Empirical Green's function analysis of some induced earthquake pairs from the Groningen gas field

Steve Oates · Jelena Schmitz · Brian Zurek · Thomas Piesold · Ewoud van Dedem

Abstract We have applied the empirical Green's function (EGF) method to 53 pairs of earthquakes, with magnitudes ranging from $M = 0.4$ to $M = 3.4$, induced by gas production from the Groningen field in the Netherlands. For a subset of the events processed, we find that the relative source time functions obtained by the EGF deconvolution show clear indications of a horizontal component of rupture propagation. The earthquake monitoring network used has dense azimuthal coverage for nearly all events such that wavelet duration times can be picked as a function of source-station azimuth and inverted using the usual Doppler broadening model to estimate rupture propagation strike, distance, and velocity. Average slip velocities have also been estimated and found to be in agreement with typical published values. We have used synthetic data, from both a simple convolutional model of the seismogram and more sophisticated finite difference rupture simulations, to validate our data processing workflow and develop kinematic models which can explain the observed characteristics of the field data. Using a measure based on the L1-norm to discriminate results of differing quality, we find that the highest quality results show very good alignment of the rupture propagation with directions of the detailed fault map, obtained from the full-field 3D seismic data. The dip direction rupture extents were estimated from the horizontal rupture propagation distances and catalogue

---

Brian Zurek Currently unaffiliated.

Highlights
- Pairs of Groningen-induced earthquakes have been analysed to reveal details of the propagation of the ruptures.
- Estimated rupture dimensions are consistent with all but one event being able to be contained within the reservoir.
- The earthquake rupture propagation directions found are consistent with the faults mapped on 3D seismic images.

S. Oates (✉)
Shell Research Limited, Aberdeen, UK
e-mail: Steve.Oates@Shell.com

J. Schmitz
ExxonMobil Technology and Engineering Company, 22777 Springwoods Village Parkway, Spring, TX 77389, USA

J. Schmitz
Formerly at Nederlandse Aardolie Maatschappij B.V. (NAM), Schepersmaat, 2, 9405-TA Assen, The Netherlands

B. Zurek
Formerly at ExxonMobil Upstream Research Company, 22777 Springwoods Village Parkway, Spring, TX 77389, USA

T. Piesold
Shell Research Limited, London, UK

E. van Dedem
Shell Global Solutions International, The Hague, The Netherlands

🖄 Springer

magnitudes showing that, for all but the largest magnitude event (the $M = 3.4$ event of 8th January 2018), the dip-direction extent is sufficiently small to be contained wholly within the reservoir.

1 Introduction

Gas production from the Groningen gas field (Fig. 1) started in 1963. It is the largest natural gas field in Europe with initial reserves estimated at 2900 billion cubic meters (bcm). The reservoir is a high quality, heavily faulted sandstone, some 200- to 300-m thick over most of the field. Annual production volumes were determined by demand between 1995 and 2014; since then production has been repeatedly cut in response to concerns about induced seismicity. Triggered tectonic events are considered to be very unlikely in the North of the Netherlands and it is now accepted that seismicity in the region is induced by gas production, from Groningen and other fields. The first production-related earthquake detected over the Groningen field occurred in 1991. Bourne and Oates (2017) reported that the number of earthquakes per unit gas production increases with cumulative gas production; Bourne et al. (2018) have shown the level

Fig. 1 Location map with earthquake epicentres (all events with $M_L \geq 1.5$ between 1995 and 2017) shown as red circles (local magnitudes in legend), the Groningen field outline, mapped geological faults, seismometer stations (blue triangles), and building locations (grey dots). Coordinates are in thousands of metres in the Dutch RD system. Originally published in van Elk et al. (2019)
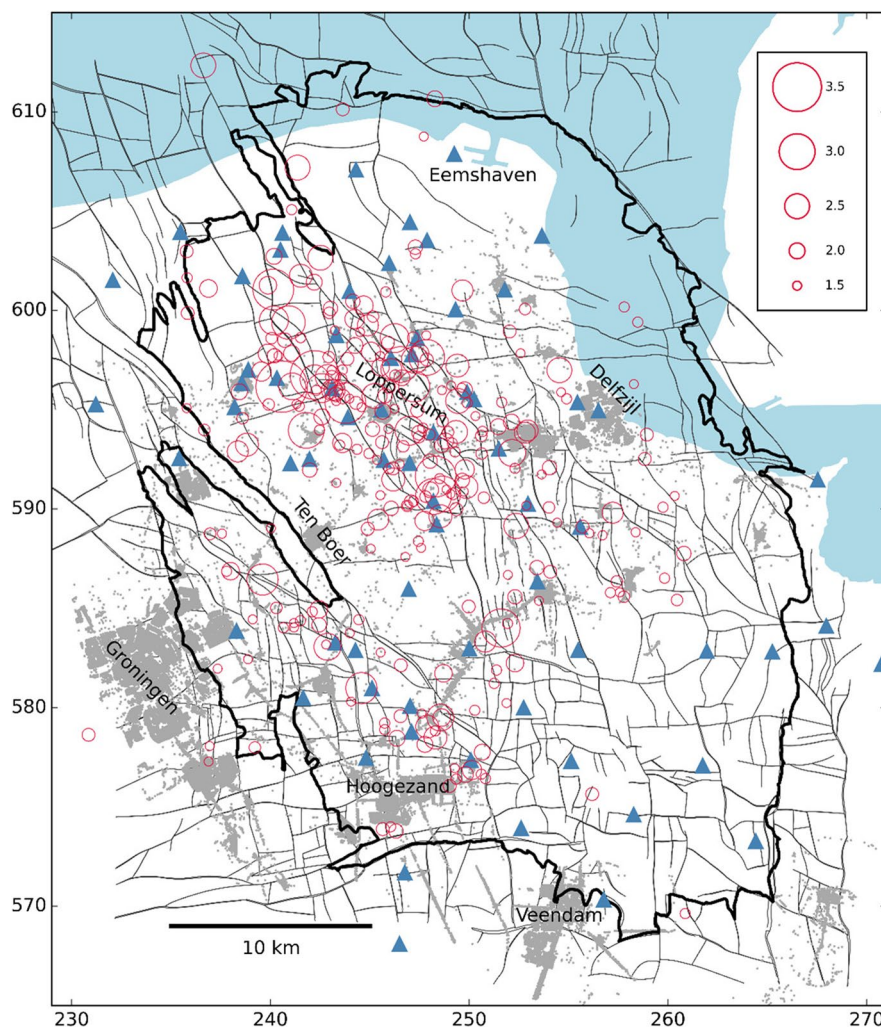
of induced seismic activity increasing exponentially with the induced stress. The largest induced earthquake in the Groningen catalogue was of M_L 3.6 and occurred in August, 2012, near the village of Huizinge, in the central part of the field. This event was felt widely, leading to many claims for minor building damage, and has triggered much public debate.

Following the emergence of induced seismicity, a network of near-surface stations was deployed over the NE Netherlands by the Royal Netherlands Meteorological Institute (KNMI). This sparse network comprised only six stations over or near the Groningen field. Each of the near-surface stations comprised three-component geophones at four depth levels in boreholes of up 300 m deep. This network's magnitude of completeness is taken to be M_L 1.5 (Dost et al. 2017). Surface accelerographs were also deployed. As part of the response to the 2012 Huizinge event, the earthquake monitoring network was densified to lower the magnitude of completeness and improve the precision of event locations (Dost et al. 2017). Starting in 2014, some 70 stations, with sensors placed every 50 m along 200 m deep boreholes, were added. The expanded network has a magnitude of completeness of about M_L 0.5, and epicentral location uncertainties are about 100–300 m. Except for some low-magnitude events, most hypocentres have been located within the Rotliegend sandstone reservoir (Spetzler and Dost 2017). To complement the KNMI network, geophones were deployed over the reservoir interval in two deep monitoring wells in the region of highest seismic activity. This has enabled detection of micro-earthquakes, with magnitudes below the detection threshold at surface. KNMI processes the data from the near-surface network, making openly available catalogues of event times, epicentres, depths, and magnitudes. Processing using full waveform inversion methods to obtain locations and moment tensors has been carried out by Willacy et al. (2020), Kühn et al. (2020), and Dost et al. (2020). Further details of the Groningen field and its induced seismicity are given by De Jager and Visser (2017) and Dost et al. (2017).

The Groningen earthquake catalogue generated by KNMI contains some 1400 located events up to the end of 2020. Event hypocenters are assumed to be located in the gas reservoir and assigned a nominal depth of 3 km below surface unless the data provides sufficient constraint for the depth to be generated by the location inversion. The densification of the network, the deployment of deep downhole geophones, and the use of full waveform inversion methods, as described above, have confirmed the assumption that events initiate in the gas reservoir (Daniel et al. 2016). The subsurface velocity profile (Willacy et al. 2019) shows the Rotliegend sandstone reservoir overlain by a thick package of Zechstein evaporites and with high-velocity carboniferous rocks below. As such, the reservoir interval which hosts the induced earthquakes acts as a wave guide, resulting in sub-horizontal ray paths (Kraaijpoel and Dost 2013; Jagt et al. 2017) and complex patterns of arrivals due to multiple reflections and mode conversions (Willacy et al. 2019).

The work of Willacy et al. (2020) established a clear connection between the mapped faults and event locations and focal mechanisms, thereby building confidence in the statistical geomechanical models used to forecast future activity and assess hazard and risk (van Elk et al. 2019). The motivation for the present work was to add information on rupture propagation to this body of knowledge to further our geomechanical understanding.

We have applied the empirical Green's function (EGF) method to strongly correlated pairs of the earthquakes, with the aim of detecting finite size rupture propagation effects in this population of small magnitude events. EGF analysis is conventionally applied to earthquake pairs with a magnitude difference of at least 1. We show here how the EGF deconvolution process can generate useful rupture propagation information for pairs of events with comparable magnitudes. The rupture directivity analysis method originally developed by Savage (1965) has been applied to the relative source time functions (RSTFs) resulting from EGF processing (see for example Folesky et al. 2016) to obtain estimates of horizontal rupture direction and length for comparison with mapped reservoir-level faulting.

Wang et al. (2014) have analysed the azimuthal variations of the durations of RSTFs for a rich set of composite events on part of the San Andreas Fault system and inverted their results for the separation and centroid time delay of the composites' sub-events. Park and Ishii (2015) present a model which can accommodate more complex patterns of duration variation from non-unilateral rupture propagation which also takes into account the dip component of the propagation

vector. Yoshido et al. (2019) applied the EGF method and subsequent directivity analysis to a swarm of earthquakes following the 2011 Tohoku-Oki earthquake and compared the rupture propagation directions obtained with mapped trajectories showing the migration of hypocentres along fault structures. Ameri et al. (2020) have applied the directivity analysis to a subset of Groningen earthquake seismogram data without EGF processing. The subset of events selected for directivity analysis by Ameri et al. (2020) has only limited overlap with the set of events analysed here. As such, their results and, in particular, their comparison of rupture directions with the mapped faults complement the results presented here.

2 Theoretical background

The review article by Hutchings and Viegas (2012) gives an overview of the development and applications of the empirical Green's function method. Here we give a concise account of the EGF method in the form in which we have applied it to the Groningen data.

The empirical Green's function method utilises spatially approximately collocated pairs of events of significantly different sizes to remove by deconvolution the effects of propagation through the subsurface (Li et al. 1995a). The aim is to give a cleaner representation of the source time function which is then suitable for further quantitative analysis. Consider two collocated events, a larger event, U_l, and a smaller event, U_g, which will be taken as the empirical Green's function. Moreover, the focal mechanism is taken to be the same for both events. We use here a simple convolutional model which expresses the seismogram as the convolution of a wavelet and a series of pulses corresponding to the discrete arrivals recorded. $S(t)$ is the source time function, $P(t)$ represents the subsurface propagation effects, and $R(t)$ and $I(t)$ represent, respectively, the effects of the recording system and the instrument response. The velocity seismograms for the two events, recorded by the same seismometer, are expressed as follows (for idealised noise-free data):

$$U_l(t) = F_l(r, \theta, \phi) M_l S_l(t) * P(t) * R(t) * I(t) \quad (1)$$

$$U_g(t) = F_g(r, \theta, \phi) M_g S_g(t) * P(t) * R(t) * I(t) \quad (2)$$

Here the source time functions are scaled by the radiation patterns in spherical coordinates and the scalar moments, $F_l(r, \theta, \phi)$ and $F_g(r, \theta, \phi)$, and M_l and M_g, respectively. The propagation term, $P(t)$, is a filter representing spreading, absorption, and scattering processes. Here, it is sufficient to recognise that the terms which have not been fully specified are nevertheless assumed to be the same for the pair of collocated events. Deconvolving $U_l(t)$ with $U_g(t)$ causes the common contributions to cancel leaving the relative source time function, $S_r(t)$:

$$U_l(t) * \left(U_g(t) \right)^{-1} = \frac{M_l}{M_g} S_l(t) * \left(S_g(t) \right)^{-1} = S_r(t). \quad (3)$$

Moreover, if the source time function of the smaller event is small enough that it can be regarded as an impulsive point source, $S_g(t) \approx \delta(t)$, then the relative source time function is an approximation of the larger event's source time function, scaled by the ratio of the moments, $S_r(t) \approx (M_l/M_g) S_l(t)$.

The deconvolution is usually implemented as a spectral division. The algorithm used here multiplies the larger (parent) event's spectrum, $\widetilde{U}_l(\omega)$, by the deconvolution function, $D(\omega)$, derived from the spectrum of the smaller (child) event, $\widetilde{U}_g(\omega)$ as follows:

$$D(\omega) = \frac{\widetilde{U}_g(\omega)}{\widetilde{U}_g^2(\omega) + N.\max(\widetilde{U}_g^2(\omega))}. \quad (4)$$

Here, N is an additive noise term used to stabilise the deconvolution outside of the signal band of the recorded seismogram (Berkhout 1977). Phase deconvolution is the corresponding subtraction of the phase spectra. The output from the amplitude deconvolution is

$$\widetilde{U}_l(\omega)D(\omega) = \frac{\widetilde{U}_g(\omega)\widetilde{U}_l(\omega)}{\widetilde{U}_g^2(\omega) + N.\max\left(\widetilde{U}_g^2(\omega) \right)} \quad (5)$$

and this only reduces to the exact EGF relative source time function in the limit of zero additive noise,

$$\lim_{N \to 0} \widetilde{U}_l(\omega)D(\omega) = \frac{\widetilde{U}_l(\omega)}{\widetilde{U}_g(\omega)}. \quad (6)$$

In practice, this additive noise term is needed to stabilise the deconvolution results (see for example

Folesky et al. 2016): it is important therefore to understand its impact on the output of the EGF process. The effect of the added noise is to attenuate the RSTF at high and low frequencies, at and beyond the edges of the signal band, whereas the effect in the time domain is to smooth the RSTF. The limit of $N \gg 1$ will also turn out to be important—in this case the spectral division of Eq. (5) reduces to the scaled product of the spectra of the two events which, in the time domain, corresponds to the scaled convolution of the seismograms or, equivalently, the cross-correlation after time reversal of one of the seismograms.

3 Description of the dataset analysed

The event data processed has been acquired over the Groningen gas field in The Netherlands (De Jager and Visser 2017). These earthquakes, in the magnitude range from $M = 0.4$ to $M = 3.4$, have been induced by gas production from the field. The seismograms were recorded using the KNMI network and the data downloaded from the publicly accessible KNMI data portal at http://rdsa.knmi.nl/dataportal (Dost et al. 2017). This network, as shown in Fig. 1, comprises some 70 stations to give a spatially dense coverage of the gas field with a generally very good azimuthal coverage for the majority of events. Each of these stations comprises an approximately 200-m-deep borehole with 4.5Hz 3C geophone units grouted in place at 200 m, 150 m, 100 m, and 50 m below the surface with, in addition, an accelerometer unit and multichannel data recorder and transmission system at surface. The data was acquired at 200sps, giving a Nyquist frequency of 100 Hz: for 200sps, the recorder has a flat frequency response from DC to 80 Hz. Between 80 and 100 Hz the signal is strongly attenuated by the recorder's anti-alias filter which gives more than 140 dB of attenuation at the output Nyquist frequency. The spectrum beyond 80 Hz is dominated by instrument noise (Spica et al. 2018). As processed here, the data had not been instrument corrected. Data from the two deep downhole monitoring wells (Daniel et al. 2016) was not used in this study.

As described by Van Dedem et al. (2018), seismograms were cross-correlated for all event pairs, using a procedure similar to that due to Arrowsmith and Eisner (2006). The maximum amplitude of the stack over all stations of the normalised cross-correlation traces was used as a discriminant to identify the event pairs with high waveform similarity for EGF processing. Examples are shown in Fig. 2 and Van Dedem et al. (2018). It should be emphasised that here we have identified event pairs or clusters on the basis of cross-correlation alone, whereas it is more common to also require small location separation and significant or large magnitude differences. We argue below that not having vanishingly small separations and large magnitude differences is not in itself a barrier to obtaining useful results from the deconvolution process. Jagt et al. (2017) describe cluster analysis of Groningen events using data from the same KNMI database over the period January, 2010–January, 2016, for the purpose of relative location refinement.

4 Simple kinematic models of extended rupture

The objective of the EGF processing applied was to detect signs of finite rupturing as distinct from the effective point source approximation often assumed for low-magnitude events. As explained above, the Groningen earthquake hypocentres are located within the Rotliegend sandstone reservoir which acts as a waveguide due to the high velocity layers directly above and below. Except for the shortest epicentral distances, the horizontal components of the source-receiver ray paths are assumed to be accounted for by propagation through the reservoir. Ray tracing (Kraaijpoel and Dost 2013; Jagt et al. 2017) and finite difference wave propagation simulations (Willacy et al. 2019) indicate that we should not expect to be able to access a significant range of take-off dip angles. We therefore assume that only the horizontal component of rupture propagation will be detectable using the EGF method. This contrasts with the work of Park and Ishii (2015) in which the dip of the take-off vector is also addressed by the inversion scheme.

To understand our results we have built simple kinematic source models which represent the larger event as a composite of multiple slip patches (such an approach is discussed by Hutchings and Viegas 2012). The motivation was to develop a conceptual framework for interpreting the field data, provide expressions for the RSTF duration as a function of station azimuth, and enable us to generate synthetic seismograms on which to validate the data processing.

