

Fitting Earthquake Spectra: Colored Noise and Incomplete Data

by Stefano Maranò, Benjamin Edwards, Graziano Ferrari, and Donat Fäh

Abstract Spectral analysis of earthquake recordings provides fundamental seismological information. It is used for magnitude calculation, estimation of attenuation, and the determination of fault rupture properties including slip area, stress drop, and radiated energy. Further applications are found in site-effect studies and for the calibration of simulation and empirically based ground-motion prediction equations.

We identified two main limitations of the spectral fitting methods currently used in the literature. First, the frequency-dependent noise level is not properly accounted for. Second, there are no mathematically defensible techniques to fit a parametric spectrum to a seismogram with gaps.

When analyzing an earthquake recording, it is well known that the noise level is not the same at different frequencies, that is, the noise spectrum is colored. The different, frequency-dependent, noise levels are mainly due to ambient noise and sensor noise. Methods in the literature do not properly account for the presence of colored noise.

Seismograms with gaps are usually discarded due to the lack of methodologies to use them. Modern digital seismograms are occasionally clipped at the arrival of the strongest ground motion. This is also critical in the study of historical earthquakes in which few seismograms are available and gaps are common, significantly decreasing the number of useful records.

In this work, we propose a method to overcome these two limitations. We show that the spectral fitting can be greatly improved and earthquakes with extremely low signal-to-noise ratio can be fitted. We show that the impact of gaps on the estimated parameters is minor when a small fraction of the total energy is missing. We also present a strategy to reconstruct the missing portion of the seismogram.

Introduction

Spectral analysis of earthquake recordings is required for numerous seismological applications. For example, determination of source parameters such as magnitude (Ottemoller and Havskov, 2003; Edwards *et al.*, 2010), earthquake stress drop (Atkinson, 1993; Hough *et al.*, 1999; Goertz-Allmann and Edwards, 2014), radiated energy (Boatwright *et al.*, 2002; Harrington and Brodsky, 2009), and slip area (Edwards *et al.*, 2015). In particular, the determination of stress drop for use in the forward modeling of earthquake ground motion (stochastic simulations, Boore, 2003) has a large impact on seismic hazard for regions of low-to-moderate seismicity (Cotton *et al.*, 2013)—because strong-motion data are not available. Path parameters that can be studied include Q and geometrical spreading (e.g., Raoof *et al.*, 1999; Allen *et al.*, 2007) and attenuation (Douglas *et al.*, 2010; Edwards *et al.*, 2011). In addition, site effects such as amplification and damping are also considered (e.g., Edwards *et al.*, 2013; Ktenidou *et al.*, 2013). Spectral analysis is also used in the calibration of simulation (Douglas *et al.*, 2013; Rietbrock *et al.*, 2013; Atkinson, 2015) and empirically based (Bora *et al.*, 2015) ground-motion prediction equations.

Two distinct approaches are used in the spectral analyses of earthquake signals: the first approach exploits model redundancy by taking ratios of spectra, and the second approach fits the recorded spectrum directly. For the latter, significantly more data are available, however, modeling these data suffers more from nonuniqueness and the influence of noise in the spectrum (Boore *et al.*, 1992). To avoid the influence of noise, authors often implement signal-to-noise ratio (SNR) criteria and only analyze data passing a defined threshold. In the past, it was more common to directly account for the influence of noise in recorded spectra (e.g., Boatwright *et al.*, 1991), however, with the improvement of modern high-quality seismic instrumentation (with low self-noise and high sensitivity), this is not commonly used.

In this work, we develop a technique for fitting a parametric spectrum to earthquake recordings and focus on both the cases of complete and incomplete data. In the first instance, the whole seismogram is observed. This is the most commonly encountered situation in modern applications, in which high quality, reliable digital instrumentation is available. In the second instance, the recorded seismograms may

exhibit some gaps. A gap is defined as a short period in which ground motion is not recorded in the seismogram. This circumstance is particularly relevant to the study of historical earthquakes, in which old instrumentation may have intermittently failed, or was unreliable. Moreover, modern instrumentation may clip around the peak motion.

In general, gaps in historical seismograms are due to the high velocity of the ground motion that does not allow the writing device to properly record the motion. For example, in mechanical instruments the extremely lightweight nib may lift up from the smoked paper due to aerodynamic effects. In the case of optical recording, the light beam may move too fast and not have sufficient time to expose the photographic paper.

As an example, [Vannoli *et al.* \(2015\)](#) study an Italian earthquake from 1930. Out of the 113 seismograms recovered from seismological observatories across Europe, only 11 were used in the final study, mostly due to the fact that incomplete seismograms could not be analyzed. Recovering a single seismogram requires a significant amount of resources and dedication. Therefore, being able to utilize incomplete seismograms is of great importance.

