) online.

**Day\_DVV.SIGMA.R.excl.txt**

**Day\_DVV.SIGMA.T.excl.txt**

**Day\_DVV.SIGMA.Z.excl.txt**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.