Mechanisms seen in the Groningen field are mainly normal faulting on high-angle faults (Willacy et al.
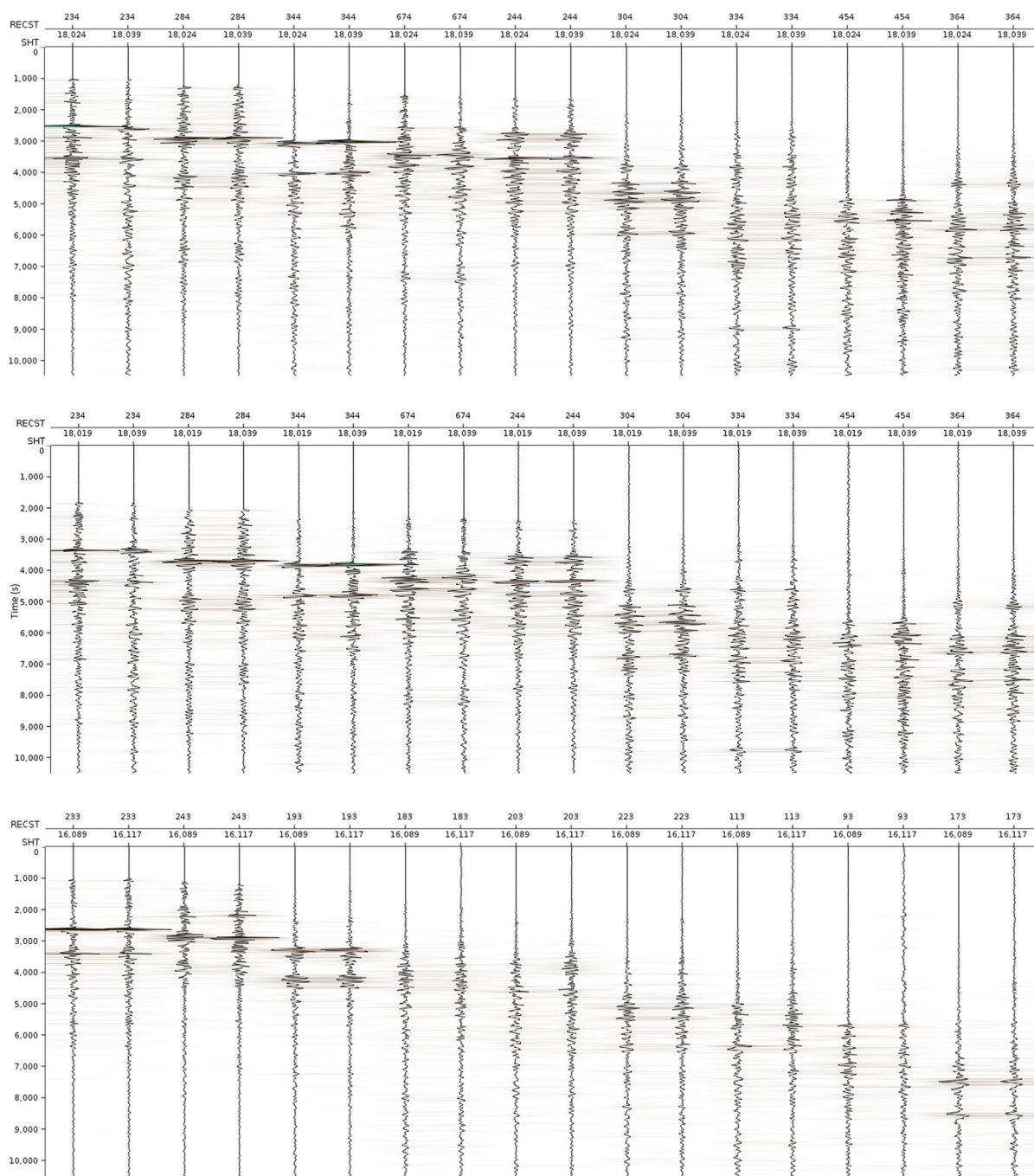
Fig. 2 Panels showing example seismogram pairs from the event pairs shown in Figs. 6(a), 7(a), 8(a), 9(a) and 10(a). The traces are from the east orientation horizontal component of the 3C geophone unit at depth 150 m: the stations with the shortest epicentral distances have been selected. The direct compressional and shear arrivals can be easily identified along with a shear interbed multiple. Trace identifiers give event ID (SHT) and receiver station number (RECST). A long gate automatic gain control (AGC) has been applied to bring individual seismogram amplitudes to a common level for visual comparison. The final panel is a zoom of part of the first panel (event pair 18024/18039) to illustrate the waveform similarity in more detail. Event separations for the event pairs shown are, respectively, 158 m (18024/18039), 71 m (18019/18039), 71 m (16117/16089), 122 m (17124/18002), and 293 m (17064/17052), as in Tables 2 and 3
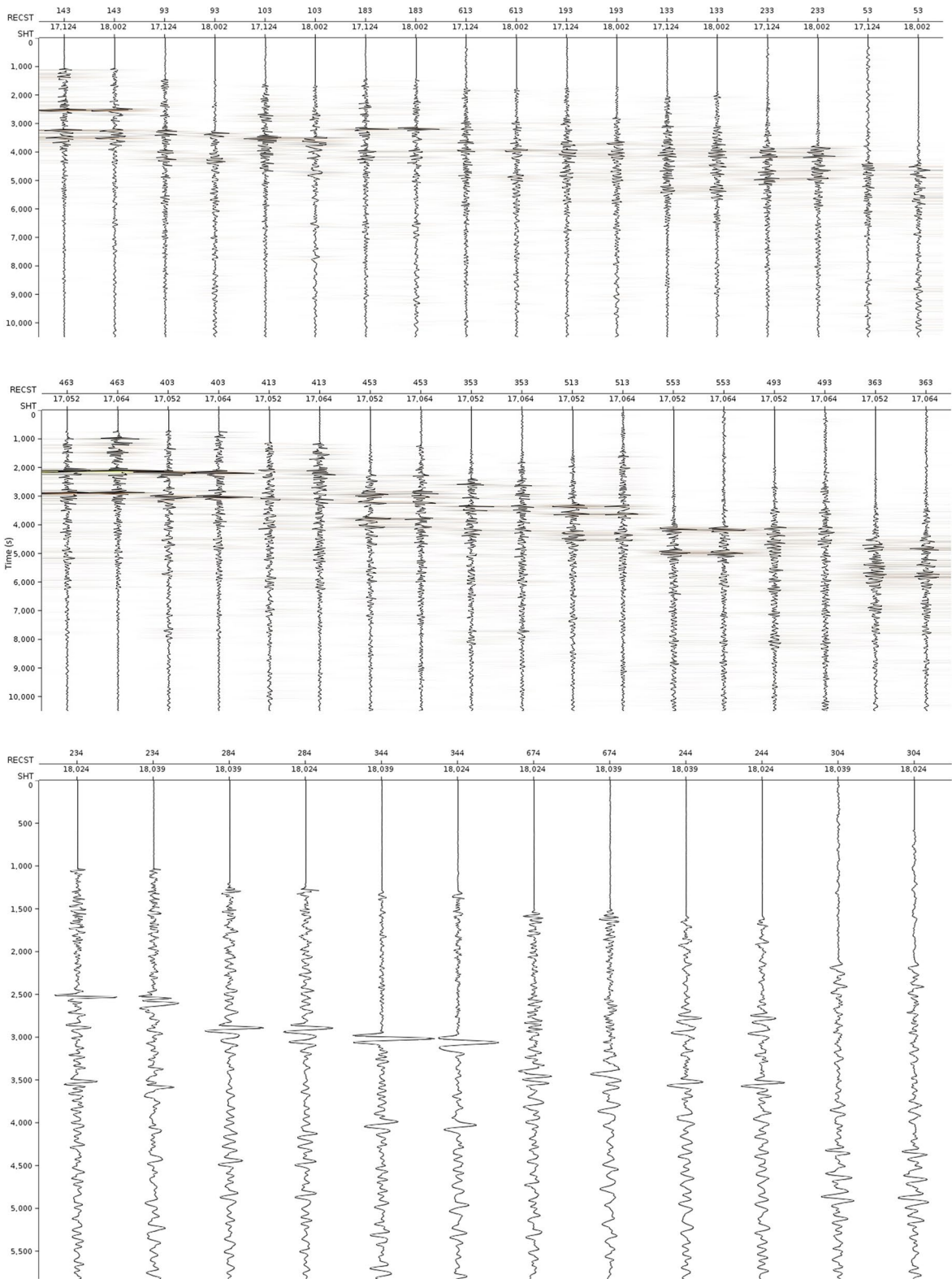
Fig. 2 (continued)

2018; Dost et al. 2020), so we consider models of mode III failure, that is rupture for which the direction of slip is perpendicular to the direction of rupture propagation (Udías et al. 2014), to model the along-strike rupture propagation. Referring to Fig. 3(a), we consider an event dominated by a starting phase and a stopping phase located at the two ends of the horizontal projection of the fault trace of length, L, and separated in time by L/ζ, where ζ is the rupture propagation velocity. Rs_i and Re_i are the epicentral distances between the source and the ith receiver station at (respectively) the start and end of the rupture process; ϕ_i is the azimuthal angle of the ith station relative to the starting phase location measured in the clockwise direction from the fault strike direction; c is the seismic velocity (P or S as required) in a homogeneous subsurface. Without loss of generality, we develop the theory in a coordinate system aligned with the fault strike direction, Φ. The cosine rule gives the epicentral distance for the stopping phase in terms of that of the starting phase, the fault trace length, and azimuthal angle, ϕ_i:

$$Re_i = \left\{ L^2 + Rs_i^2 - 2LRs_iCos\phi_i \right\}^{1/2}. \tag{7}$$

If the rupture propagation distance is small compared with the source-station offset ($Rs_i \gg L$), as will be the case for the geometries we consider here, then dropping higher order terms in L/Rs_i gives

$$Re_i \approx Rs_i \left\{ 1 - \left(\frac{2LCos\phi_i}{Rs_i} \right) \right\}^{1/2} \approx Rs_i - LCos\phi_i. \tag{8}$$

In the Appendix it is shown that the angle-dependent duration of the RSTF in the case of unilateral rupture propagation is

$$\Delta t\left(x_i\right) \approx \frac{L}{\zeta} + w_t - \frac{L}{c}Cos\phi_i. \tag{9}$$

To accommodate realistic band-limited pulses and duration measured between zero crossings, the width, w_t, of the individual pulses has been added. This is essentially the expression given by Savage (1965). It can be seen as an example of angle-dependent Doppler broadening.

For the case in which a pair of parallel ruptures, both comprising multiple slip patches, is processed
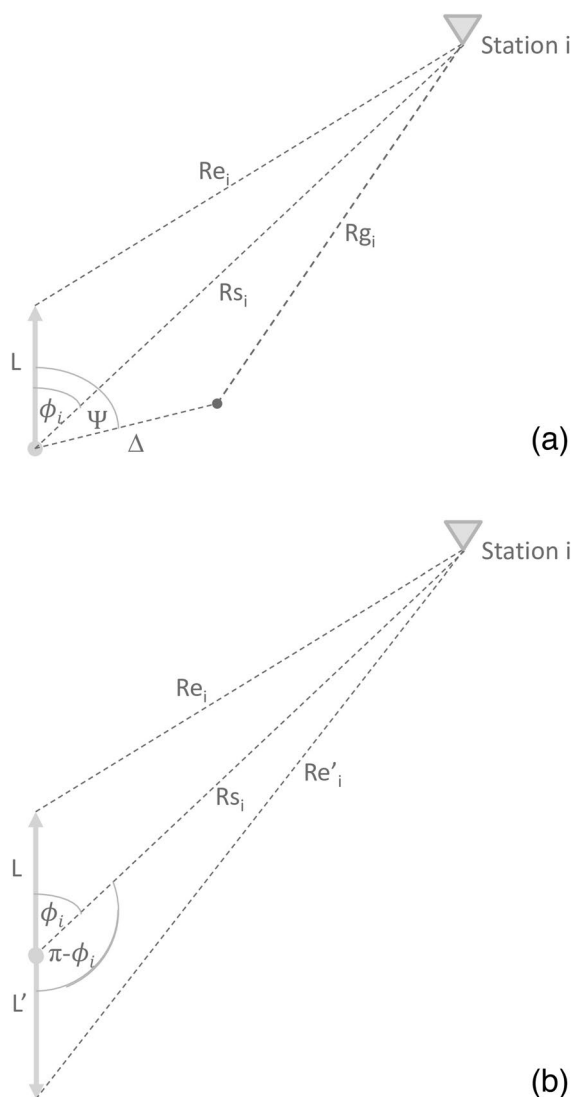
Fig. 3 The geometries considered for unilateral (**a**) and bilateral (**b**) rupture propagation. These diagrams show the definitions of angles and distances in the horizontal plane from a rupture to an individual seismometer station. Rs_i, Re_i and Re'_i are epicentral distances to, respectively, the start and end points of the ruptures; Rg_i is the epicentral distance to the child event; L and L' are rupture propagation distances; Δ is the offset of the smaller child event from the larger parent event; ϕ_i is the station azimuth; Ψ is the azimuth of the offset vector of the child event relative to the strike direction of the larger event. For the case of bilateral rupture propagation (**b**) the child event's source-receiver distance, azimuth, and offset relative to the hypocentre (i.e. initiation point) of the parent event are not shown on the diagram for reasons of simplicity but are defined as for the unilateral case (**a**)

with the EGF deconvolution workflow with the application of high additive noise, the duration expression is of the same form but with the rupture propagation distance replaced by the sum of the rupture propagation distances of the two events (as shown in the Appendix). This implies that the usual assumption for EGF processing, that the smaller event be considered to be a single pulse, is not strictly necessary for this application if a generalisation of the EGF deconvolution operator is admitted. Stabilisation of the deconvolution process is required to avoid inflating noise outside the signal band and this leads naturally to a deconvolution operator which combines the pure EGF spectral division with the convolution of parent and child seismograms. In the low N limit, the spectral division removes the effects of common operators (propagation, instrument response, band pass filters, etc.) leading to optimal recovery of duration model parameters for the parent when the child is a much smaller rupture. In the high N limit the seismogram convolution extracts an approximate RSTF from which useful estimates of the duration model parameters for ruptures of comparable size can still be made. An intermediate value of additive noise combines the effects of low and high N with limited impact on the measured duration. The usual restriction to event pairs with a magnitude difference of at least 1.0 (Hutchings and Viegas 2012) would rule out most of the event pairs identified by the cross-correlation criterion in the Groningen catalogue. We argue (see Appendix) that by adjusting the stabilising noise in the deconvolution operator useful results can also be obtained by applying the EGF deconvolution workflow to event pairs with similar magnitudes.

In the following sections we will show for some representative event pair examples how a compelling qualitative visual fit of the convolutional model data to the field data can be obtained using the rupture propagation parameter values obtained by processing the real data. It will be demonstrated to what extent the simple synthetics, when processed using the same workflow as used for the field data, return the input parameters. This we feel is an important validation of the processing scheme.

More sophisticated kinematic rupture models have been built using a finite difference code which models both the propagation of the rupture itself and the propagation of seismic waves away from the evolving rupture. The rupture is modelled as a composite failure,

comprising a number of separate slip patches with additional stochastic heterogeneity distributed over a fault plane embedded in a 3D subsurface model such that components of rupture propagation in both the dip and strike directions can be modelled. Propagation of the rupture across the fault plane is governed by causal rules which determine the triggering of a slip patch by its neighbours. As a full waveform model this approach generates seismograms which better represent real seismogram data while still being based on a simple kinematic description of the rupture process. Details of the finite difference kinematic rupture simulation approach are given in Zurek et al. (2017), Zurek and deMartin (2019), and Edwards et al. (2019). See also Graves and Pitarka (2016). Examples from this modelling approach are shown below.
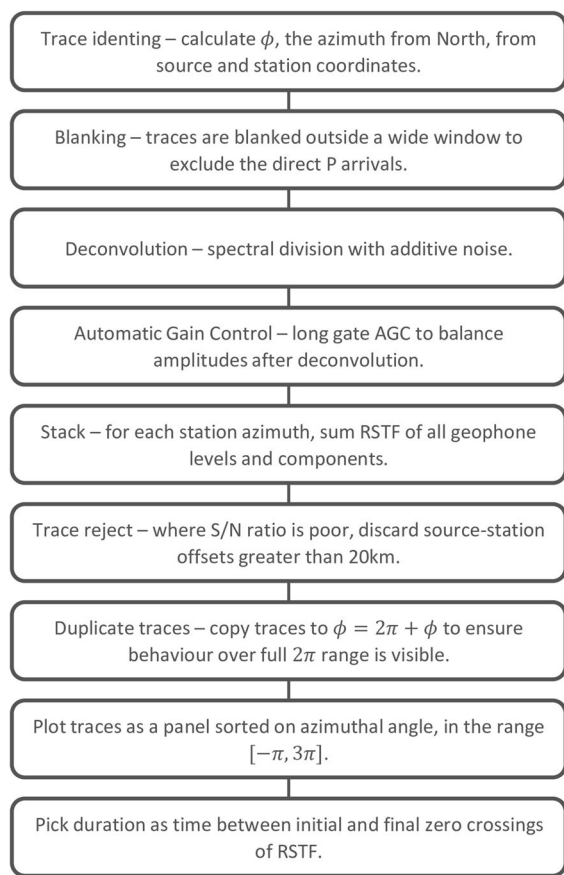
5 Data processing workflow

Our data analysis workflow is similar to that described by Tomic et al. (2009). The main steps in the EGF processing sequence applied are summarised in Fig. 4; certain aspects are described in more detail here.

Seismic data processing made use of established proprietary software (Willacy et al. 2020). As explained above, the deconvolution is implemented as a spectral division with additive noise to stabilise the process. Figure 5 illustrates the effect of varying the additive noise level on the time domain seismograms for event pair 18024/18039. The plots show the RSTFs obtained using a range of noise values between $N = 1$ and $N = 0.0001$ in the spectral division. Notice how the higher levels of additive noise smooth the traces. Although this stabilises the picking process, mis-picks still occur as shown in this example and have been manually excluded from the parameter inversion process described below. Note also how the lower N values better suppress amplitudes before and after the picked start and end times of the RSTF.

After the deconvolution step, a long gate automatic gain control (AGC) is applied to the RSTF seismogram data. This leaves the zero crossings, and hence the measured duration, unchanged but dramatically enhances the trace-to-trace consistency of the azimuth panels, helping visual interpretation.

We obtain the best results by summing the RSTFs for all four geophone levels (200 m, 150 m, 100 m,

Fig. 4 Summary of data processing steps in the empirical Green's function workflow applied

and 50 m) and components (X, Y, Z) of a station. Since the spectral division removes all effects common to both seismograms, including path and coupling effects, the RSTF should be approximately the same for all depth levels and components of the same station justifying stacking to suppress noise (also see Appendix).

The RSTF traces are duplicated by copying to azimuth values which are greater by 2π and then displayed on the interval $[-\pi, 3\pi]$. This is purely for display purposes and ensures that periodic effects which may occur close to the ends of the original azimuth range are not obscured.

RSTF durations are obtained as the time delays between picked zero crossings of the RSTF as shown. Obvious mis-picks were removed by hand excluding them from the inversion process. The

azimuthal dependency of the duration is fit using the Doppler broadening expression as derived above. Excel's Solver was used to minimise the sum of the absolute values of the residuals between the modelled and observed durations. This gives values for the parameters L, ζ and the fault strike, Φ; the pulse width, w_t, needs to be estimated separately.