In both the case of complete and incomplete data, we are interested in appropriately accounting for the correlated nature of the noise. It is well established that not all the frequencies should have the same weight in determining the spectral fit of earthquake recordings. Certain frequencies exhibit a higher ambient noise level than others ([Cauzzi and Clinton, 2013](#)). In addition, the spectra exhibit energy spreads unevenly across frequencies—in which the frequency of recorded energy is defined by a combination of earthquake magnitude, fault kinematics, and attenuation.

Concerning the case in which we have complete data, a common approach for fitting spectral models to a recorded seismogram is to perform a nonlinear least-squares fit in the frequency domain. More precisely, the fit of the model to the logarithm of the spectral amplitude of data is evaluated with a least-squares criterion. The underlying implicit assumption is that the spectral amplitudes follow a lognormal distribution. Such an assumption is, however, incompatible with time-domain measurements following a Gaussian distribution.

The presence of a different noise level at different frequencies is typically addressed in the literature with a threshold mechanism. Let SNR_ω denote the SNR at angular frequency ω . The SNR_ω is computed as the ratio of Fourier amplitude and noise level at discrete frequencies. Whenever SNR_ω exceeds a given threshold, the data at that frequency are used in the fit; otherwise, if it is below the threshold, the data are completely ignored. For example, a common choice is to consider frequencies with $\text{SNR}_\omega > 3$ (see, e.g., [Oth *et al.*, 2011](#)). Hard thresholding using SNR is suboptimal for two reasons. First, certain frequencies are completely ignored despite the fact that they still carry useful information. Second, the frequencies in use are given equal importance, whereas their different SNR_ω suggests that they should be given different levels of importance. Other authors address this by weighting the least-squares misfit penalty

function using the SNR. For example, [Boatwright *et al.* \(1991\)](#) used SNR-based weighting, limiting the maximum weight to the equivalent of $\text{SNR}_\omega = 2$. This approach does account for low-SNR regions, but also uniformly weights data above this arbitrary threshold.

Concerning the incomplete data case, we are not aware of any rigorous methods to fit a spectrum to a seismogram with gaps. A group of techniques used in the analysis of unevenly spaced data are referred to as least-squares spectral analysis ([Lomb, 1976](#); [Scargle, 1982](#)). In these techniques, one seeks a least-squares fit of sinusoids to measurements. It will be clear that our approach has some similarities with these techniques.

In [Čadek \(1987\)](#), it is reported how problematic the presence of gaps in historical recordings is; and that practically all the recordings considered present gaps, for example, due to minute marks, which were used in historical seismograms for timing. [Čadek \(1987\)](#) reported that several different methods are used in order to complete the missing part of the recording. Unfortunately, details of the attempted approaches and of the evaluation are not presented. However, it is mentioned that the best results are obtained through filling the gaps by copying a portion of the signal in the vicinity of the gap.

The central task addressed in this work is fitting a parametric spectral model to an observed seismogram. The proposed approach relies on the maximum-likelihood (ML) criterion. The likelihood function is used to evaluate the goodness of fit of a spectrum to the seismogram and to estimate model parameters of interest. In contrast with most of the existing methods, which are working in the frequency domain, our approach formulates the likelihood in the time domain. The actual computation is then performed more efficiently in the frequency domain. The proposed approach improves methods present in the literature in two different ways: modeling the frequency-dependent noise level due to the correlated nature of the noise and accounting for the presence of gaps in the seismogram.

By modeling the correlated noise present in the seismogram, we are able to appropriately weigh the impact of different frequencies on the misfit. Frequencies with high SNR are given more importance in determining the final best-fitting model. This enables us to use the entire signal and not merely selected frequencies that satisfy some quality criterion. In practice, the presence of, for example, high-frequency instrument noise will not significantly affect our estimation. Moreover, we are able to use seismograms with extremely low SNR that would not be usable with traditional methods.

Within the same framework we also address the presence of gaps in the seismograms. We do not make any assumption on the signal during the portions that were not observed. Only the portion of the seismogram that is observed is used to compute the likelihood, and the unobserved portion is simply ignored.

We emphasize that the proposed approach allows the computation of the likelihood of the data for both a paramet-

ric spectrum and a nonparametric spectrum. A parametric spectrum is typically described in closed form and has few parameters (e.g., Brune, 1970; Madariaga, 1976). A nonparametric spectrum may be computed numerically, for example, with a numerical modeling of an earthquake rupture.