To generate the synthetic seismograms used, arrivals from an assumed source location at the actual station locations were generated with travel times derived for a constant velocity subsurface, as shown in the Appendix. The synthetic seismograms were then band pass filtered to match the frequency content of the field data and white noise was added to represent the seismograms' post-AD converter noise floor (see Havskov and Alguacil 2004; Spica et al. 2018).

6 Description of results

A total of 53 event pairs with high cross-correlation values have been processed using the workflow described above. RSTFs for some representative example event pairs are shown in Figs. 6(a), 7(a), 8(a), 9(a) and 10(a). The results presented here are for traces windowed on the S arrivals. In all cases, the Doppler broadening expression for the unilateral propagation model was fitted to the arrival picks of the RSTFs generated (Figs. 6(b), 7(b), 8(b), 9(b) and 10(b)). The Generalized Reduced Gradient Nonlinear solver functionality in Excel was used to minimise the L1 sum of the residuals by varying the three parameters $\{A_1, A_2, \Phi\}$ describing the azimuthal dependence of the wavelet duration:

$$\Delta t(x_i) \approx A_1 - A_2 Cos(\phi_i - \Phi) \qquad (10)$$

where $A_1 = L/\zeta + w_t$, $A_2 = L/c$, and Φ is the fault strike.

There is a well-developed 3D model of seismic velocities over the Groningen field based on full-field 3D seismic data coverage and numerous well logs (Edwards et al. 2019). This model has been used in locating seismic events as well as being the basis for the finite difference rupture simulations described above. Processing of these synthetic data leads to a shear velocity of $Vs = 2435$ m/s to match
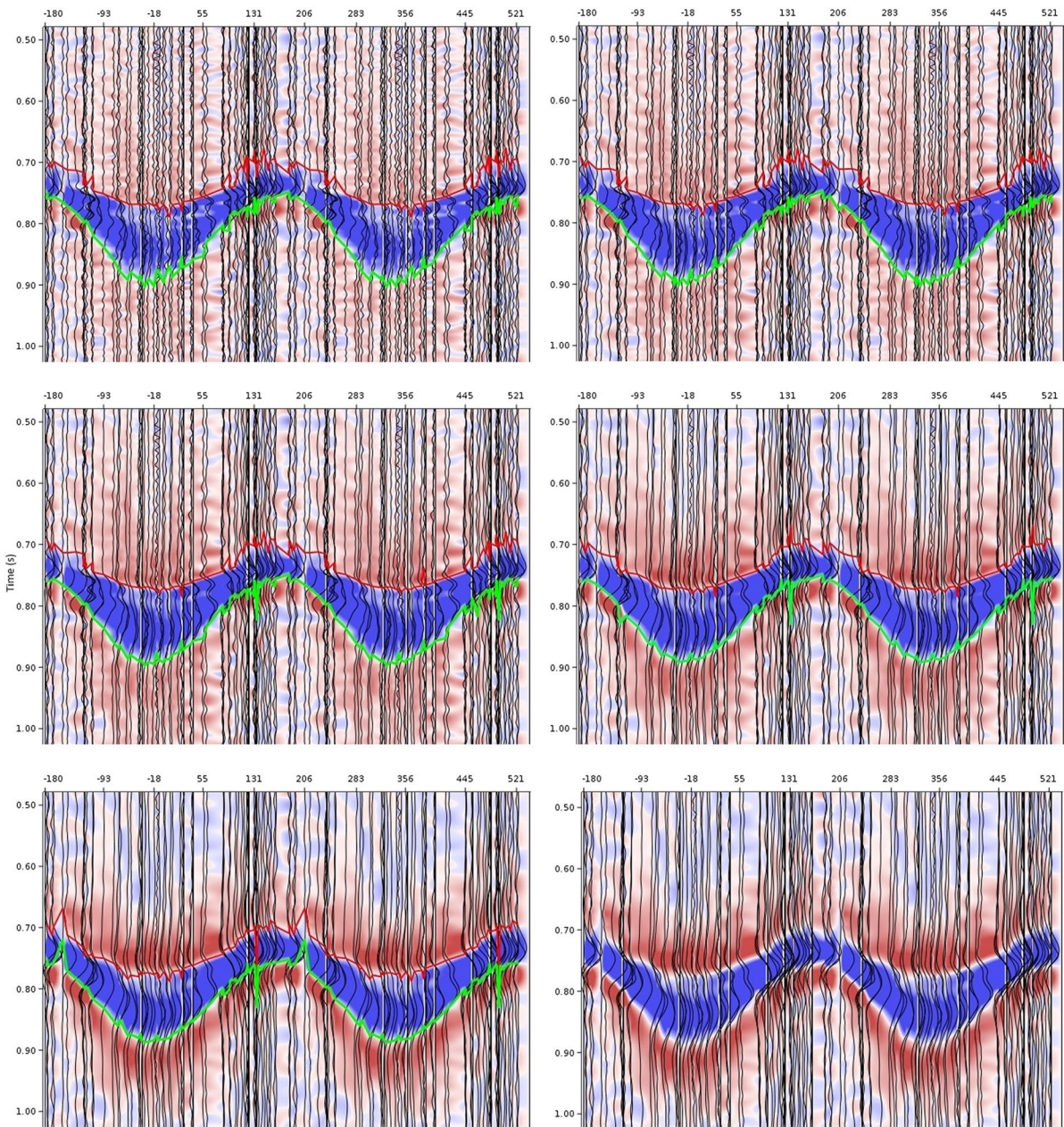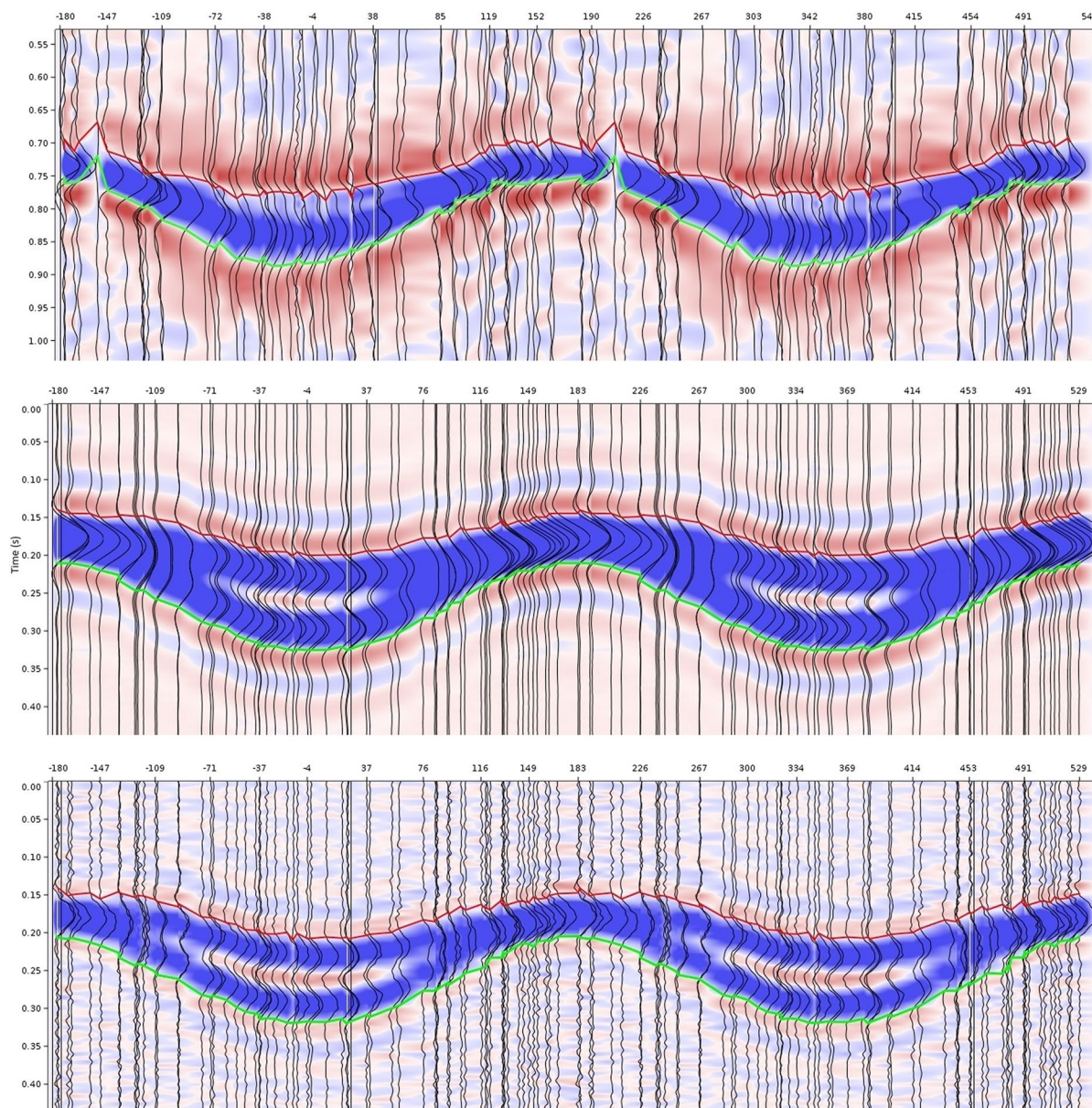
Fig. 5 EGF output for event pair 18024/18039 using a range of deconvolution noise values $N=0.0001$ (top left), 0.001 (top right), 0.01 (centre left), 0.1 (centre right), and 1.0 (bottom row, with and without time picks). The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. Here and in subsequent plots, each trace gives the RSTF for a single station (stacked over the four depth levels and three components); the background colours indicate trace amplitudes (interpolated between traces) using a standard red-blue colour scale for negative and positive values; the red and green curves are the picked zero crossings from which the RSTF duration is determined; the trace identifiers give the azimuths ϕ_i of the stations in ° clockwise from N

the inverted value of the rupture propagation length with the model input value. We take this shear velocity as being a representative value for sub-horizontal propagation in the high-velocity evaporite layers on top of the reservoir interval and have also used it in generating the convolutional model synthetics.

(a)

Fig. 6 a EGF trace output for event pair 18024/18039 with horizontal source-station offsets restricted to 20 km (top) and the accompanying convolutional model synthetic data processed with $N=1$ (middle) and $N=0.0001$ (bottom). The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. **b** Picked duration between the zero crossings of the RSTF traces (top) for event pair 18024/18039 (blue data points) with Doppler model fit (orange curve) and the accompanying convolutional model data for, respectively, $N=1$ and $N=0.0001$ (middle and bottom). See results summarised in Tables 2 and 3. Results are shown on the interval $[-\pi, 3\pi]$ as for the seismogram displays

The excellent azimuthal coverage of the array enables us to obtain reliable estimates of the strike direction, Φ, and distance, L, of rupture propagation in cases where L is sufficiently large (the dependence of the quality of the results on L is discussed further below). To determine ζ we need to also know w_t, the
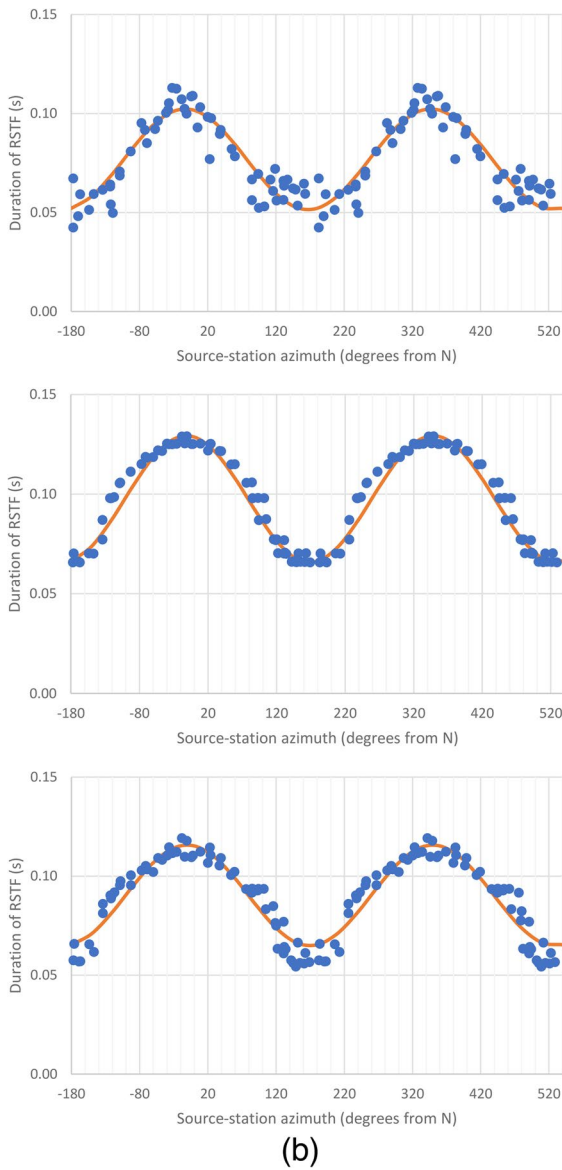
Fig. 6 (continued)

width of the individual pulses in the RSTF. This is difficult to determine accurately as it would require reliable picking of the maxima of the individual constituent pulses which are strongly affected by noise and tuning. The pulse half-width should be no smaller than the rise time of the constituent slip patches but will also depend on other factors, especially the band width of the data and the level of additive noise used in the deconvolution. Park and Ishii (2015) acknowledge the difficulty in determining the rise time and

so neglect it, such that they underestimate ζ. Here we identify approximate bounds on the pulse width. A lower bound is provided by the band limitation of the data: as explained above, a steep anti-alias (high cut) filter from 80 to 100 Hz has been applied. If the bandwidth is $w_f = 80$ Hz, then the minimum pulse width is determined by a frequency-time uncertainty relationship, $w_f w_t \geq 1$, such that $w_t \geq 1/80 = 0.0125$ s. An upper bound is provided by inspection of the data: assuming $\zeta \leq c$ implies $w_t \leq A_1 - A_2$. Inspection of the variation of duration with azimuth for all events gives $w_t \leq 0.03$ s (this can be verified from the data in Tables 2 and 3). $w_t = 0.019$ s is the largest value consistent with the assumption that $\zeta \leq c$, whereas $w_t = 0.017$ s corresponds to $\zeta \leq 0.9c$ as is often observed (see for example Yoshido et al. 2019). $w_t = 0.017$ s is the value that has been chosen to estimate ζ for all cases in Tables 2 and 3.

A quality measure, ξ, for the parameter inversion was defined as

$$\xi = 1 - \Sigma/\Sigma_0 \tag{11}$$

where Σ is the L1 sum of the residuals for the best fit azimuth-dependent Doppler model and Σ_0 is the corresponding sum obtained by comparing the data to the azimuth-independent average, $\Delta t(x_i) = A_1$. Values of ξ range between 1 and 0 and can be used as an attribute to objectively rank the quality of the EGF inversion results. The highest values indicate the results for which the Doppler broadening expression provides the most convincing fit to the picked duration.

Figures 11 and 12 summarise the results of the Doppler fitting process for all event pairs. Notice how the quality metric, ξ, depends on the rupture propagation distance, L, and the average magnitude for the event pair: L determines amplitude of the azimuthal variation of the duration and the average magnitude will largely determine the signal-to-noise ratio of the seismogram data. There is a weak trend to be seen in the plot of ξ against the separation between the catalogue locations of the events comprising a pair: the largest separations correspond to low values of ξ, but for small values of the separation there is a large spread of quality measure values presumably reflecting its dependence on a number of factors (L, differences in source mechanism, etc.). Catalogue location uncertainties will play a part in this assessment—it is
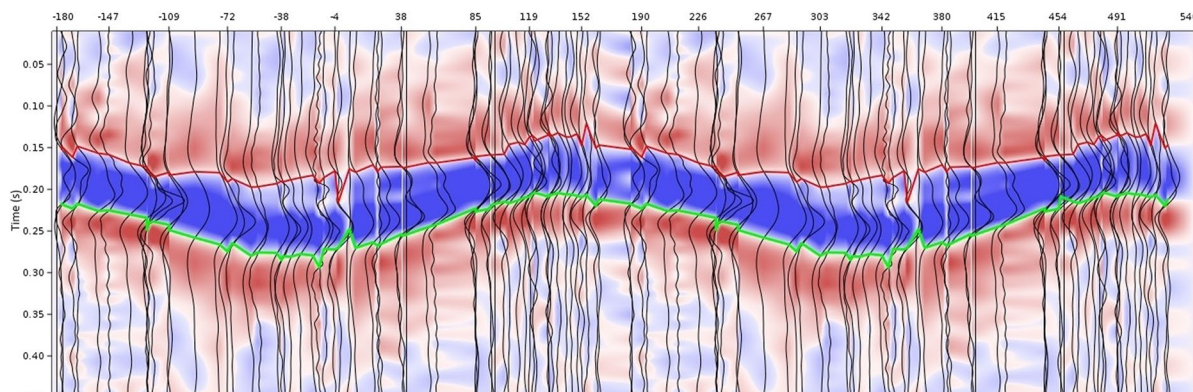
(a)
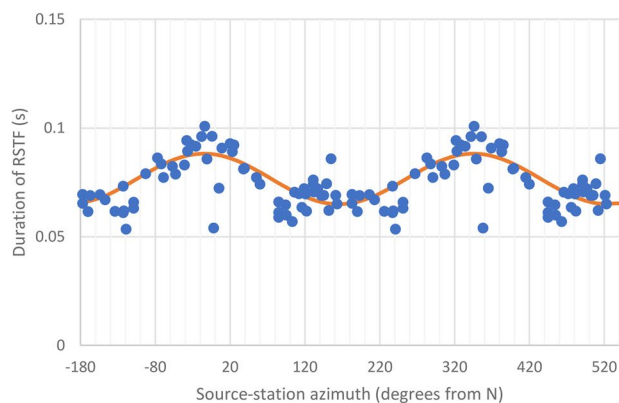
(b)