Outline and Notation

The next sections of this article are organized as follows. In the [Problem Statement and System Model](#) section, the spectrum fitting task is formulated mathematically. In the [Proposed Approach](#) section, the proposed approach is presented. Finally, in the [Numerical Examples and Applications](#) section, several representative applications on real recordings are showcased. Details pertaining to the numerical implementation of the proposed techniques are discussed in the [Appendix](#).

We denote deterministic scalars, vectors, and matrices as a , \mathbf{a} , and \mathbf{A} , respectively. Random variables are always capitalized and the corresponding realization is lowercase, disregarding whether they are scalars or vectors. For instance, we denote with $\mathbf{X} \sim \mathcal{N}(\mathbf{m}_X, \mathbf{V}_X)$ a vector random variable \mathbf{X} , distributed according to a multivariate Gaussian distribution with mean vector \mathbf{m}_X and covariance matrix \mathbf{V}_X . We denote with $\mathcal{N}(\mathbf{x}; \mathbf{m}_X, \mathbf{V}_X)$ the value of a multivariate Gaussian distribution with mean vector \mathbf{m}_X and covariance matrix \mathbf{V}_X evaluated at \mathbf{x} , that is,

$$\begin{aligned} \mathcal{N}(\mathbf{x}; \mathbf{m}_X, \mathbf{V}_X) &= |\mathbf{V}_X|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m}_X)^\top \mathbf{V}_X^{-1}(\mathbf{x} - \mathbf{m}_X)\right). \end{aligned} \quad (1)$$

Problem Statement and System Model

In this work, we present a method for fitting a parametric spectral model $|\tilde{u}(\omega, \boldsymbol{\theta})|$ to observed time-domain measurements y_1, y_2, \dots, y_N . The model is described with a known function of the parameters $\boldsymbol{\theta}$. In most spectral fitting applications, the phase spectrum is not specified (Hanks and McGuire, 1981; Boore, 2003). As an example application of the proposed method, we present the fitting of the Brune far-field source model to the seismogram (Brune, 1970). The Brune model is a widespread approach in the analysis of earthquakes. We note that the proposed method can work with any other parametrization of the spectral amplitudes $|\tilde{u}(\omega, \boldsymbol{\theta})|$ as a function of the parameters $\boldsymbol{\theta}$.

When fitting the spectrum, we also need to consider that the noise is correlated. Correlated noise is equivalent to the noise spectral density not being flat, that is, colored noise. By properly modeling the correlated noise, we can appropriately weigh how the different frequencies affect the spectral fit. In practice, frequencies with high SNR drive the resulting fit, whereas frequencies with low SNR are less important. In this way, it is possible to account for the high-frequency noise of certain instruments, for the microseisms, and to fit spectra to earthquake recordings with very low SNR. As explained in

the following, to fit the time series, it is necessary to jointly estimate both the spectral parameters of interest $\boldsymbol{\theta}$, controlling the amplitude spectrum and the nuisance parameters $\boldsymbol{\phi}$, which control the phase spectrum. In addition to the case of complete data, we are interested in the accounting for the presence of gaps in the seismogram. This is particularly relevant to the study of historical earthquakes, in which instruments may have not recorded parts of the seismogram. Throughout this article, we refer to this circumstance as the incomplete data case.

Measurement and Noise Model

Consider the following measurement model:

$$Y_k = u(t_k) + Z_k, \quad (2)$$

for $k = 1, \dots, N$. Here, Y_k and Z_k are scalar random variables modeling the measurement and the additive noise, respectively, at time t_k . The time-domain noiseless signal at time t_k is denoted with $u(t_k)$. Samples are uniform $t_k = (k-1)T_s$, with T_s as the sampling interval. In vector notation, we denote the noiseless signal as $\mathbf{u} = (u(t_1), u(t_2), \dots, u(t_N))^\top$, the noisy measurements as $\mathbf{Y} = (Y_1, Y_2, \dots, Y_N)^\top$, and additive Gaussian noise as \mathbf{Z} , that is,

$$\mathbf{Y} = \mathbf{u} + \mathbf{Z}, \quad (3)$$

in which $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_Z)$.

We model the additive noise \mathbf{Z} with a multivariate Gaussian distribution. The noise is assumed to be wide-sense stationary and thus is completely described by the (zero) mean vector and covariance matrix \mathbf{V}_Z . The covariance matrix \mathbf{V}_Z is, in general, not diagonal. Given the autocovariance function of the noise $\gamma(\tau)$, the noise covariance matrix is the Toeplitz matrix $\boldsymbol{\Gamma}$ with the element in position j, k being

$$[\boldsymbol{\Gamma}]_{j,k} = \gamma(|t_j - t_k|). \quad (4)$$

In the complete data case, that is, when the seismogram has no gaps, $\mathbf{V}_Z = \boldsymbol{\Gamma}$.

In the incomplete data case, only $n < N$ measurements are observed. We introduce the measurement matrix $\mathbf{P} \in \mathbb{R}^{n \times N}$ having full row-rank, with exactly one entry as 1 on each row and the remaining entries as 0. In other words, because in the complete data case $\mathbf{P} = \mathbf{I}$, then in the incomplete data case the row in \mathbf{P} corresponding to a missing measurement is removed, resulting in a nondiagonal matrix.

The system model in the case of incomplete data is therefore

$$\mathbf{Y} = \mathbf{P}\mathbf{u} + \mathbf{Z}, \quad (5)$$

in which $\mathbf{Y}, \mathbf{Z} \in \mathbb{R}^n$ are random variables. \mathbf{P} is the known measurement matrix. The vector \mathbf{u} is the noiseless time-domain signal. In this case, $\mathbf{V}_Z = \mathbf{P}\boldsymbol{\Gamma}\mathbf{P}^\top$ is not Toeplitz. We emphasize that the missing measurements simply do not appear in equation (5).

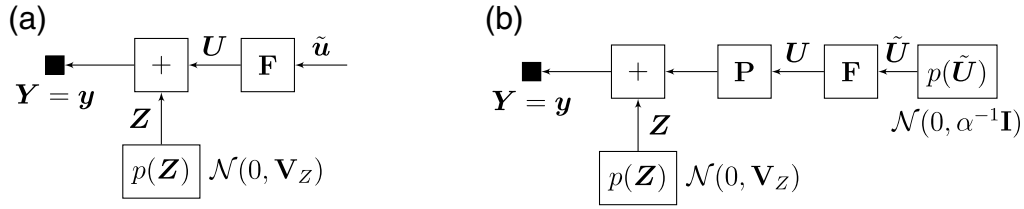


Figure 1. Factor graphs for the (a) complete data and (b) incomplete data cases. The graphs reflect equations (3) and (5). See text for description of all the mathematical quantities involved. (a) Factor graph of $\mathbf{Y} = \mathbf{F}\tilde{\mathbf{u}} + \mathbf{Z}$. (b) Factor graph of $\mathbf{Y} = \mathbf{P}\mathbf{F}\tilde{\mathbf{u}} + \mathbf{Z}$.

Signal Model

In this section, we relate the noiseless time-domain signal $u(t)$ with the frequency-domain spectrum $\tilde{u}(\omega)$. The spectrum may be parametrized with $\boldsymbol{\eta}$, that is, $\tilde{u}(\omega) = \tilde{u}(\omega, \boldsymbol{\eta})$. With $\boldsymbol{\eta} = (\boldsymbol{\theta}, \boldsymbol{\phi})^\top$, in which $\boldsymbol{\theta}$ parametrizes the Fourier amplitudes and $\boldsymbol{\phi}$ is a vector of M_i elements representing the phase spectrum. The real signal $u(t)$ is sampled at N uniform time instants with sampling interval T_s . We consider the spectrum of the signal at frequencies $\omega_m = 2\pi m/NT_s$ for $m = 0, \dots, M_r$. We have that $M_r = \lfloor N/2 \rfloor + 1$ and $M_i = N - M_r$, with $\lfloor x \rfloor$ the integer part of x . Because of Hermitian symmetry, only positive frequencies are modeled.

The noiseless signal \mathbf{u} is related with its spectrum $\tilde{\mathbf{u}}$ via a linear transformation $\mathbf{F} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ as

$$\mathbf{u} = \mathbf{F}\tilde{\mathbf{u}}. \quad (6)$$

The effect of the matrix \mathbf{F} is akin to an inverse Fourier transform. The m th column of \mathbf{F} is

$$\frac{2}{NT_s} [\cos(\omega_m t_1), \cos(\omega_m t_2), \dots, \cos(\omega_m t_N)]^\top \quad \text{if } m \in [1, M_r] \quad (7)$$

$$-\frac{2}{NT_s} [\sin(\omega_m t_1), \sin(\omega_m t_2), \dots, \sin(\omega_m t_N)]^\top \quad \text{if } m \in [M_r + 1, N]. \quad (8)$$

The first column ($m = 1$) and, when N is even, the column $m = M_r$ are further normalized by a factor of $1/\sqrt{2}$. The normalization by T_s ensures that the amplitudes of the discrete spectrum correspond to the amplitudes of the continuous Fourier transform. Given the construction of \mathbf{F} , the vector $\tilde{\mathbf{u}}$ is arranged so that the first M_r entries correspond to the real part of the inverse discrete Fourier transform (DFT) of \mathbf{u} and the last M_i entries to the imaginary part. This construction also ensures that the inverse is easily found as

$$\mathbf{F}^{-1} = \frac{NT_s^2}{2} \mathbf{F}^\top. \quad (9)$$

Such design of \mathbf{F} enables us to avoid complex numbers in our derivation and dealing with complex Gaussian distributions. In practice, \mathbf{F} can be implemented exploiting the fast Fourier transform (FFT) algorithm.