Fig. 7 **a** EGF trace output for event pair 18019/18039. Compare with the output for event pair 18024/18039 in Fig. 6(a). The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. **b** Picked duration be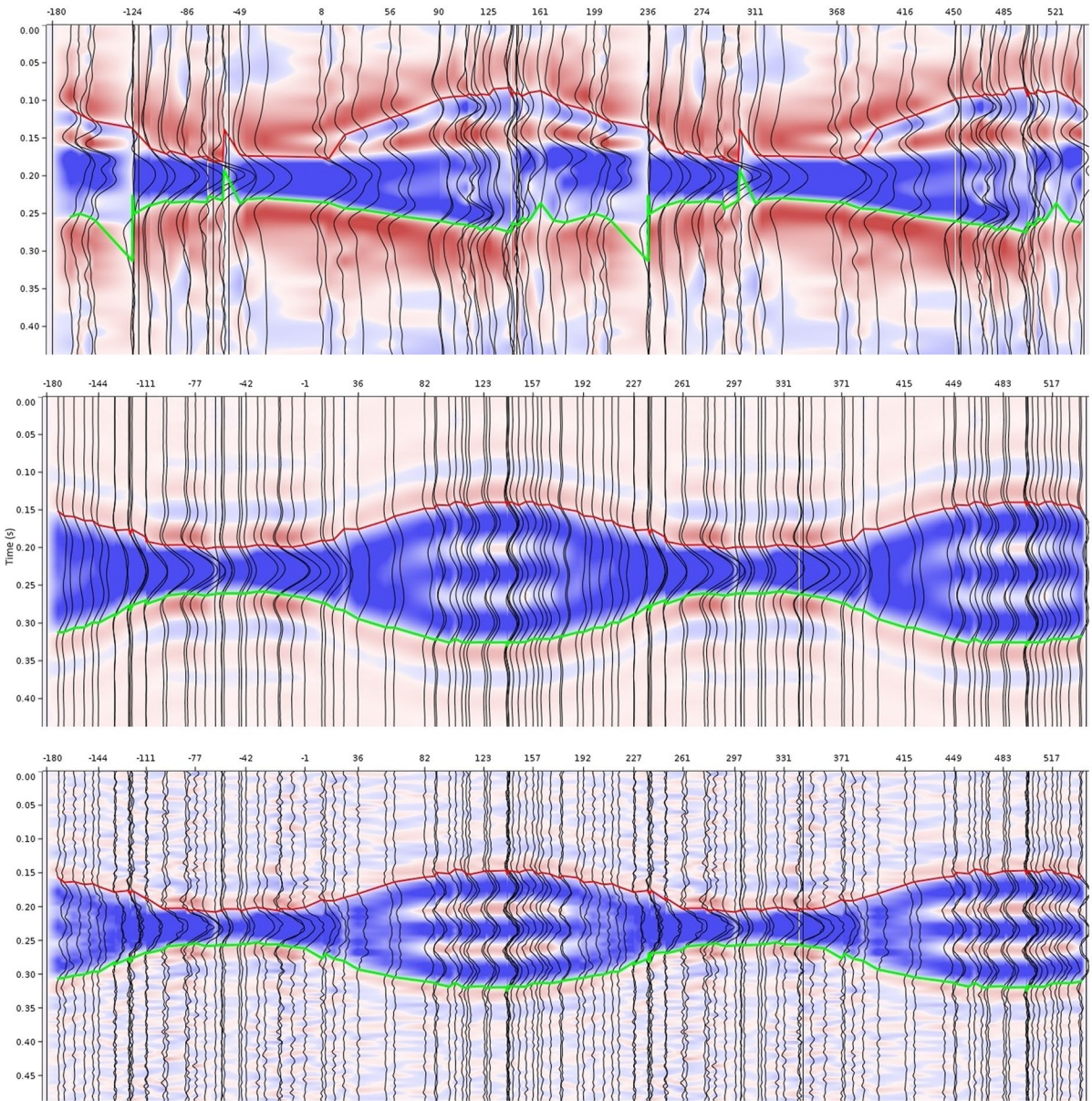tween the zero crossings of the RSTF traces for event pair 18019/18039 (blue data points) with Doppler model fit (orange curve). Compare with the output for event pair 18024/18039 in Fig. 6(b). Results are shown on the interval $[-\pi, 3\pi]$ as for the seismogram displays

unlikely that events with very large physical separations will have high seismogram cross-correlation values. We expect to be able to refine relative location estimates using the azimuthal variation of the RSTF arrival times as explained in Appendix B.

For a subset of the processed event pairs, convolutional model synthetics were generated as described above to demonstrate a reasonable visual match between the synthetic and field data RSTFs and explore to what extent the deconvolution data processing workflow recovers the input parameters. Table 1 gives the input and inverted values of rupture propagation strike and horizontal distance. Inversion results are given for deconvolution noise levels $N = 1$, the chosen default for the field data, and $N = 0.0001$, a

very low noise level which would only serve to stabilise the spectral division. Source offset ranges (Δ) and angles (Ψ) are defined above Eqs. (21) and (22) and were determined by a simple trial and error process to obtain a good visual match to the corresponding real data RSTF. In Appendix B it is shown that the duration is independent of Δ and Ψ. Note how the lower noise value generally gives a better recovery of the input parameters. With a vanishingly small N, all convolutional factors common to both events cancel as shown in Eqs. (4)–(6): this applies to filters applied in processing as well as to propagation and instrumentation effects. On the real data cases we found there is a trade-off between stability of the arrival picking (achieved with higher N) and the good temporal

(a)

Fig. 8 a EGF trace output for event pair 16117/16089 (top) and the accompanying convolutional model synthetic data processed with $N=1$ (middle) and $N=0.0001$ (bottom). In this case a three-phase model (starting and stopping phases and one intermediate phase) was used. The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. **b** Picked duration between the zero crossings of the RSTF traces (top) for event pair 16117/16089 (blue data points) with Doppler model fit (orange curve) and the accompanying convolutional model data for $N=1$ (middle) and $N=0.0001$ (bottom). See results summarised in Tables 2 and 3. Results are shown on the interval $[-\pi, 3\pi]$ as for the seismogram displays

(b)
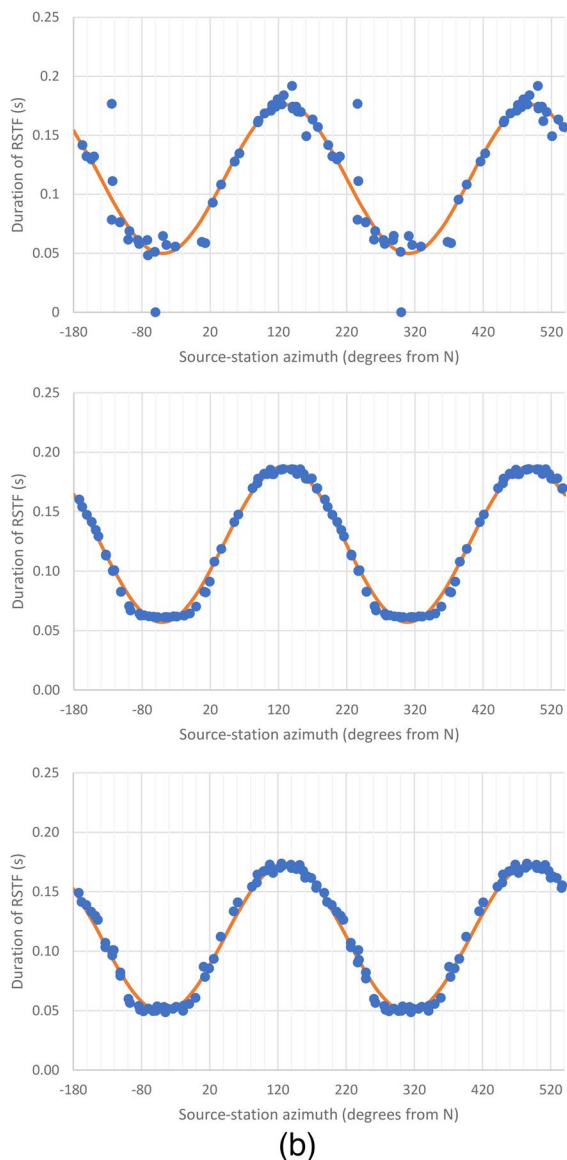
Fig. 8 (continued)

resolution (lower N) needed for accurate recovery of the underlying rupture characteristics. Processing of the convolutional model synthetics allows the impact of the choice of N value to be assessed. Notice also how the duration is dependent on the width of the individual constituent pulses, but the ratio $A_2 = L/c$ which determines the range of the azimuthal variation of the duration is independent of this.

In Fig. 13 we demonstrate consistency between the convolutional model and finite difference simulation for

a simple case of an event pair with the larger rupture a unilaterally propagating two-phase rupture, with the starting and stopping phases each of the same slip patch size as the smaller rupture. The magnitudes in the finite difference simulation are 3.4 and 1.7; source coordinates are {245,493, 597,742, 2950} and {245,585, 597,620, 2950}, respectively; rupture dimensions are 355 m×355 m and 50 m×50 m, respectively. The strike of the rupture propagation vector is $-53°$ with respect to N and the horizontal rupture propagation distance is 305 m. Rupture propagation velocity is $\zeta = 2435$ m/s. The EGF deconvolution workflow and Doppler parameter inversion process give a rupture vector strike direction of $-52°$ and a horizontal rupture propagation distance of 306 m if the effective shear velocity is taken as 2435 m/s.

More sophisticated finite difference simulations of causally connected slip patches stochastically distributed across the fault plane were also run. An example of the RSTF obtained from one such model is shown in Fig. 14. This simulation was for a model of bilateral rupture propagation initiated at the middle of the rupture surface. Parent and child horizontal rupture patch lengths are again 355 m and 50 m, respectively, such that $L=152.5$ m; the strike of the rupture propagation vector is $-53°$ from N. The rupture velocity varies stochastically across the rupture surface with values distributed about $\zeta = 2400$ m/s. The subsurface velocity model is unchanged. Due to the more complex nature of the composite rupture propagation and the resulting stochastic distribution of slip across the rupture plane, the RSTF shows greater variability than the simple convolutional model synthetics. In Fig. 14 the picked duration is compared with the simple bilateral model of Eq. (25) with the same input parameter values and arguably shows the same approximate periodicity. A tentative conclusion is that this extra richness of modelled rupture behaviour is not required to explain the clearest cases of Doppler broadening shown here (that is those cases with the highest values of ξ).

7 Interpretation and discussion of results

For the highest quality results, the azimuthal Doppler broadening of the source time function is clear to see on the seismic panels, as shown in Figs. 6(a), 7(a), 8(a), 9(a) and 10(a). In most cases we also see the superposed undulation of the arrival times with
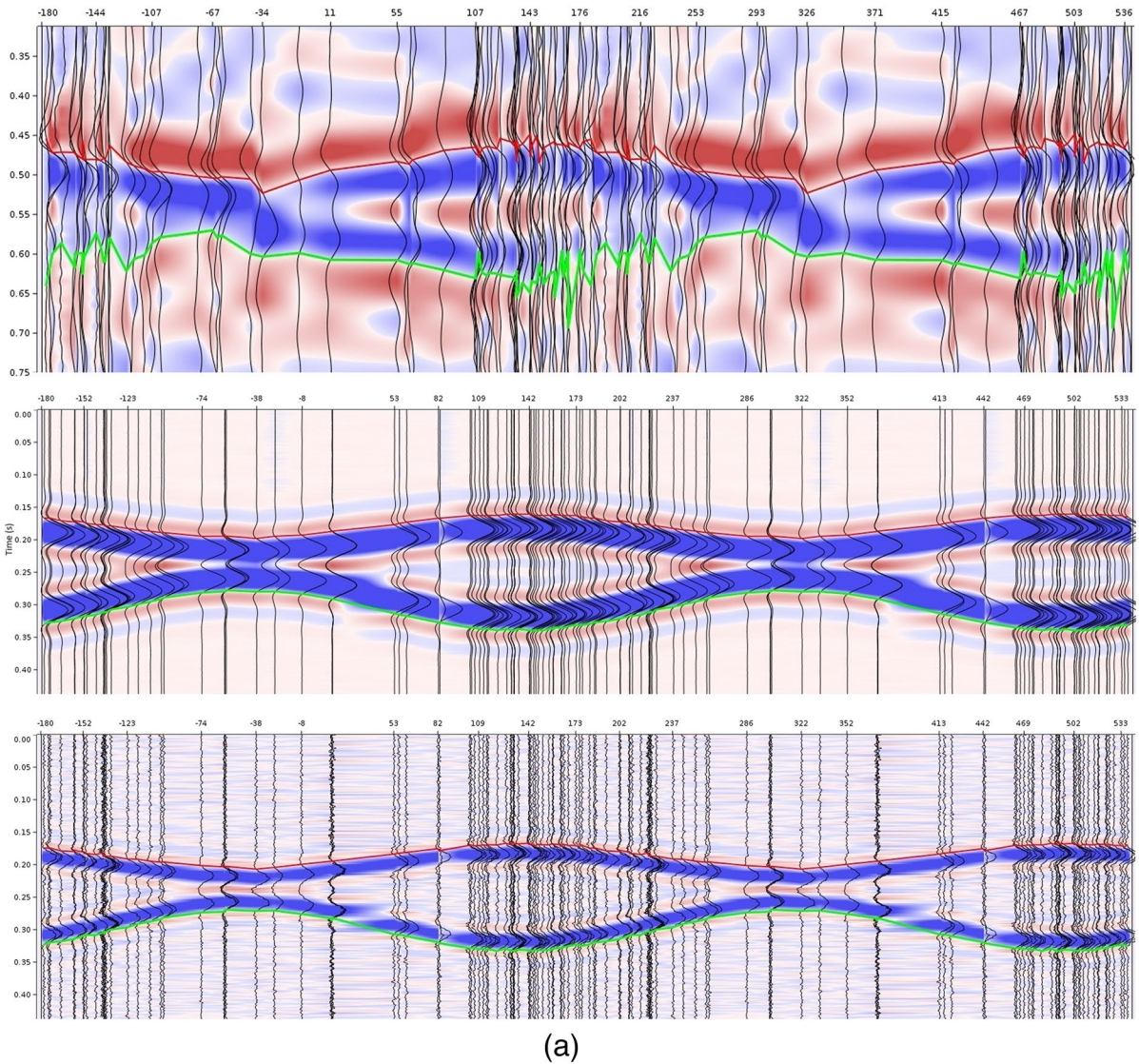
(a)

Fig. 9 **a** EGF trace output for event pair 17124/18002 (top) and the accompanying convolutional model synthetic data with $N=1$ (middle) and $N=0.0001$ (bottom). The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. **b** Picked duration between the zero crossings of the RSTF traces (top) for event pair 17124/18002 (blue data points) with Doppler model fit (orange curve) and the accompanying convolutional model data for $N=1$ (middle) and $N=0.0001$ (bottom). See results summarised in Tables 2 and 3. Results are shown on the interval $[-\pi, 3\pi]$ as for the seismogram displays

azimuth, characteristic of an offset between the two events. The propagation directions correspond very well with the details of the underlying fault network (Fig. 15) obtained by interpreting the 3D data and refined using an automatic discontinuity detection algorithm applied to seismic attribute volumes (Kortekaas and Jaarsma 2017). For the 10 highest quality event pair results, characterised by $\xi \geq 0.5$, only 15062/15065 does not show a convincing alignment

of the rupture vector with a nearby feature of the fault map. This, the most westerly of the event pairs on the first subplot of Fig. 15, has a poor azimuthal coverage due to being close to the edge of the monitoring array. The correspondence with the mapped fault directions worsens as we go to lower values of the quality measure ξ, as shown in the other subplots of Fig. 15.

Where events form correlated clusters of more than two events, it is interesting to observe the
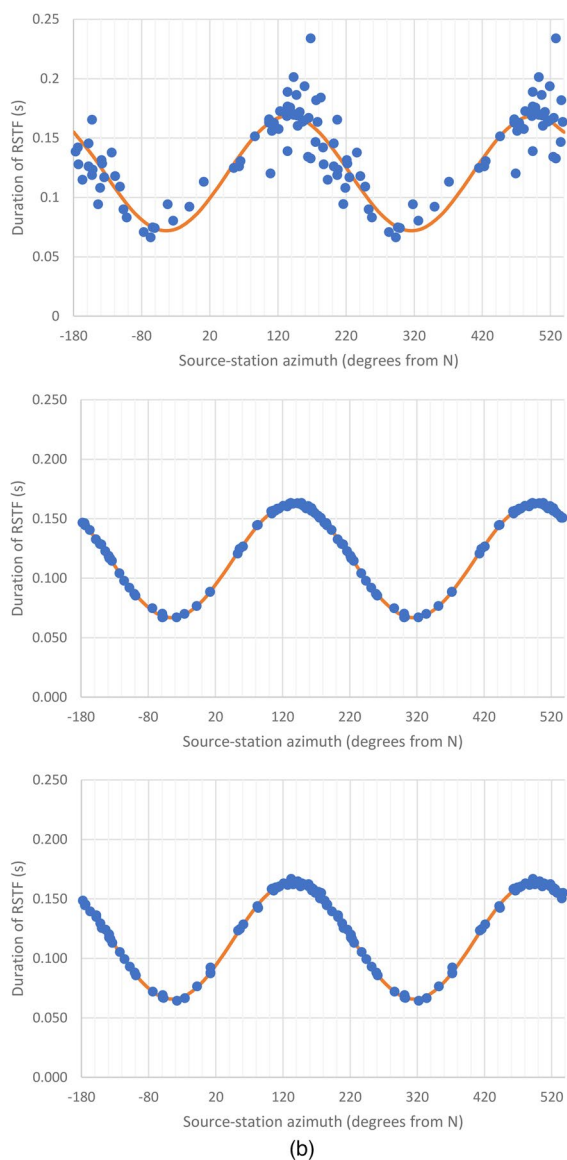
Fig. 9 (continued)

occurrence of opposing rupture propagation directions. See especially the cluster with event IDs 18039/18024/18019/18096. Tables 2 and 3 shows that these results are however not contradictory—where there is more than one analysis for a given parent event, the rupture azimuths and lengths obtained are consistent. We assume that correlated events will have rupture propagation vectors parallel or anti-parallel to each other, aligning with a common host fault. It can be seen that adding π to the rupture strike Φ is equivalent to changing the sign of the corresponding

rupture distance L in Eq. (29). Having aligned all rupture directions in a cluster in this way, the set of coefficients of the cosine terms in Eq. (29) from all event pairs in the cluster can be inverted for the individual event rupture distances. In the case of the event cluster 18039/18024/18019/18096, the problem is over-determined such that average values can be obtained. These cluster-derived values, as listed in Table 4 in the Appendix, are then preferred to the raw summed values in Tables 2 and 3. Figures 11, 12 and 15 however show compilations of the Doppler inversion results for the event pairs and so do not include these cluster refinements for individual rupture propagation lengths.