Proposed Approach

In this section, we outline the proposed approach for both the complete data case and the incomplete data case. Some details regarding the implementation, necessary in order to be able to perform the computations efficiently, are discussed in the [Appendix](#). The main hindrance being that for a large number of samples N , naive implementation of certain matrix operations becomes slow or prohibitive.

In [Figure 1](#), a factor graph provides a graphical depiction of the quantities described in this section. A factor graph is a graphical representation of the joint probability density function (PDF) of all the quantities of interest ([Loeliger et al., 2007](#)). Boxes represent joint PDFs. The edges connecting the boxes represent random variables, either observed or unknown. A random variable appears as an argument of a certain PDF if, and only if, the corresponding edge is connected to the corresponding box. The sum-product algorithm is a popular algorithm for solving inference tasks on factor graphs ([Kschischang et al., 2001](#); [Loeliger et al., 2007](#)).

In the following, we do not discuss the estimation of the covariance matrix \mathbf{V}_Z . Many approaches for the estimation of covariance matrices exist. In this work, we simply consider the sample covariance matrix computed from a sufficiently long portion of noise observed before the earthquake's first arrival.

Complete Data

In the complete data case, all N samples of the signal are observed and the noise covariance matrix is $\mathbf{V}_Z = \mathbf{I}$. The measurement matrix \mathbf{P} will not appear in this section pertaining to complete data.

We are interested in computing the likelihood of measurements $\mathbf{Y} = \mathbf{y}$ for a given nonparametric spectrum $\tilde{\mathbf{u}}$. From equations (3) and (6), the conditional PDF of \mathbf{Y} given $\tilde{\mathbf{u}}$ is

$$p(\mathbf{Y}|\tilde{\mathbf{u}}) = \mathcal{N}(\mathbf{Y}; \mathbf{F}\tilde{\mathbf{u}}, \mathbf{V}_Z). \quad (10)$$

Using a given measurement \mathbf{y} of \mathbf{Y} , it is possible to compute the likelihood of the observations $p(\mathbf{Y} = \mathbf{y}|\tilde{\mathbf{u}})$ from equation (10). Using basic probability rules or message passing on the factor graph of [Figure 1a](#), we can rewrite the likelihood in a more convenient form as

$$p(\mathbf{Y} = \mathbf{y}|\tilde{\mathbf{u}}) = \beta \mathcal{N}(\mathbf{F}^{-1}\mathbf{y}; \tilde{\mathbf{u}}, \mathbf{F}^{-1}\mathbf{V}_Z\mathbf{F}^{-\top}), \quad (11)$$

in which β is a constant. This latter expression enables us to evaluate the likelihood function for a different $\tilde{\mathbf{u}}$ without the

Table 1
Algorithms for Parametric Spectrum Fitting, Complete and Incomplete Data Case

Algorithm 1 (complete data)	
1: $\hat{\mathbf{u}}_{\text{ML}} = \arg \max_{\tilde{\mathbf{u}}} p(Y = y \tilde{\mathbf{u}})$	Find ML estimate of spectrum $\tilde{\mathbf{u}}$
2: $\hat{\boldsymbol{\phi}} = \angle \hat{\mathbf{u}}_{\text{ML}}$	Get initial phases estimate
3: repeat	
4: $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} p(Y = y \boldsymbol{\theta}, \hat{\boldsymbol{\phi}})$	Find best fitting $\hat{\boldsymbol{\theta}}$ by maximizing (11)
5: $\hat{\boldsymbol{\phi}} = \arg \max_{\boldsymbol{\phi}} p(Y = y \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})$	Refine estimates of $\boldsymbol{\phi}$ and improve fit
6: until Convergence criterion is satisfied	
Algorithm 2 (incomplete data)	
1: $\hat{\alpha}_{\text{ML}} = \arg \max_{\alpha} p(Y = y \alpha)$	Find ML estimate of regularization constant α
2: $\tilde{\mathbf{u}}_{\text{MAP}} = \arg \max_{\tilde{\mathbf{u}}} p(\tilde{\mathbf{u}} Y = y, \hat{\alpha}_{\text{ML}})$	Find MAP estimate of spectrum $\tilde{\mathbf{u}}$
3: $\hat{\boldsymbol{\phi}} = \angle \tilde{\mathbf{u}}_{\text{MAP}}$	Get initial phases estimate
4: repeat	
5: $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} p(Y = y \boldsymbol{\theta}, \hat{\boldsymbol{\phi}})$	Find best fitting $\hat{\boldsymbol{\theta}}$ by maximizing (27)
6: $\hat{\boldsymbol{\phi}} = \arg \max_{\boldsymbol{\phi}} p(Y = y \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})$	Refine estimates of $\boldsymbol{\phi}$ and improve fit
7: until Convergence criterion is satisfied	

MAP, maximum *a posteriori*.

need of evaluating the matrix–vector product $\mathbf{F}\tilde{\mathbf{u}}$ appearing in equation (10). From equation (11), it is clear that the ML estimate of $\tilde{\mathbf{u}}$ is simply

$$\tilde{\mathbf{u}}_{\text{ML}} = \mathbf{F}^{-1}\mathbf{y}, \quad (12)$$

which is closely related to the DFT of the measurements \mathbf{y} . We observe that this result does not depend on the noise covariance matrix \mathbf{V}_Z .

We now consider the parametric form of the spectrum $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(\boldsymbol{\eta})$. Because the likelihood is invariant to reparameterization, the likelihood of the observation for the parameter vector $\boldsymbol{\eta}$ can be computed from equation (11) as

$$\ell(\boldsymbol{\eta}) = p(Y = y|\tilde{\mathbf{u}}(\boldsymbol{\eta})). \quad (13)$$

An ML estimate of $\boldsymbol{\eta}$ is found by maximizing the likelihood

$$\boldsymbol{\eta}_{\text{ML}} = \arg \max_{\boldsymbol{\eta}} \ell(\boldsymbol{\eta}). \quad (14)$$

Such maximization is, in general, nonlinear and depends on the parametric spectrum model. ML estimation from equation (14) requires a joint maximization over $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. In our implementation, we choose to split equation (14) into two separate maximizations: the first over $\boldsymbol{\theta}$ and the second over $\boldsymbol{\phi}$. Although this choice can lead to suboptimal results, it makes it easier to change the parametrization of the spectrum in the implementation. These two optimizations can be cyclically repeated until a convergence criterion is satisfied. A sensible initial value for the $\boldsymbol{\phi}$ maximization can be obtained from equation (12). When the noise covariance matrix is diagonal, this initial guess is exact.

From equation (13), it can be shown that the ML estimation is equivalent to a (nonlinear) generalized least-squares fit

$$\hat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta}} (\mathbf{y} - \mathbf{F}\tilde{\mathbf{u}}(\boldsymbol{\eta}))^T \mathbf{V}_Z^{-1} (\mathbf{y} - \mathbf{F}\tilde{\mathbf{u}}(\boldsymbol{\eta})). \quad (15)$$

This relation indicates the similarity between the proposed ML method and least-squares spectral analysis methods (Lomb, 1976; Scargle, 1982).

Algorithm 1 in Table 1 summarizes a possible implementation of the proposed approach. The spectrum $\tilde{\mathbf{u}}_{\text{ML}}$ is computed and an initial estimate $\hat{\boldsymbol{\phi}}$ for the phases is obtained from it. The likelihood function is maximized, considering as optimization variable only $\boldsymbol{\theta}$, and using the initial guess for the phases. This optimization can, in general, be performed numerically using, for example, gradient descent methods. Whether the numerical optimization is able to achieve the optimal solution depends on the specific functional form of the parametric spectrum and, possibly, on an initial guess for the parameter vector $\boldsymbol{\theta}$. In addition, when the parametric spectrum is known analytically, a more sophisticated optimization can be devised. After an ML estimate of $\boldsymbol{\theta}$ is obtained, phases are optimized. This latter optimization can also be performed numerically, perhaps supplying an explicit expression of the gradient to the numerical optimization routine. The two latter optimization steps can be repeated cyclically to improve the fit.

Incomplete Data

In the presence of gaps, we observe only $n < N$ samples. This is modeled with the measurement matrix $\mathbf{P} \in \mathbb{R}^{n \times N}$ as appeared in equation (5). First, we seek an estimate of the nonparametric spectrum $\tilde{\mathbf{u}}$. This is needed to obtain an initial estimate of the signal phases. This is an underdetermined problem. Second, we consider the estimation of parameters for a parametric spectrum as in the previous section.

Nonparametric: Maximum *A Posteriori* Estimation

The estimation of the spectrum $\tilde{\mathbf{u}} \in \mathbb{R}^N$ from observations $\mathbf{y} \in \mathbb{R}^n$ is an underdetermined problem. We choose to treat the

spectrum as a random variable \tilde{U} , in contrast with the previous section in which the spectrum was treated as an unknown deterministic value. We introduce a prior distribution on \tilde{U} in order to regularize the underdetermined problem.