From a geomechanical perspective, the observed range of horizontal rupture propagation distances, L, helps constrain estimates of the dimensions of the ruptures. Using the value of L and the catalogue magnitudes, we estimate the dip direction extent of the rupture, L_D, using the usual simple relationship between rupture area, stress drop, and seismic moment (Beresnev 2001). We take the rupture length in the horizontal direction to be the sum of the slip patch length of the smaller event (assumed to be square) and the EGF-derived rupture propagation distance. We then estimate the rupture extent in the dip direction as the length required to give the catalogue (moment) magnitude of the larger event for an assumed value of stress drop. To do this we take the effective relationship for the shear modulus in terms of the average normal faulting slip displacement U and stress drop $\Delta\sigma$ derived from seismological observations as $\mu \approx L_D \Delta\sigma / U$ (Beresnev 2001), thereby assuming that the shape-dependent factor in the stress–strain relationship is approximately 1 (Kanamori and Anderson 1975; Noda et al. 2013). The first subplot in Fig. 12 shows the estimated dip-direction rupture extent against the magnitude of the larger event for an assumed stress drop of 7 MPa (Edwards et al. 2019). For all but the largest magnitude event (the $M = 3.4$ event of 8th January 2018), the estimated dip-direction rupture extent is sufficiently small to be contained wholly within the reservoir which is some 200 m thick over most of the field (De Jager and Visser 2017). This largest magnitude event (18002) has been located with full waveform inversion (Willacy et al. 2018) its location placing it on a mapped fault with fault plane dip varying between 67 and 84°. Of the two focal mechanism solutions for the slip vector, the one
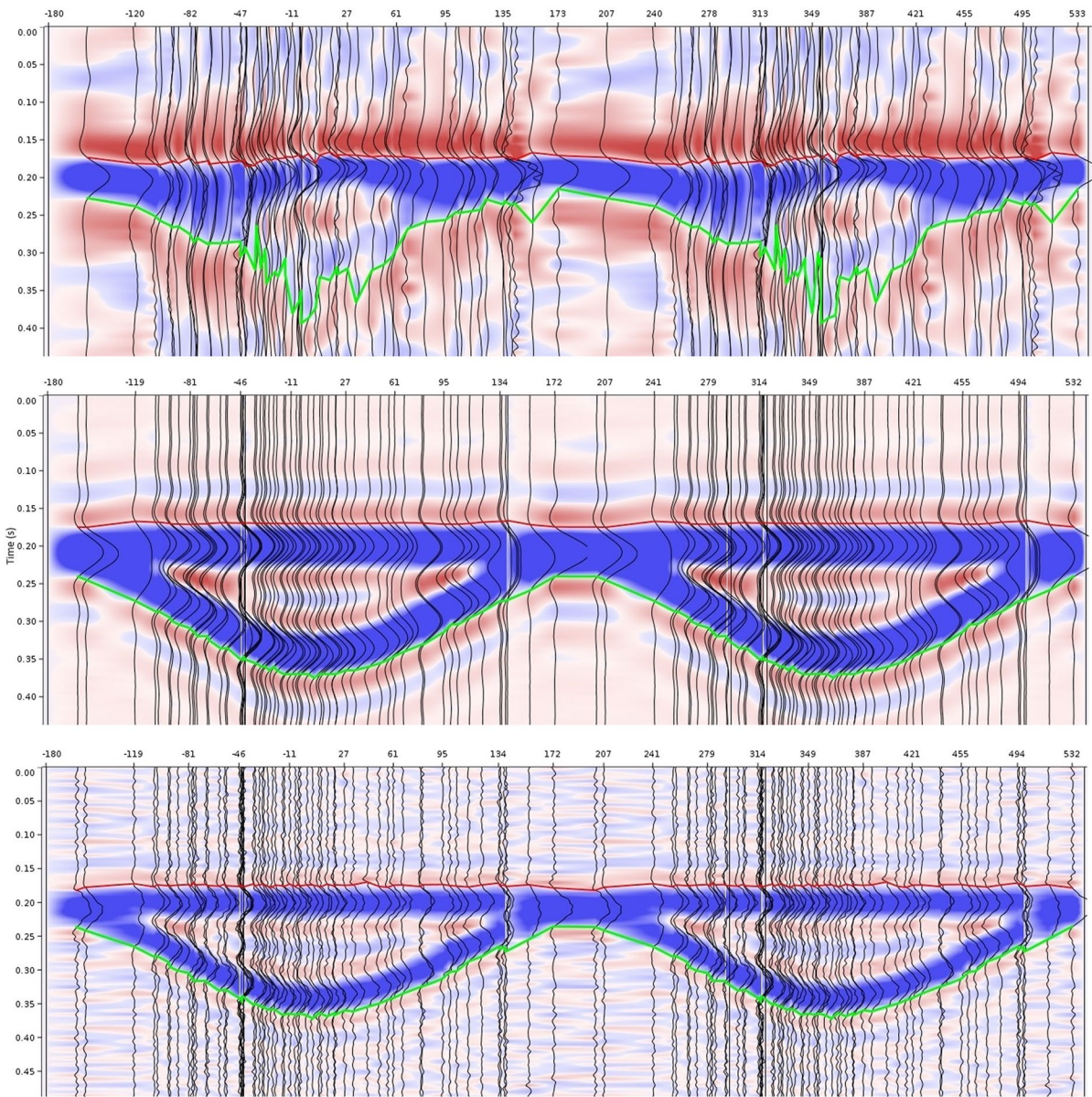
(a)

Fig. 10 **a** EGF trace output for event pair 17064/17052 (top) and the accompanying convolutional model synthetic data with $N=1$ (middle) and $N=0.0001$ (bottom). The traces have been duplicated and displayed on the interval $[-\pi, 3\pi]$ to ensure that periodic effects are not obscured. **b** Picked duration between the zero crossings of the RSTF traces (top) for event pair 17064/17052 (blue data points) with Doppler model fit (orange curve) and the accompanying convolutional model data for $N=1$ (middle) and $N=0.0001$ (bottom). See results summarised in Tables 2 and 3. Results are shown on the interval $[-\pi, 3\pi]$ as for the seismogram displays

which is consistent with the mapped fault has a dip of 73°. From Table 3, the minimum estimate of the dip direction rupture extent for the event is 380 m. With a dip angle of 67° or greater, this gives a vertical

projection of at least 350 m. This contrasts with the mapped Rotliegend reservoir thickness at the event location of about 275 m and the throw of about 40 m, totalling 315 m. These estimates of rupture dimensions
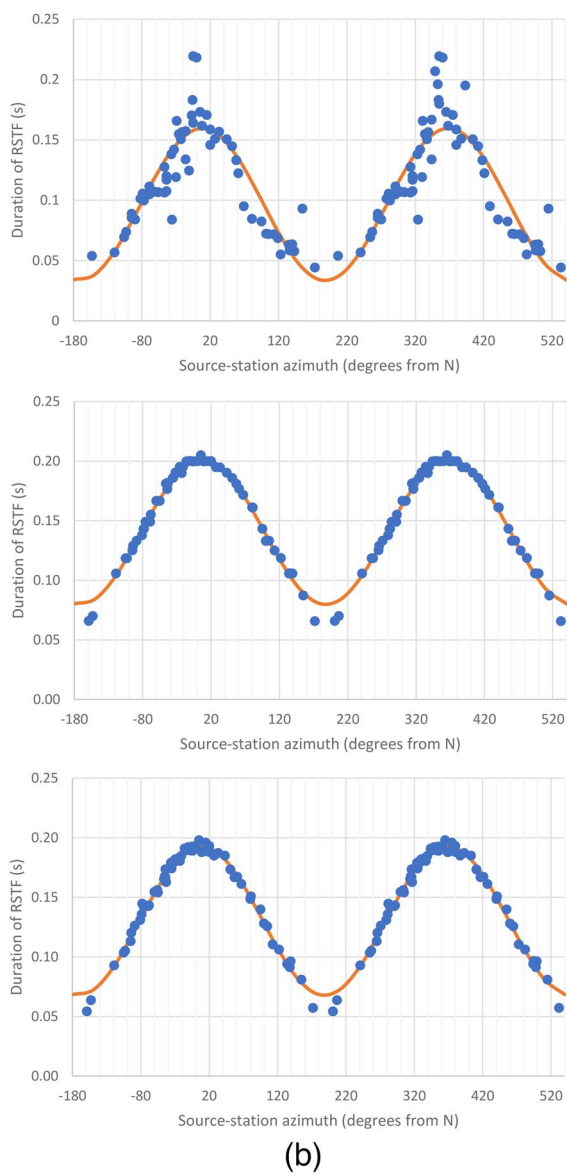
(b)

Fig. 10 (continued)

have been made using an assumed value of stress drop taken from the work by Edwards et al. (2019), which is based on a subset of events from the Groningen catalogue: there is considerable scatter in those stress drops but we find that repeating the exercise with their lower and upper logic tree branch values of 5 MPa and 10 MPa does not change the general conclusion regarding containment of the events in the reservoir.

An alternative approach may be to use an empirical scaling relationship between fault length and width,

such as presented by Leonard (2010), obtained by fitting slip models to a large number of seismological datasets (see also Leonard 2012). Leonard (2010) shows a good match to his compendium of normal and reverse dip-slip fault data with $L_D = CL_S^{2/3}$ for $C = 1.7$. It can however be shown that estimating L_D from this simple relationship requires wildly varying stress drops if the relationship between moment, rupture area, and stress drop is to be preserved. Moreover, Leonard (2010) does not consider data for faults shorter than 2 km raising the question of whether the relationships obtained are representative of reservoir-scale seismicity. On the other hand, the stress drops presented by Edwards et al. (2019) are estimated from Groningen-induced earthquake data and so should be applicable.

The EGF relative source time functions are (in principle) free from the amplitude effects of spreading and absorption thanks to the deconvolution and are therefore well suited to the determination of corner frequencies and related quantities. Beresnev (2002) shows that, for a source time function which radiates a frequency-squared far-field spectrum, the corner frequency f_c can be related to the maximum slip velocity, without further assumptions, as follows (where e is the base of natural logarithms, A is the rupture area, and M_0 the scalar seismic moment):

$$v_{max} = \frac{2\pi f_c U}{e} = \frac{2\pi f_c M_0}{e\mu A}. \tag{12}$$

For a finite rupture such as we consider here comprised of multiple slip patches, causally linked, we should not assume the simple source time function given by Beresnev (2002). Nevertheless, the spatially averaged slip displacement divided by the time over which the slip develops should be an approximate average slip velocity, $\bar{v}$. The measured duration of the RSTF is the sum of the finite duration of the constituent pulses after deconvolution and the rupture propagation time; this will span the processes of initiation, propagation, and arrest such that, $\bar{v} \geq U/\Delta t$. Again using $\mu \approx L_D\Delta\sigma/U$, we estimate a lower bound on the slip velocity from the RSTF duration (the value perpendicular to the rupture vector has been used):

$$\bar{v} \geq \frac{\Delta\sigma L_D}{\mu\Delta t}. \tag{13}$$

Figure 12 illustrates the range of slip and rupture velocity values obtained, assuming $w_t = 0.017$ s and
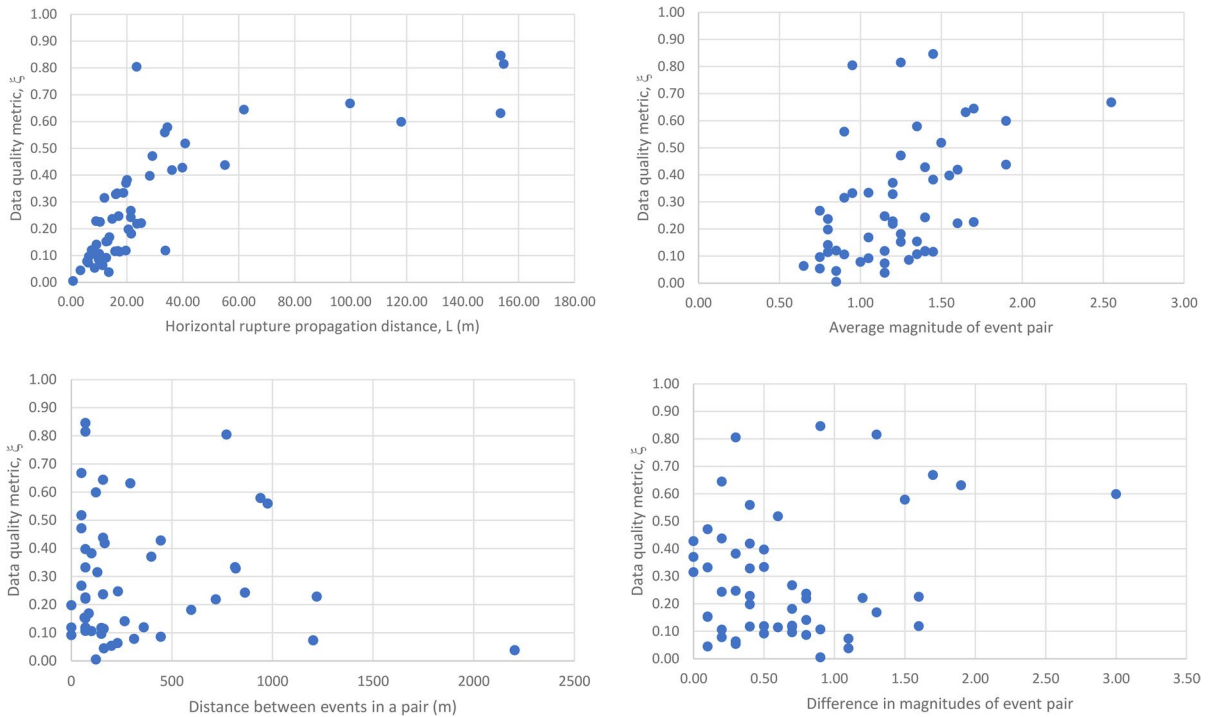
Fig. 11 Plots summarising the quality of the inversion of the picked RSTF durations for the Doppler broadening model parameters. In these plots the data quality metric, ξ, on the vertical axis is plotted as a function of the horizontal rupture propagation distance, L (top left), average magnitude of the event pairs (top right), distance between child and parent in each event pair (bottom left), and the difference in magnitude of the event pairs (bottom right). Notice how the quality metric, ξ, is strongly correlated with the rupture propagation length, L, and (to a lesser degree) correlated with the average magnitude; there is a much less clear correlation with event separation. The difference in magnitudes within the pair, normally taken as a discriminant when selecting event pairs for EGF analysis, is not a determinant of the quality of the sinusoidal fitting process

$\Delta\sigma = 7$ MPa. Note how rupture velocity increases with horizontal rupture distance—the higher rupture propagation velocities are only reached on the longer ruptures. Weng and Ampuero (2019) have presented a model of bounded rupture propagation with an inertial term giving rise to acceleration of the rupture, which could explain this behaviour. Slip velocities, in line with the estimate given by Beresnev (2001), are generally about 3 orders of magnitude smaller than the rupture propagation velocity, ζ. Recent dynamic rupture simulations by Buijze et al. (2019) and laboratory measurements by Hunfeld et al. (2017) also support the range of slip velocity values given here but emphasise the sensitivity of the slip velocity to the detailed variation of fault friction properties.

The kinematic model seismograms have been used as a guide to our interpretation of the data, helping to visualise the effect of a range of scenarios,

demonstrating qualitative as well as quantitative correspondence with the real data by comparison of input and inverted parameter values. Figures 6(a), 8(a), 9 (a) and 10(a) show comparison of RSTF output generated for real and synthetic data for some selected event pairs. It should be emphasised that we have in all cases fitted a model of unilateral rupture propagation to the data. In the most striking cases (event pairs with the highest values of ξ), the sinusoidal variation of the duration with a period of 2π seems to be inescapable. None of the event pairs processed show clear evidence for a superposed second sinusoidal variation (see Eq. (26)) characteristic of bilateral propagation (event pairs 18019/18024 and 18066/18065 are possible exceptions). Several authors, for example Folesky et al. (2016), have noted a predominance of unilateral rupture propagation. King and Nabelek (1985) argue that ruptures tend to run between bends and junctions
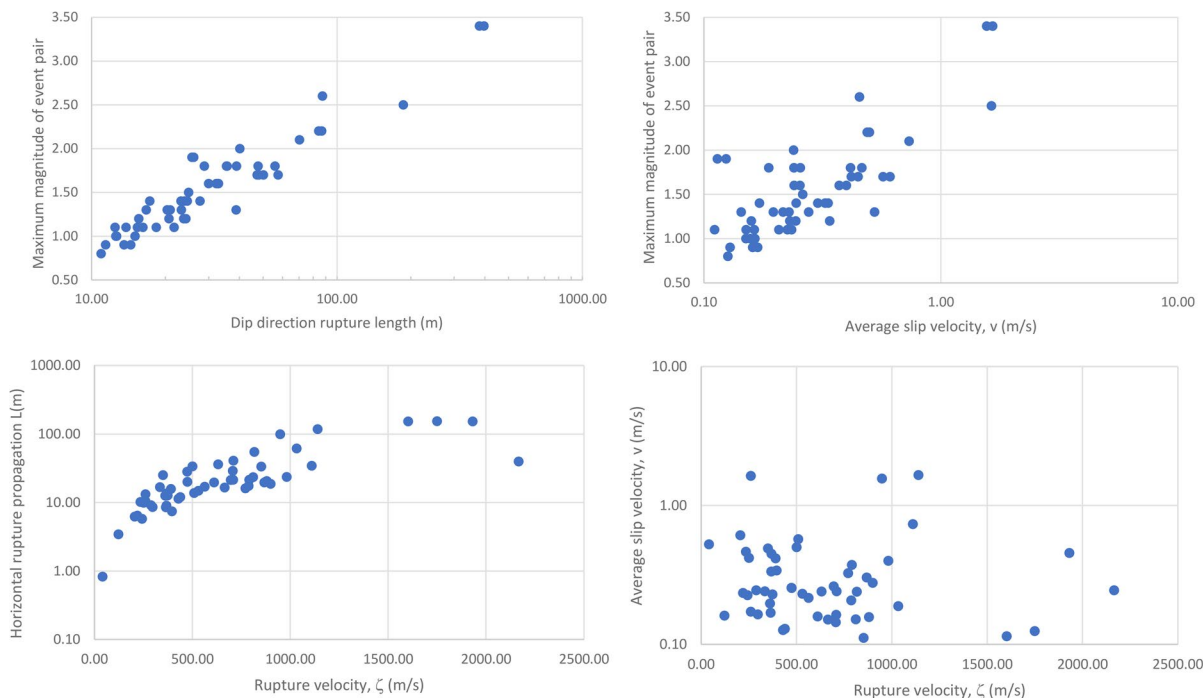
Fig. 12 Plots summarising the results of the inversion of the picked RSTF durations for the Doppler broadening model parameters and related rupture parameter estimates. In the top row, maximum magnitude of the pair of events is plotted on the vertical axis against the estimated values of (left) dip direc- tion rupture length, L_D, and (right) average slip velocity, $\bar{v}$. In the bottom row rupture propagation velocity, ζ, is plotted on the horizontal axis against (left) the horizontal rupture propa- gation distance, L, and (right) the average slip velocity, $\bar{v}$

Table 1 Summary of key model input parameters and inversion results for deconvolution noise levels $N=1$ and $N=0.0001$, for synthetic seismogram data corresponding to four representative event pairs