The noise covariance matrix is $\mathbf{V}_Z = \mathbf{W}_Z^{-1} = \mathbf{P}\mathbf{P}^\top$. We observe that, in contrast with the complete data case, this is not a Toeplitz matrix. We treat the spectrum as a random variable \tilde{U} with a Gaussian prior $p(\tilde{u})$

$$\tilde{U} \sim \mathcal{N}(\mathbf{0}, \alpha^{-1}\mathbf{I}), \quad (16)$$

in which α is a regularization parameter to be determined. Using Bayes theorem, the likelihood of equation (10), and the prior of equation (16) we find that the posterior probability is

$$p(\tilde{u}|Y = y, \alpha) \propto p(Y = y|\tilde{u})p(\tilde{u}|\alpha) \quad (17)$$

$$= \gamma_{\tilde{u}} \exp(-\tilde{u}^\top \mathbf{W}_{\tilde{u}} \tilde{u} / 2 + \tilde{u}^\top \mathbf{W}_{\tilde{u}} \mathbf{m}_{\tilde{u}}), \quad (18)$$

in which

$$\gamma_{\tilde{u}} = \mathcal{N}(\mathbf{0}; \mathbf{m}_{\tilde{u}}, \mathbf{W}_{\tilde{u}}^{-1}) \mathcal{N}(\mathbf{0}; \mathbf{0}, \alpha^{-1}\mathbf{I}) \quad (19)$$

$$\mathbf{W}_{\tilde{u}} \mathbf{m}_{\tilde{u}} = \mathbf{F}^\top \mathbf{P}^\top \mathbf{W}_Z \mathbf{y} \quad (20)$$

$$\mathbf{W}_{\tilde{u}} = \mathbf{F}^\top \mathbf{P}^\top \mathbf{W}_Z \mathbf{P} \mathbf{F} + \alpha \mathbf{I}. \quad (21)$$

For a given α , the maximum *a posteriori* (MAP) estimate of \tilde{u} is

$$\tilde{u}_{\text{MAP}} = \mathbf{W}_{\tilde{u}}^{-1} \mathbf{W}_{\tilde{u}} \mathbf{m}_{\tilde{u}}. \quad (22)$$

Observe that for $\mathbf{P} = \mathbf{I}$ and $\alpha = 0$, the above estimate coincides with the estimate in equation (12).

The value of α to use in equation (22) can be estimated from the data with the ML criterion. The likelihood of the observations as a function of α is

$$p(Y = y|\alpha) = \int p(Y = y, \tilde{u}|\alpha) d\tilde{u} \quad (23)$$

$$= \int p(Y = y|\tilde{u}) p(\tilde{u}|\alpha) d\tilde{u} \quad (24)$$

$$= \mathcal{N}\left(\mathbf{y}; \mathbf{0}, \mathbf{V}_Z + \frac{2}{\alpha N T_s^2} \mathbf{I}\right). \quad (25)$$

By maximizing equation (25), we can obtain an ML estimate $\hat{\alpha}_{\text{ML}}$ of α to use in equation (22).

The observed MAP estimation of \tilde{u} is equivalent to a generalized least-squares fit with ridge regression regularization:

$$\tilde{u} = \arg \min_{\tilde{u}} (\mathbf{y} - \mathbf{P}\mathbf{F}\tilde{u})^\top (\mathbf{P}^\top \mathbf{W}_Z \mathbf{P})^{-1} (\mathbf{y} - \mathbf{P}\mathbf{F}\tilde{u}) + \alpha \|\tilde{u}\|^2. \quad (26)$$

Parametric: ML Estimation

ML estimation of $\boldsymbol{\eta}$ in the incomplete case is analogous to the complete case. To find initial phase estimates, $\hat{\boldsymbol{\phi}}$ is necessary to find the MAP estimate of \tilde{u} . In addition, because the covariance matrix $\mathbf{V}_Z = \mathbf{P}\mathbf{P}^\top$ is no longer Toeplitz, some simplification may be beneficial in the numerical implementation. These details are described in the [Appendix](#).

From equations (3) and (6), the conditional PDF of \mathbf{Y} given $\boldsymbol{\eta}$ is

$$p(Y|\boldsymbol{\eta}) = \mathcal{N}(\mathbf{Y}; \mathbf{P}\mathbf{F}\tilde{u}(\boldsymbol{\eta}), \mathbf{V}_Z). \quad (27)$$

An ML estimate of $\boldsymbol{\eta}$ is found by maximizing the likelihood $\arg \max_{\boldsymbol{\eta}} p(\mathbf{Y} = \mathbf{y}|\boldsymbol{\eta})$.