		Input parameter values				Inversion results $N=1$		Inversion results $N=0.0001$	
Event IDs		Φ	L	Ψ	Δ	Φ	L	Φ	L
		(° from N)	(m)	(° from N)	(m)	(° from N)	(m)	(° from N)	(m)
Child	Parent								
18024	18039	167	62	10	75	168	77	169	62
16117	16089	−50	154	−50	71	−51	160	−49	152
17124	18002	−45	118	−45	44	−45	117	−45	120
17064	17052	188	154	-	0	188	148	188	152

in a fault network such that individual events tend to be characterised by unilateral rupture propagation. Referring to Fig. 15 it can be seen that many of the plotted rupture propagation vectors do initiate on a recognisable fault junction. The Haskell finite rupture model may then be a good approximate description, although it is accepted that it violates the physical requirements of continuity at the rupture boundary and finite rupture propagation velocity in the direction perpendicular to the length. The model of Imanishi and Takeo (2002) suggests an alternative explanation for the observation of apparent unilateral propagation which avoids the physical difficulties of the Haskell model. At higher frequencies, seismograms will be dominated by radiation from accelerating and decelerating phases of the rupture—the starting and stopping phases. Imanishi and Takeo (2002) show how a rupture propagates from an initiation point and then radiates strongly from stopping phases where the rupture front is tangent to the eventual boundary of the
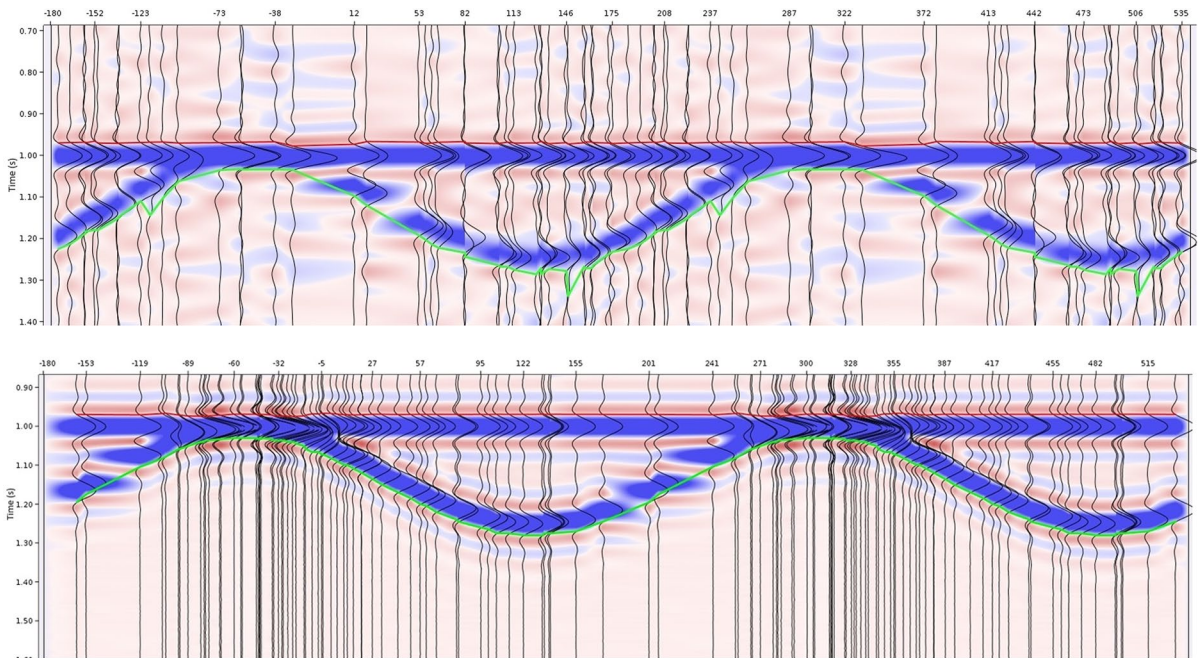
Fig. 13 EGF trace output for a synthetic event pair with magnitudes 3.4 and 1.7, generated using the finite difference approach (upper plot) described here. Rupture dimensions are 355 m × 355 m and 50 m × 50 m and the strike of the rupture propagation vector is −53° with respect to N. Rupture propagation velocity is $\zeta = 2435$ m/s. Considering the large rupture as consisting of a starting and stopping phase, each with the same slip patch size as the small rupture, this implies a horizontal rupture distance of 305 m. The EGF deconvolution workflow and Doppler parameter inversion process give a rupture vector strike direction of −52° and a horizontal rupture propagation distance of 306 m if the effective shear velocity is taken as 2435 m/s. The lower plot shows the comparison with the convolutional model with the same input model parameters, both processed with $N = 1$

rupture. This is a potential means of linking a natural description of rupture propagation with the RSTFs we see but would be strongly dependent on the specifics of the geometry of the fault boundary relative to the rupture initiation point. Nevertheless, as a conceptual model, it could explain why the very simple model we use here fits the data well. This however comes at the price of obscuring the connection between the observed horizontal rupture propagation length and the rupture extent. If the Imanishi and Takeo (2002) model were adopted as an interpretational paradigm, the observed value of L would have to be seen as a lower limit on the horizontal rupture propagation distance. Horizontal rupture propagation over greater distances could occur but without additional stopping phases being detectable by our network of sensors. Strike- and dip-direction extents calculated as above would then be seen as lower and upper bounds, respectively. Our observations would then not conflict with the possibility of all ruptures being contained

within the reservoir but, equally, it would not be possible to rule out propagation of the largest event into the Carboniferous interval below the Rotliegend reservoir.

Savage (1965) argues that, for unilateral rupture propagation, the azimuthal variation of the amplitude of first motion is the inverse of the expression for Doppler broadening of the duration of the source time function. Folesky et al. (2016) use this result to interpret seismic data from the Basel geothermal project. Despite the excellent azimuthal coverage of the Groningen array, we were unable to invert measured amplitudes of the RSTFs for the parameters $\{L, \zeta, \Phi\}$ with any degree of confidence. Noise, differences in azimuths of the focal mechanisms of the paired events, and loss of true relative amplitude in the deconvolution workflow are expected to disrupt the underlying azimuthal dependence of measured amplitudes of the RSTF. The duration, measured between zero crossings, is however
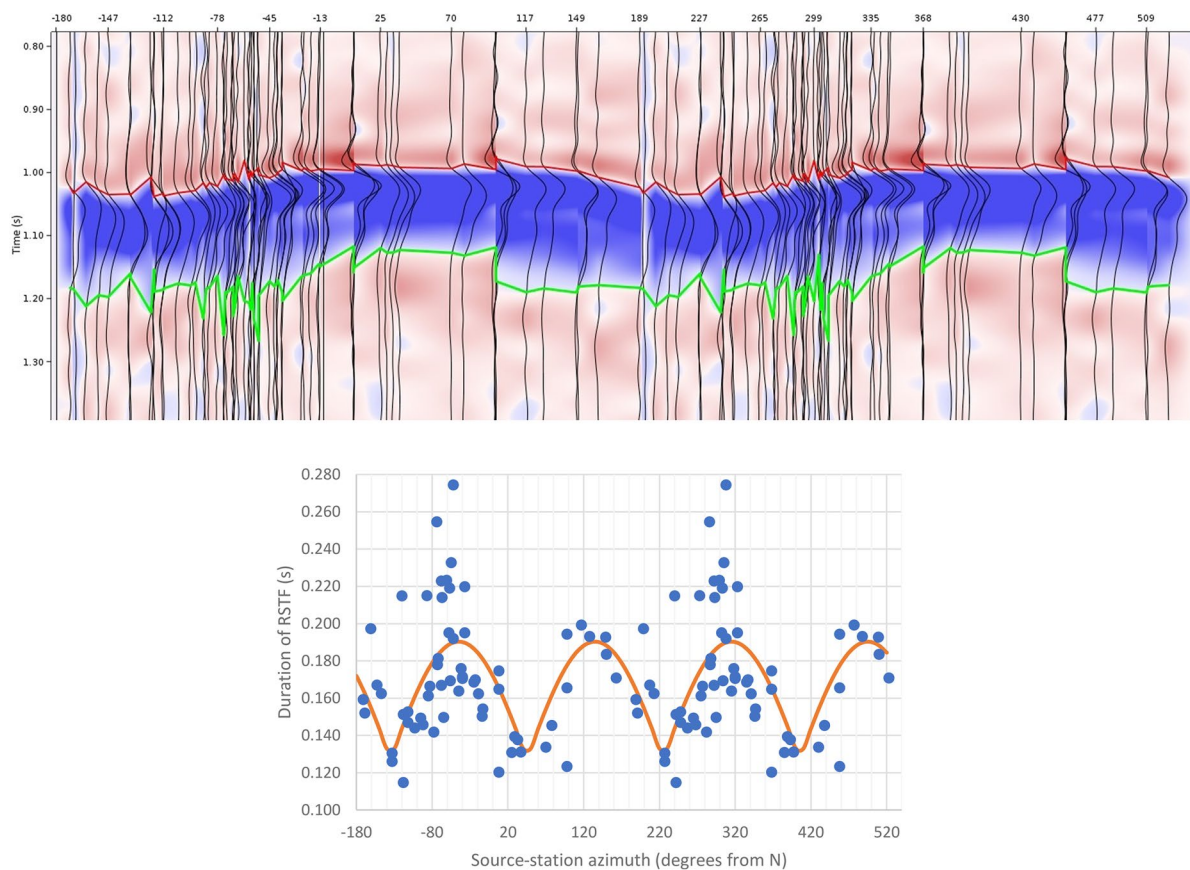
Fig. 14 The upper plot shows the EGF trace output for a finite difference simulation of horizontal bilateral rupture propagation. The lower plot shows the picked durations as a function of azimuth and the fit to the bilateral propagation model of Eq. (25) obtained with parameters $L = 153$ m, $c = 2435$ m/s, $\Phi = 135°$ from N in reasonable agreement with the model input parameters detailed in the text. Comparing the bi-lateral model in Eq. (25) with the data according to Eq. (11) gives $\xi = 0.15$ for this choice of parameters (the azimuth independent mean duration is $L/\zeta + w_t + 2L/\pi c$ for the bilateral model). In comparing the model to the data, values of duration greater than 0.20 s have been rejected as outliers originating from the regions of unstable picks visible on the trace display

largely insensitive to these effects. Abercrombie (2015) presents an analysis of the factors affecting source parameter estimation from spectral analysis of RSTFs and concludes that high-quality EGFs, with only minimal offsets between parent and child events, are needed. Abercrombie et al. (2017) show robust results from analysis of the RSTF duration determined for pairs of azimuths by cross-correlation with a modelled stretch applied.

Ameri et al. (2020) selected the 21 events in the Groningen catalogue with $M_L \geq 2.0$ for a spectral directivity analysis, not using the EGF method. Of these only five feature in our subset of highly correlated event pairs (Table 2) and only for two of these five common events were directivity

effects observed by Ameri et al. (2020). For these two events (20170527_152900 (ID 17052) and 20180108_140052 (ID 18002)), their results are in fair agreement with ours: Ameri et al. (2020) find (respectively) $\Phi = 203°$ and $283°$ compared with our results of $\Phi = 188°$ and $305°$. For two of the other three events, Ameri et al. (2020) observed no directivity effect; in our analyses these events belong to pairs which were assigned low values of the quality parameter, $\xi \leq 0.23$. The remaining common event is interesting. Ameri et al. (2020) put it in an "unknown" category of events for which "… the limited number of records does not allow to identify whether rupture directivity occurred…". In our analysis, this is the anomalous event in the west of the field
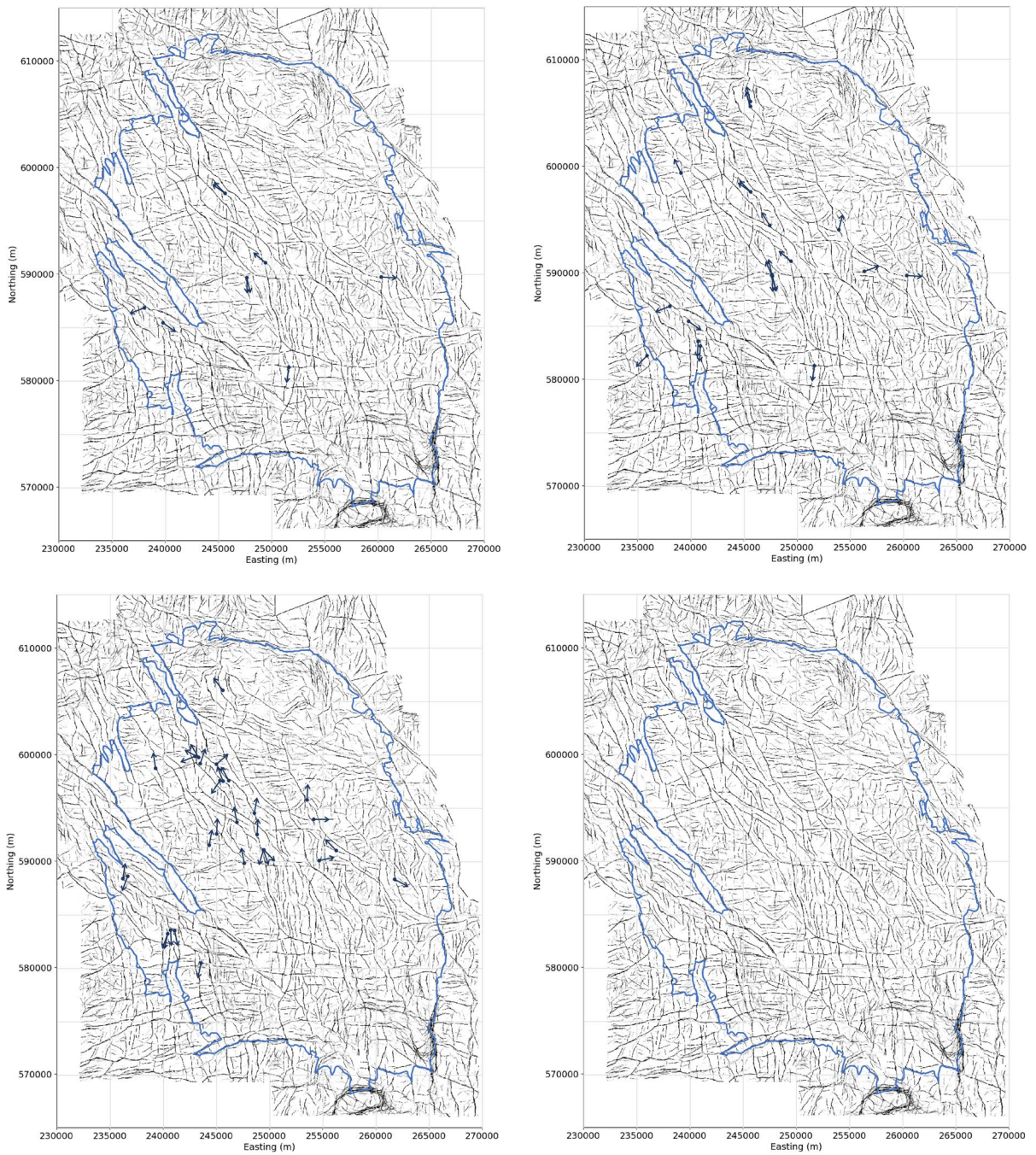
Fig. 15 Scaled horizontal rupture vectors overlain on the fault map. In each of these plots, the base map shows the detailed fault interpretation generated by Kortekaas and Jaarsma (2017). The outline of the gas field is shown as a blue contour. The vectors give only the horizontal rupture propagation direction (vectors are drawn with a constant length), obtained from the inversion of the picked durations. The top left plot shows the 10 highest quality event pairs for which $\xi \geq 0.5$; top right, the 22 event pairs with $\xi \geq 0.25$; bottom left, all 31 remaining event pairs with $0.25 > \xi \geq 0$; bottom right, the base fault map and field outline without events
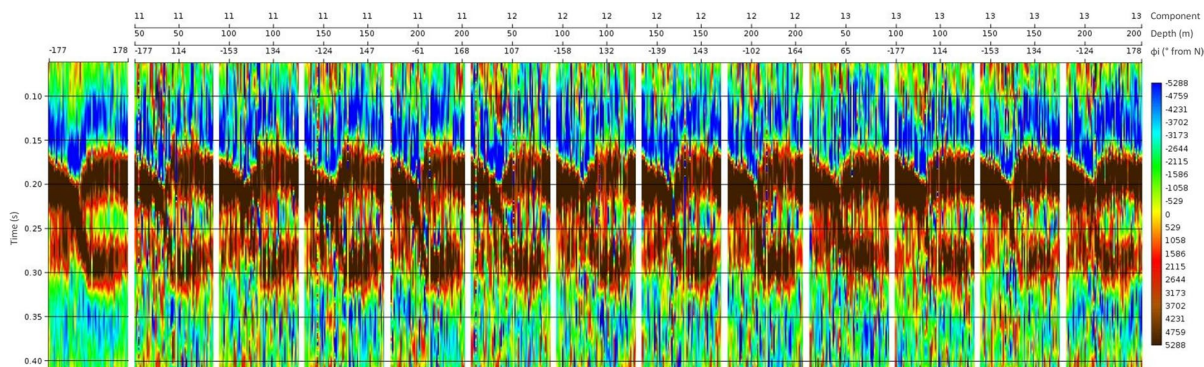
Fig. 16 Panels showing summed RSTFs (left-most panel) and RSTFs for the 12 individual geophones at each station location for event pair 17123/18002. The trace identifiers give the component number, geophone depth, and station azimuth. Each panel shows the range of station azimuths on the interval −180° to +180° with the groups of blank trace slots indicating the break between the data from neighbouring geophones. Wiggle traces are not shown; the colour scale indicates trace amplitude

(top left panel of Fig. 15) which has a high value of the quality parameter, $\xi = 0.58$, but a rupture propagation strike which does not align with a feature on the fault map. As proposed above, and consistent with the categorisation by Ameri et al. (2020), this is we believe due to poor azimuthal coverage of the event location. Overall, we see at least qualitative agreement with the work of Ameri et al. (2020) although the overlap of the subsets of data analysed is limited.

8 Conclusions

We have applied an empirical Green's function deconvolution workflow to 53 pairs of correlated events in the Groningen-induced earthquake catalogue. Measurements of the azimuthal variation of the duration of the relative source time functions (the Doppler effect) have been inverted to give the horizontal rupture propagation azimuth and distance. A quality measure, derived from the L1-norm of the residuals of the data compared to the Doppler model, enables us to rank the individual results. We have interpreted our results in the light of a simple convolutional model of composite rupture and show that finite difference simulations generate results which are consistent with this. The finite difference kinematic rupture simulations can be used to explore the characteristics of a wider range of more sophisticated models of causally connected slip patches distributed across a fault plane.

We see clear evidence of the rupture propagation for many of the event pairs analysed. For the highest quality results, we observe convincing alignment of the propagation vectors with the detailed underlying fault map obtained from the 3D seismic data volume covering the field. The relative source time functions are consistent with a simple kinematic model of unilateral rupture propagation between dominant starting and stopping phases. Event pairs 16117/16089 and 16118/16089 have a distinct intermediate phase in the RSTF which can be best explained by a three-phase parent event.

For all but one of the parent events analysed (the event of 8th January 2018), the estimated dip direction rupture size is sufficiently small to be contained within the reservoir. Our estimates of slip velocities (of the order of 0.1 to 1 m/s) and rupture velocities (up to about 0.9 of the shear velocity) are consistent with other published studies but depend on some assumptions. We have obtained useful results from the Doppler analysis by relaxing the usual restrictions that EGF event pairs be precisely co-located and with a magnitude difference of at least 1. We admit all pairs showing a high degree of seismogram correlation and assume that in the few cases where such pairs have a large separation between catalogue locations, there is scope for relative location refinement.

We find that the duration of the relative source time function, between picked zero crossings, is a particularly robust attribute which is independent of the offset between the pair of events. We show how the azimuthal variation of the arrival times of the

RSTF's zero crossings could be inverted to estimate the offset between the events, emphasising the connection between the EGF deconvolution workflow and other correlation-based analyses such as double differencing and full waveform inversion. Additive noise is required to stabilise the deconvolution, but this also means that the removal of the path effects by the deconvolution is then no longer exact: this limits the accuracy with which input parameters can be recovered from synthetic data. On the other hand, a high level of additive noise allows useful approximate results to be generated for a broader class of event pairs. An intermediate value $N \approx 1$ seems to be a *Goldilocks* value combining these benefits.

The dataset analysed is available in the public domain and, given its excellent azimuthal coverage of the source region, would reward further processing efforts. Our results leave open questions, some of which could be addressed by systematically exploring the effect of varying the parameters of the deconvolution workflow and potentially embedding the method in a broader full waveform processing and inversion scheme. This would require greater automation of the method which was beyond the scope of our study. We have interpreted our results in the light of a simplistic unilateral rupture model. Although this model produces synthetic data displaying the main characteristics seen in the field data, it is purely kinematic and gives no insight into the rupture dynamics. Extending the 2D dynamic rupture simulations of Buijze et al. (2019) into 3D would improve our understanding of the relevant mechanisms. Although we found time domain amplitude variations with azimuth much less stable than the corresponding duration variations, frequency domain analysis using the directivity function method (Udías et al. 2014) may be sufficiently robust to provide essentially independent estimates of the rupture propagation parameters.

Author contributions   S.O. set-up and ran most of the data processing, developed the convolutional rupture models and drafted the manuscript.

J.S. played a significant role in adapting the methodology for this application and interpreting the results, and also contributed to writing the manuscript.

B.Z. developed and carried out the finite difference kinematic rupture simulations, and contributed to writing the manuscript.

T.P. initially applied the Empirical Green's Function method on a subset of the correlated events thereby demonstrating that the data is amenable to this analysis, and contributed to reviewing and editing the manuscript.

E.v.D. set-up and ran the search for the subset of correlated events, supplied the seismogram data in a suitable form and contributed to reviewing and editing the manuscript.

Declarations

Competing interests   The authors declare no competing interests.

Appendix A: catalogue of event pairs and summary of results of analysis

Tables 2 and 3 catalogue all event pairs processed with the EGF deconvolution workflow and the results derived. "Child" and "parent" refer to smaller and larger events, respectively. Table 4 gives the refinements of rupture propagation lengths for correlated clusters of three or more events.

Appendix B: angle dependent durations of RSTFs for simple kinematic models

Here we give derivations of expressions for the duration of the RSTF as a function of the azimuthal angle for simple model scenarios as illustrated in Fig. 3(a) and (b). A number of the below expressions can be found in Savage (1965), Li et al. (1995b), and Cesca et al. (2011). The objective here is to provide a self-contained presentation with a consistent notation which clearly states the assumptions being made.

Table 2 Event pair catalogue. Dates and times are given in the format YYYYMMDD, HHMMSS; time is truncated to remove the decimal seconds. Event locations and magnitudes are taken from the full waveform inversion catalogue of Willacy et al. (2018) if available and otherwise from the online KNMI cata-logue (https://www.knmi.nl/nederland-nu/seismologie/aardb evingen). Coordinates are *N*, *E*, *D* in metres using the RDS coordinate system standard for The Netherlands (*Rijksdrie-hoeksstelsel*)

| Child | Parent | Child | | Parent | | Child | Parent | Child | | | Parent | | |
Event ID		Date	Time	Date	Time	Magnitude		*N* (m)	*E* (m)	*D* (m)	*N* (m)	*E* (m)	*D* (m)
15062	15065	20150704	110547	20150707	030900	0.6	2.1	586751	237064	3000	586878	237996	3000
15106	15107	20151029	060510	20151029	080846	0.8	1.1	589197	259772	3000	589765	260294	3000
15078	15122	20150802	013243	20151215	074355	0.6	1.7	588555	238367	3000	588406	236168	3000
16011	15123	20160218	081944	20151225	041936	0.4	1.3	592550	244950	2900	592600	245000	3000
16029	16031	20160325	094639	20160402	004753	0.7	1.1	584891	238965	3000	585462	239756	3000
16043	16042	20160516	222437	20160516	203841	0.4	1.1	590150	256250	2900	590150	256300	2900
16072	16058	20160824	184423	20160718	085811	0.6	1.7	599230	243974	3000	599880	242963	3000
16117	16089	20161230	030553	20161101	001228	1.0	1.9	591100	249450	2900	591100	249400	2950
16118	16089	20161230	030607	20161101	001228	0.6	1.9	591100	249450	2900	591100	249400	2950
16117	16090	20161230	030553	20161101	005746	1.0	2.2	591100	249450	2900	591100	249400	2950
16118	16090	20161230	030607	20161101	005746	0.6	2.2	591100	249450	2900	591100	249400	2950
16093	16092	20161108	112533	20161108	112317	0.9	1.4	594550	248550	2900	594550	248550	2900
17023	16107	20170226	213948	20161207	015249	1.4	1.8	594300	247450	2950	594450	247400	3000
16118	16117	20161230	030607	20161230	030553	0.6	1.0	591100	249450	2900	591100	249450	2900
17011	17009	20170204	032737	20170204	025652	0.6	0.9	599134	244774	3000	599138	244974	3000
17010	17012	20170204	031126	20170205	154933	0.8	1.3	598704	239522	3000	599364	239045	3000
17016	17017	20170214	205204	20170215	120137	0.8	1.6	599112	243576	3000	599774	243298	3000
18014	17017	20180212	104453	20170215	120137	0.9	1.6	599200	243450	2950	599774	243298	3000
17037	17031	20170410	233713	20170404	100044	1.1	1.8	588350	261700	2900	588350	261750	2850
18015	17038	20180217	081730	20170413	103355	1.0	1.4	604900	245200	2850	605606	245584	3000
17050	17049	20170520	182544	20170520	024052	0.4	1.1	591000	256150	2850	591050	256250	2950
17064	17052	20170707	021936	20170527	152900	0.7	2.6	581340	251389	3000	581300	251600	2800
17038	17053	20170413	103355	20170530	034428	1.4	1.4	605606	245584	3000	606051	245575	3000
18015	17053	20180217	081730	20170530	034428	1.0	1.4	604900	245200	2850	606051	245575	3000
16009	17057	20160205	211138	20170607	002610	0.4	1.2	583550	240900	2900	583586	240657	3000
18028	17057	20180331	064403	20170607	002610	1.2	1.2	583550	241050	2950	583586	240657	3000
17063	17068	20170705	231902	20170716	210320	1.0	1.3	588415	236701	3000	588636	236631	3000
17075	17076	20170813	202619	20170814	031859	0.5	1.2	597900	245100	2900	597600	245300	2900
17087	17086	20170905	060958	20170905	053024	1.2	1.3	595800	253450	2800	595800	253500	2850
17124	17123	20171222	200611	20171222	194028	0.4	1.7	597500	245650	3000	597550	245600	2950
18001	17127	20180101	144651	20171228	140030	0.8	1.3	590100	254650	3000	590100	254650	3000
17115	17128	20171201	210541	20171229	231548	1.3	1.4	597602	246201	3000	597601	246135	3000
17123	18002	20171222	194028	20180108	140052	1.7	3.4	597550	245600	2950	597600	245600	2950
17124	18002	20171222	200611	20180108	140052	0.4	3.4	597500	245650	3000	597600	245600	2950
18005	18004	20180120	081921	20180117	043739	0.9	1.0	582200	235900	3250	582250	235850	3250
17016	18014	20170214	205204	20180212	104453	0.8	0.9	599112	243576	3000	599200	243450	2950
18096	18019	20181222	045411	20180301	130337	1.2	1.3	589700	247600	2750	589750	247600	2750
18019	18024	20180301	130337	20180325	110948	1.3	1.6	589750	247600	2750	589850	247600	2750
18096	18024	20181222	045411	20180325	110948	1.2	1.6	589700	247600	2750	589850	247600	2750
16009	18028	20160205	211138	20180331	064403	0.4	1.2	583550	240900	2900	583550	241050	2950
18019	18039	20180301	130337	20180506	163906	1.3	1.8	589750	247600	2750	589700	247650	2750

Table 2 (continued)

Child Event ID	Parent	Child Date	Time	Parent Date	Time	Child Parent Magnitude		Child N (m)	E (m)	D (m)	Parent N (m)	E (m)	D (m)
18024	18039	20180325	110948	20180506	163906	1.6	1.8	589850	247600	2750	589700	247650	2750
18096	18039	20181222	045411	20180506	163906	1.2	1.8	589700	247600	2750	589700	247650	2750
18064	18049	20180806	161523	20180627	143252	0.9	1.7	580966	243511	3000	580521	243519	3000
18066	18065	20180809	080155	20180808	025529	1.8	2.0	593900	253900	2900	594050	253900	2950
19018	18081	20190330	015136	20181020	045325	0.5	0.8	593900	254050	2800	593975	254136	3000
16064	19017	20160807	204022	20190329	133720	1.3	1.5	599357	238645	3000	598750	239250	2900
19047	19023	20190625	160931	20190416	021507	0.9	0.9	583032	240801	3000	583144	240866	3000
18059	19025	20180728	030948	20190425	192706	0.9	1.8	591500	244300	2800	591550	244300	2850
19034	19035	20190519	173833	20190519	224211	0.8	1.0	592550	248800	2900	592550	248800	3000
19036	19039	20190520	004942	20190523	021037	0.5	1.1	583200	240250	2950	583247	240396	3000
19038	19039	20190523	015208	20190523	021037	0.9	1.1	583250	240100	3100	583247	240396	3000
19044	19043	20190609	053430	20190609	050015	0.9	2.5	593700	246850	2900	593700	246900	2950

Synthetic seismograms are generated using the convolutional model of the seismic trace. Considering an EGF pair of events, the seismogram of the larger composite event (direct P or S arrivals only) can be written as a sum of the two slip patch contributions

$$S_l(x_i, t) = F_l(r_i, \theta_i, \phi_i) M_l W(x_i, t) * \left[\delta\left(t - \frac{Rs_i}{c}\right) + \delta\left(t - \frac{L}{\zeta} - \frac{Re_i}{c}\right) \right] \tag{14}$$

where $W(x_i, t)$ is the source wavelet assumed common to both phases, and x_i represents the coordinates of the ith station as shown in Fig. 3(a) and (b). It is assumed here that the ray arrives at the geophone stations approximately vertically after initial sub-horizontal propagation from the source (see Kraaijpoel and Dost 2013; Jagt et al. 2017); the travel time of the sub-vertical leg of the propagation will, to a good approximation, be common to all such rays at a given station and hence will drop out of the analysis. The smaller event is assumed to be a single phase with the same source wavelet and focal mechanism, collocated with the starting phase of the larger event

$$S_g(x_i, t) = F_g(r_i, \theta_i, \phi_i) M_g W(x_i, t) * \left[\delta\left(t - \frac{Rs_i}{c}\right) \right]. \tag{15}$$

Performing the EGF deconvolution as described above and making use of the sift property of the delta function under convolution, $f(t) * \delta(t - T) = f(t - T)$, give

$$S_l(x_i, t) * S_g^{-1}(x_i, t) = \frac{M_l}{M_g}\left(\delta(t) + \delta\left(t + \frac{Rs_i}{c} - \frac{L}{\zeta} - \frac{Re_i}{c}\right) \right). \tag{16}$$

For this idealised RSTF comprising two delta functions, the duration is

$$\Delta t(x_i) = \frac{Re_i}{c} - \frac{Rs_i}{c} + \frac{L}{\zeta} \approx \frac{L}{\zeta}\left(1 - \frac{\zeta}{c} Cos\phi_i\right). \tag{17}$$

This is the expression given by Savage (1965) and others (e.g. Li et al. 1995b) for the angle-dependent event duration in the case of unilateral rupture propagation. It can be seen as an example of angle-dependent Doppler broadening. For realistic band-limited pulses and duration measured between zero crossings, the width, w_t, of the individual pulses needs to be added to this expression for the duration:

$$\Delta t(x_i) \approx \frac{L}{\zeta} + w_t - \frac{L}{c} Cos\phi_i. \tag{18}$$

Table 3 Event pair EGF deconvolution results. "Offset" refers to the separation between the pair of events according to the catalogue locations. "Number of azimuths" is the number of distinct station azimuths used in the inversion of the picked duration on the interval $[-\pi, 3\pi]$. All other column notation is as used in the main text. Estimates of derived quantities were made using the formulae presented in the main text with stress drop $\Delta\sigma = 7.0$ MPa, average effective shear velocity $V_S = 2435$ m/s, average density $\rho = 2350$ kg/m^3 and $w_t \approx 0.017$ s

Child Event ID	Parent	Offset (m)	Φ (° from N)	L (m)	$A_1 = L/\zeta + w_t$ (s)	ξ	Number of azimuths	ζ (m/s)	$\bar{v}$ (m/s)	L_S (m)	L_D (m)
15062	15065	941	248.74	34.53	0.0481	0.58	36	1110.14	0.73	45.32	70.27
15106	15107	771	93.86	23.51	0.0460	0.80	30	810.69	0.15	37.10	13.81
15078	15122	2204	9.09	13.59	0.0539	0.04	30	368.22	0.45	24.38	48.02
16011	15123	122	2.51	0.83	0.0371	0.01	62	41.19	0.52	9.40	38.75
16029	16031	976	125.09	33.58	0.0564	0.56	68	852.25	0.11	45.69	12.45
16043	16042	50	70.87	21.46	0.0473	0.27	82	708.40	0.16	30.04	15.35
16072	16058	1202	−111.98	6.22	0.0473	0.07	42	205.33	0.61	17.02	57.48
16117	16089	71	−49.71	153.67	0.1129	0.85	84	1602.43	0.11	170.78	25.63
16118	16089	71	−50.58	154.68	0.1054	0.82	84	1749.81	0.12	165.48	26.04
16117	16090	71	161.51	25.15	0.0888	0.22	86	350.33	0.49	42.27	86.49
16118	16090	71	196.81	33.82	0.0846	0.12	80	500.33	0.50	44.62	84.18
16093	16092	0	9.75	19.70	0.0397	0.12	90	867.80	0.30	34.95	23.89
17023	16107	166	327.30	36.14	0.0742	0.42	82	631.74	0.24	63.26	35.43
16118	16117	0	135.66	20.60	0.0404	0.20	70	880.35	0.16	31.40	12.63
17011	17009	200	49.29	8.52	0.0404	0.05	52	364.21	0.17	19.32	13.55
17010	17012	814	−25.88	18.81	0.0379	0.33	80	900.02	0.28	32.40	20.88
17016	17017	718	−62.44	23.67	0.0411	0.22	85	982.08	0.40	37.26	32.68
18014	17017	596	−31.54	21.59	0.0443	0.18	91	790.71	0.37	36.84	32.87
17037	17031	71	118.28	15.82	0.0575	0.12	101	390.50	0.42	35.02	47.62
18015	17038	818	343.19	16.11	0.0379	0.33	54	770.69	0.32	33.22	24.50
17050	17049	150	315.18	6.46	0.0465	0.10	64	219.15	0.23	15.04	21.69
17064	17052	293	188.24	153.56	0.0965	0.63	122	1931.61	0.45	165.68	87.16
17038	17053	445	346.03	39.86	0.0354	0.43	108	2166.36	0.24	66.98	17.26
18015	17053	1220	322.58	9.01	0.0415	0.23	60	367.73	0.33	26.12	27.63
16009	17057	265	176.85	9.17	0.0488	0.14	64	288.29	0.24	17.74	23.74
18028	17057	398	183.27	19.72	0.0493	0.37	78	610.63	0.16	41.27	15.57
17063	17068	232	200.82	17.07	0.0473	0.25	118	563.34	0.22	34.18	20.33
17075	17076	361	213.31	7.44	0.0358	0.12	56	395.69	0.34	17.06	24.21
17087	17086	71	2.91	12.65	0.0520	0.15	89	361.42	0.20	34.19	20.32
17124	17123	87	−21.50	13.81	0.0441	0.17	128	509.46	0.57	22.38	50.12
18001	17127	0	77.14	12.69	0.0509	0.09	124	374.23	0.23	26.28	23.18
17115	17128	66	−29.30	13.21	0.0676	0.15	84	261.06	0.17	37.38	23.10
17123	18002	50	305.36	99.80	0.1222	0.67	128	948.65	1.56	138.11	380.05
17124	18002	122	-45.01	118.10	0.1206	0.60	128	1139.94	1.65	126.67	396.83
18005	18004	71	224.60	16.57	0.0419	0.33	74	665.47	0.15	31.82	12.55
17016	18014	162	379.22	3.43	0.0451	0.04	60	122.18	0.16	17.03	14.43
18096	18019	50	−22.78	29.17	0.0583	0.47	124	706.33	0.14	50.71	16.69
18019	18024	100	−12.25	20.08	0.0593	0.38	124	474.62	0.25	44.25	29.99
18096	18024	150	348.86	16.81	0.0673	0.12	108	334.27	0.24	38.36	32.21
16009	18028	158	168.82	14.85	0.0450	0.24	64	530.48	0.23	23.43	20.66
18,019	18039	71	165.21	28.21	0.0766	0.40	120	473.31	0.26	52.38	38.94

Table 3 (continued)

Child Event ID	Parent	Offset (m)	Φ (° from N)	L (m)	$A_1 = L/\zeta + w_t$ (s)	ξ	Number of azimuths	ζ (m/s)	$\bar{v}$ (m/s)	L_S (m)	L_D (m)
18024	18039	158	167.15	61.90	0.0769	0.64	110	1033.35	0.19	96.04	28.76
18096	18039	50	179.36	40.91	0.0746	0.52	122	710.21	0.24	62.45	35.66
18064	18049	445	191.15	9.93	0.0566	0.09	126	250.88	0.42	25.19	47.25
18066	18065	158	14.06	55.12	0.0845	0.44	126	816.54	0.24	98.10	40.19
19018	18081	230	90.70	11.37	0.0435	0.06	88	429.11	0.13	20.99	10.94
16064	19017	863	351.17	21.43	0.0478	0.24	80	695.71	0.26	45.60	24.86
19047	19023	129	183.45	12.02	0.0444	0.32	82	438.57	0.13	27.27	11.41
18059	19025	71	10.19	10.19	0.0605	0.11	126	234.26	0.46	25.44	55.87
19034	19035	100	3.35	8.61	0.0460	0.11	106	296.82	0.16	22.20	15.02
19036	19039	161	193.21	17.48	0.0392	0.11	76	787.54	0.21	27.11	16.16
19038	19039	312	199.98	5.81	0.0409	0.08	80	242.99	0.23	21.06	18.33
19044	19043	71	349.22	10.47	0.0572	0.23	132	260.46	1.63	25.72	186.12

Intermediate slip patches can be inserted between the starting and stopping phases without changing the azimuthal dependence of duration of the RSTF. Moreover, if there is an offset between the two nominally collocated sources, we find that this introduces a second azimuthally variant term to the arrival times but does not change the duration for the assumed homogeneous subsurface. If

Rg_i is the epicentral distance between the smaller source and the ith receiver station (Fig. 3(a)) and $Rg_i \neq Rs_i$, then

$$S_g\left(x_i, t\right) = F_g\left(r_i, \theta_i, \phi_i\right) M_g W\left(x_i, t\right) * \left[\delta\left(t - \frac{Rg_i}{c}\right)\right] \tag{19}$$

and

$$S_l\left(x_i, t\right) * S_g^{-1}\left(x_i, t\right) = \frac{M_l}{M_g}\left(\delta\left(t + \frac{Rg_i}{c} - \frac{Rs_i}{c}\right) + \delta\left(t + \frac{Rg_i}{c} - \frac{L}{\zeta} - \frac{Re_i}{c}\right)\right). \tag{20}$$

The terms Rg_i/c cancel when calculating the duration so that the RSTF duration is independent of the offset of the smaller source. In reality, the cancellation of propagation terms shown in Eq. (3) will deteriorate with increasing offset, consistent with the findings of Abercrombie (2015). If the distance of the smaller source from the epicentre of the larger source is Δ and the azimuth of the offset vector relative to the strike direction of the larger event is Ψ, following the above steps we can express the arrival times of the two pulses as

$$t_1\left(x_i\right) = (Rs_i - Rg_i)/c \approx (\Delta/c)Cos(\Psi - \phi_i) \tag{21}$$

and

$$t_2\left(x_i\right) = L/\zeta + (Re_i - Rg_i)/c \approx (L/\zeta)(1 - (\zeta/c)Cos\phi_i) + (\Delta/c)Cos(\Psi - \phi_i). \tag{22}$$

This highlights the close connection between EGF and relative event location methods. The above expressions could be used to invert the azimuthally varying arrival times for the source offset parameters, Δ and Ψ, extending the scope of the usual EGF workflow to also generate relative event location refinements.

We can similarly construct the RSTF for bi-lateral propagation, with a central starting phase giving rise to a pair of stopping phases at opposite ends of the fault trace, Re_i and Re'_i from the ith station (Fig. 3(b)):

$$S_l(t) = F_l(r_i, \theta_i, \phi_i) M_l W(x_i, t) * \left[\delta\left(t - \frac{Rs_i}{c}\right) + \delta\left(t - \frac{L}{\zeta} - \frac{Re_i}{c}\right) + \delta\left(t - \frac{L}{\zeta} - \frac{Re'_i}{c}\right) \right] \tag{23}$$

$$S_l(x_i, t) * S_g^{-1}(x_i, t) = \frac{M_l}{M_g} \left(\delta(t) + \delta\left(t + \frac{Rs_i}{c} - \frac{L}{\zeta} - \frac{Re_i}{c}\right) + \delta\left(t + \frac{Rs_i}{c} - \frac{L}{\zeta} - \frac{Re'_i}{c}\right) \right). \tag{24}$$

Table 4 Refinement of rupture propagation lengths for correlated clusters of three or more events. The strike direction, Φ, is a cluster average where rotations of 180° have been applied where needed to approximately align all rupture vectors with the first named event. After this realignment with the common direction, the set of coefficients of the cosine terms in Eq. (29) from all event pairs in the cluster can be inverted for the individual event rupture distances. Strike and dip direction rupture patch sizes are estimated in the same way as in Table 2 and 3 above: L_S is taken to be the sum of the slip patch length of the smaller event (assumed square) and the rupture propagation distance; L_D is then derived from the relationship between moment and area

Cluster	Event ID	Average Φ (° from N)	L (m)	L_S (m)	L_D (m)
C1	16090	164.66	19.19	29.99	102.68
	16089	−48.21	164.48	175.28	25.30
	16117	148.23	8.39 (±3.42)	19.19	16.16
	16118	148.23	12.22 (±3.42)	23.02	7.39
C2	17017	−24.92	20.92	34.51	33.96
	18014	−24.92	0.68	14.27	15.77
	17016	−24.92	2.76	16.35	12.39
C3	17053	337.27	16.38	33.49	24.41
	17038	337.27	23.48	40.59	22.17
	18015	157.27	7.37	24.48	14.31
C4	17057	176.31	7.02	15.60	25.32
	18028	176.31	12.70	21.28	21.68
	16009	176.31	2.15	10.73	7.67
C5	18002	-40.38	102.05	110.62	424.65
	17123	139.62	2.25	10.82	72.08
	17124	−40.38	16.06	24.63	5.06
C6	18039	168.59	54.68 (±5.34)	76.22	32.28
	18024	167.70	1.68 (±5.34)	23.22	41.40
	18019	346.19	21.35 (±5.34)	42.89	18.15
	18096	347.89	13.36 (±5.34)	34.90	16.92

The duration of the RSTF is obtained as the maximum value of the time difference between pairs of arrivals. For this idealisation of bilateral rupture propagation, we obtain the following expression for the duration as a function of the azimuthal angle:

$$\Delta t(x_i) \approx \frac{L}{\zeta}\left(1 + \frac{\zeta}{c}|Cos\phi_i|\right) + w_t. \tag{25}$$

This can be further generalised to cover the case of bilateral propagation where the fault trace lengths in the two directions are different (L and L' say) and the rupture propagation velocities are also allowed to be different (ζ and ζ'). In this case of asymmetric bilateral propagation, the duration is

$$\Delta t(x_i) \approx max\left\{ \frac{L}{\zeta}\left(1 - \frac{\zeta}{c}Cos\phi_i\right), \frac{L'}{\zeta'}\left(1 + \frac{\zeta'}{c}Cos\phi_i\right) \right\} + w_t. \tag{26}$$

Note that this expression reduces to the previous cases of symmetric bilateral propagation for $L = L'; \zeta = \zeta'$ and unilateral propagation for $L' = 0$. A distinguishing characteristic of bilateral propagation is that the duration will have two local maxima aligned with the two arms of the rupture. In the case that $L = L'$ and $\zeta = \zeta'$, the duration is expected to be periodic on the interval $[0, \pi]$, rather than $[0, 2\pi]$ as in the unilateral case.

A further generalisation to consider in support of the event pairs analysed is the case of an event pair of similar magnitudes. In this case we should consider the possibility that the smaller event should also be treated as an extended rupture. Primed variables $\{R's_i, R'e_i, L', \zeta', \phi'_i, r'_i, \theta'_i\}$ refer to the smaller event in such a scenario and are the analogues of $\{Rs_i, Re_i, L, \zeta, \phi_i, r_i, \theta_i\}$ in Fig. 3(a). The seismogram for the smaller event is

$$S_g(x_i, t) = F_g(r'_i, \theta'_i, \phi'_i) M_g W(x_i, t) * \left[\delta\left(t - \frac{R's_i}{c} \right) + \delta\left(t - \frac{L'}{\zeta'} - \frac{R'e_i}{c} \right) \right]. \tag{27}$$

Although it can be demonstrated numerically that this composite seismogram has at least an approximate inverse, the convolutional inverse cannot be written using a simple algebraic expression such as used to obtain Eq. (16). Referring to Eq. (5), it can be seen that the limit of very high additive deconvolution noise ($N \gg 1$) can be expressed in simple algebraic terms since the output generated by the deconvolving function reduces to a scaled convolution of the pair of seismograms.

$$\begin{aligned} S_l(x_i, t) * S_g(x_i, t) \propto & \left(\delta\left(t - \frac{R's_i}{c} - \frac{Rs_i}{c} \right) + \delta\left(t - \frac{R's_i}{c} - \frac{L}{\zeta} - \frac{Re_i}{c} \right) + \delta\left(t - \frac{Rs_i}{c} - \frac{L'}{\zeta'} - \frac{R'e_i}{c} \right) \right. \\ & \left. + \delta\left(t - \frac{L'}{\zeta'} - \frac{R'e_i}{c} - \frac{L}{\zeta} - \frac{Re_i}{c} \right) \right) * W(t) * W(t) \end{aligned} \tag{28}$$

and hence, in this limit,

$$\Delta t(x_i) \approx \frac{L}{\zeta}\left(1 - \frac{\zeta}{c} Cos\phi_i \right) + \frac{L'}{\zeta'}\left(1 - \frac{\zeta'}{c} Cos\phi'_i \right) + w_i. \tag{29}$$

If $\phi'_i \approx \phi_i$ and $\zeta' = \zeta$ this reduces to

$$\Delta t(x_i) \approx \frac{(L + L')}{\zeta}\left(1 - \frac{\zeta}{c} Cos\phi_i \right) + w_t. \tag{30}$$

That is, for this idealisation in which a pair of parallel composite ruptures are processed with the EGF deconvolution workflow with the application of very high additive noise, the duration expression is of the same form as before but proportional to the sum of the rupture propagation distances of the two events.

In the main body of the text, it is explained that the RSTF should be approximately the same for all depth levels and components of the same surface station, since the spectral division removes all effects common to both seismograms, including path and coupling effects. In the data processing workflow the RSTFs are summed over all geophone levels (50 m, 100 m, 150 m, 200 m) and components (X, Y, Z) of a station to improve the signal to noise ratio of the result. To demonstrate this, Fig. 16 compares the summed RSTFs with the individual RSTFs for each geophone for event pair 17123/18002. A different style of display from the previous figures was found useful here. The RSTF data are sorted on component type (11 = X, 12 = Y, 13 = Z), geophone depth (50 m, 100 m, 150 m, 200 m), and then station azimuth in degrees. The plot shows the full range of station azimuths between ± 180° for each depth/ component combination. The leftmost panel shows the stacked RSTFs. Because of the compressed horizontal scale, wiggle traces have not been shown and a colour scale which was found to accentuate the varying trace amplitudes has been chosen. The individual RSTFs show the same overall form and azimuth dependence, and the visible noise is reduced by the summation as expected.

Tables 2, 3 and 4.

References

Abercrombie RE (2015) Investigating uncertainties in empirical Green's function analysis of earthquake source parameters. J Geophys Res: Solid Earth 120:4263–4277

Abercrombie RE, Poli P, Bannister S (2017) Earthquake directivity, orientation, and stress drop within the subducting plate at the Hikurangi margin, New Zealand. J Geophys Res: Solid Earth 122(10):10176–10188

Ameri G, Martin C, Oth A (2020) Ground-motion attenuation, stress drop, and directivity of induced events in the Groningen gas field by spectral inversion of borehole records. Bull Seismol Soc Am 110(5):2077–2094

Arrowsmith SJ, Eisner L (2006) A technique for identifying microseismic multiplets and application to the Valhall field. North Sea Geophysics 71(2):v31–v40

Beresnev IA (2001) What we can and cannot learn about earthquake sources from the spectra of seismic waves. Bull Seismol Soc Am 91(2):397–400

Beresnev IA (2002) Source parameters observable from the corner frequency of earthquake spectra. Bull Seismol Soc Am 92(5):2047–2048

Berkhout AJ (1977) Least-squares inverse filtering and wavelet deconvolution. Geophysics 42(7):1369–1383

Bourne SJ, Oates SJ (2017) Extreme threshold failures within a heterogeneous elastic thin-sheet and the spatial-temporal development of induced seismicity within the Groningen gas field. J Geophys Res: Solid Earth 122(10):10299–10320

Bourne SJ, Oates SJ, van Elk J (2018) The exponential rise of induced seismicity with increasing stress levels in the Groningen gas field and its implications for controlling seismic risk. Geophys J Int 213(3):1693–1700

Buijze L, van den Bogert PAJ, Wassing BBT, Orlic B (2019) Nucleation and arrest of dynamic rupture induced by reservoir depletion. J Geophys Res: Solid Earth 124(4):3620–3645

Cesca S, Heimann S, Dahm T (2011) Rapid directivity detection by azimuthal amplitude spectra inversion. J Seismolog 15:147–164

Daniel G, Fortier E, Romijn R, Oates S (2016) Location results from borehole microseismic monitoring in the Groningen gas reservoir, Netherlands. 6th EAGE Workshop on Passive Seismic, Muscat, Oman

De Jager J, Visser C (2017) Geology of the Groningen field – an overview. Geol Mijnbouw/Netherlands J Geosci 96(5):s3–s15

Dost B, Ruigrok E, Spetzler J (2017) Development of seismicity and probabilistic hazard assessment for the Groningen gas field. Geol Mijnbouw/Netherlands J Geosci 96(5):s235–s245

Dost B, van Stiphout A, Kühn D, Kortekaas M, Ruigrok E, Heimann S (2020) Probabilistic moment tensor inversion for hydrocarbon-induced seismicity in the Groningen gas field, the Netherlands, Part 2: Application. Bull Seismol Soc Am 110(5):2112–2123

Edwards B, Zurek B, van Dedem E, Stafford PJ, Oates S, van Elk J, deMartin B, Bommer JJ (2019) Simulations for the development of a ground motion model for induced seismicity in the Groningen gas field, The Netherlands. Bull Earthq Eng 17:4441–4456

Folesky J, Kummerow J, Shapiro SA, Häring M, Asanuma H (2016) Rupture directivity of fluid-induced microseismic events: observations from an enhanced geothermal system. J Geophys Res: Solid Earth 121:8034–8047

Graves R, Pitarka A (2016) Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations. Bull Seismol Soc Am 106(5):2136–2153

Havskov J, Alguacil G (2004) Instrumentation in earthquake seismology. Springer, Dordrecht

Hunfeld LB, Niemeijer AR, Spiers CJ (2017) Frictional properties of simulated fault gouges from the seismogenic Groningen gas field under in situ P-T chemical conditions. J Geophys Res: Solid Earth 122:8969–8989

Hutchings L, Viegas G (2012) Application of empirical Green's functions in earthquake source, wave propagation and strong ground motion studies. In: D'Amico S (ed) Earthquake Research and Analysis - New Frontiers in Seismology. InTech, London, pp 87–140

Imanishi K, Takeo M (2002) An inversion method to analyze rupture processes of small earthquakes using stopping phases. J Geophys Res 107(B3):2048

Jagt L, Ruigrok E, Paulssen H (2017) Relocation of clustered earthquakes in the Groningen gas field. Neth J Geosci 96(5):s163–s173

Kanamori H, Anderson DL (1975) Theoretical basis of some empirical relations in seismology. Bull Seismol Soc Am 65(5):1073–1095

King G, Nabelek J (1985) Role of fault bends in the initiation and termination of earthquake rupture. Science 228:984–987

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

Kortekaas M, Jaarsma B (2017) Improved definition of faults in the Groningen field using seismic attributes. Neth J Geosci 96(5):s71–s85

Kraaijpoel DA, Dost B (2013) Implications of salt-related propagation and mode conversion effects on the analysis of induced seismicity. J Seismolog 17(1):95–107

Kühn DS, Heimann MP, Isken E, Ruigrok E, Dost B (2020) Probabilistic moment tensor inversion for hydrocarbon-induced seismicity in the Groningen gas field, The Netherlands, Part 1: Testing. Bull Seismol Soc Am 110(5):2095–2111

Leonard M (2010) Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. Bull Seismol Soc Am 100(5A):1971–1988

Leonard M (2012) Erratum to Earthquake fault scaling: self-consistent relating of rupture length, width, average displacement, and moment release. Bull Seismol Soc Am 102(6):2797–2797

Li Y, Toksoz M, Rodi W (1995a) Source time functions of nuclear explosions and earthquakes in central Asia determined using empirical Green's functions. J Geophys Res 100:659–674

Li Y, Doll C, Toksöz MN (1995b) Source characterization and fault plane determination for M_{bLg} = 1.2 to 4.4 earthquakes in the Charlevoix Seismic Zone, Quebec, Canada. Bull Seismol Soc Am 85(6):1604–1621

Noda H, Lapusta N, Kanamori H (2013) Comparison of average stress drop measures for ruptures with heterogeneous stress change and implications for earthquake physics. Geophys J Int 193(3):1691–1712

Park S, Ishii M (2015) Inversion for rupture properties based upon 3-D directivity effect and application to deep earthquakes in the Sea of Okhotsk region. Geophys J Int 203(2):1011–1025

Savage JC (1965) The effect of rupture velocity upon seismic first motions. Bull Seismol Soc Am 55(2):263–275

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

Spica ZJ, Nakata N, Liu X, Campman X, Tang Z, Beroza GC (2018) The ambient seismic field at Groningen gas field: an overview from the surface to reservoir depth. Seismol Res Lett 89(4):1450–1466

Tomic J, Abercrombie RE, do Nascimento AF (2009) Source parameters and rupture velocity of small $M \leq 2.2$ reservoir induced earthquakes. Geophys J Int 179(2):1013–1023

Udías A, Madariaga R, Buforn E (2014) Source mechanisms of earthquakes: theory and practice. Cambridge University Press, Cambridge

van Dedem E, Willacy C, Blokland JW, Piesold T, Minisini S (2018) Full waveform event location and cluster analysis for Groningen induced seismicity. 80th EAGE Conf Exhibition 2018, Copenhagen, Denmark

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

Wang E, Rubin AM, Ampuero JP (2014) Compound earthquakes on a bimaterial interface and implications for rupture mechanics. Geophys J Int 197(2):1138–1153

Weng H, Ampuero JP (2019) The dynamics of elongated earthquake ruptures. J Geophys Res: Solid Earth 124:8584–8610

Willacy C, van Dedem E, Minisini S, Li J, Blokland JW, Das I, Droujinine A (2018) Application of full-waveform event location and moment-tensor inversion for Groningen induced seismicity. Lead Edge 37(2):92–99

Willacy C, van Dedem E, Minisini S, Li J, Blokland JW, Das I, Droujinine A (2019) Full-waveform event location and moment tensor inversion for induced seismicity. Geophysics 84(2):KS39–KS57

Willacy C, Blokland JW, van Dedem E (2020) Automating event location monitoring for induced seismicity. Lead Edge 39(7):505–512

Yoshido K, Saito T, Emoto K, Urata Y, Sato D (2019) Rupture directivity, stress drop, and hypocenter migration of small earthquakes in the Yamagata-Fukushima border swarm triggered by upward pore-pressure migration after the 2011 Tohoku-Oki earthquake. Tectonophysics 769:228184

Zurek B, deMartin B (2019) Using waveform simulations to help constrain kinematics of small earthquakes at the Groningen gas field. AGU Fall Meet 2019, San Francisco, USA

Zurek B, Burnett W, Dedontney N, Gist G (2017) The effect of modeling kinematic finite faults on deterministic formulation of ground motion prediction equations – Groningen an induced seismicity case study. SEG Int Exposition 87th Annu Meet, 2017, Houston, USA