Algorithm 2 in Table 1 summarizes a possible implementation of the proposed approach for the incomplete data case. First, an appropriate value for the regularization constant α is found with the ML criterion. Using this value, we find an MAP estimate of the nonparametric spectrum and then an initial estimate for the phases. Similar to the complete data case, an ML estimate for $\boldsymbol{\theta}$ is found, keeping the phases fixed to the initial guess. After an ML estimate of $\boldsymbol{\theta}$ is obtained, phases are optimized.

Numerical Examples and Applications

In this section, we present different examples and applications to show the capabilities of the method presented in this article. We consider fitting a simple parametric spectrum to earthquake recordings using the ML technique described in the previous section.

We choose to use the spectral model proposed in [Brune \(1970\)](#), augmented by the attenuation factor from [Anderson and Hough \(1984\)](#), to fit the shape of earthquake spectra. We do not consider a factor to account for geometrical spreading, rather we simply fit the observed spectrum at a given epicentral distance. To simplify the analysis, we also neglect the presence of site effects.

With these assumptions, the amplitude of the velocity spectrum is described by

$$|\tilde{u}(\omega, \boldsymbol{\theta})| = \frac{\omega \Omega}{1 + (\omega/2\pi f_c)^2} e^{-\omega t^* / 2} \quad (28)$$

with parameter vector $\boldsymbol{\theta} = (\Omega, f_c, t^*)^\top$. The parameters Ω , f_c , and t^* represent the long-period spectral displacement plateau, the source corner frequency, and the anelastic attenuation, respectively. Equation (28) only specifies the amplitude spectrum, not the phase spectrum.

It is important to note that the approach proposed in this article is not limited to fitting the Brune model of equation (28) but can be used with any other spectral model. We consider the Brune model merely for the sake of exposition. The Brune model has been extensively used in the literature, has a simple parametrization, and is easy to understand. Our approach could also be used to determine other parameters

influencing the spectral shape, for example, by using other source models (Savage, 1972; Boatwright, 1982), site effects (Drouet *et al.*, 2008), and can also form the basis of other analyses, such as determining geometrical spreading (Edwards *et al.*, 2008).

An estimate for the parameter vector θ is found using Algorithms 1 and 2 in Table 1, outlined in the [Proposed Approach](#) section. We fit the Brune model to the recording from a single horizontal component.

We consider recordings from the Swiss Seismic Networks (SDSNet and SSMNet) of the 15 October 2014 Diemtigen (M_w 3.0) earthquake (see also [Data and Resources](#)). The epicenter was located at 46.65° N and 7.57° E. The depth of the hypocenter is 9.5 km. The origin time is 19:36:32 UTC. Different stations are used, some with high SNR, some with low SNR. Both velocimeter and accelerometer recordings are used. Figure 2 shows the position of the stations and the position of the epicenter within Switzerland.

The signal plus noise-to-noise ratio (SNNR) is computed as the ratio between the power of the event over the power of the pre-event. The SNNR and SNR are related as

$$\text{SNNR} = \frac{S + N}{N} = \text{SNR} + 1. \quad (29)$$

Complete Data

Example

As a first example, we consider fitting the Brune model of equation (28) to a complete seismogram using the procedure described in the [Complete Data](#) section. The seismogram is a recording of the Diemtigen earthquake at the SBAT HG strong-motion station. The results are summarized in Figure 3. Figure 3a,b compares the observed seismogram and the ML fit in the time domain. As discussed in the [Proposed Approach](#) section, the likelihood is equivalent to a least-squares fit in the time domain. The residual error between observed seismogram and fit is weighted by the covariance matrix.

Figure 3c compares the DFT of the spectrum with the ML estimate of the parametric spectrum. The parametric spectrum matches well with the DFT of the signal. We emphasize that the high-frequency noise visible between 30 and 55 Hz does not appear to significantly affect the parametric fit, despite the fact that all frequencies are effectively used in the procedure. Robustness to noise is provided by the noise covariance matrix \mathbf{V}_Z . The matrix \mathbf{V}_Z is estimated using a portion of noise before the event, and is not shown here.

Figure 3d shows slices of the log-likelihood function (13) as a function of θ . The ML estimate is pinpointed with a white cross, that is, θ_{ML} . Because the log likelihood is proportional to the least-squares misfit, this plot quantifies the trade-off between the model parameters. Instead of Ω , the variation in estimated magnitude is shown, that is, $\Delta M_w = (2/3) \log(\Omega/\Omega_{\text{ML}})$, in which Ω_{ML} is the ML estimate of Ω . For a reliable estimate of the magnitude it would

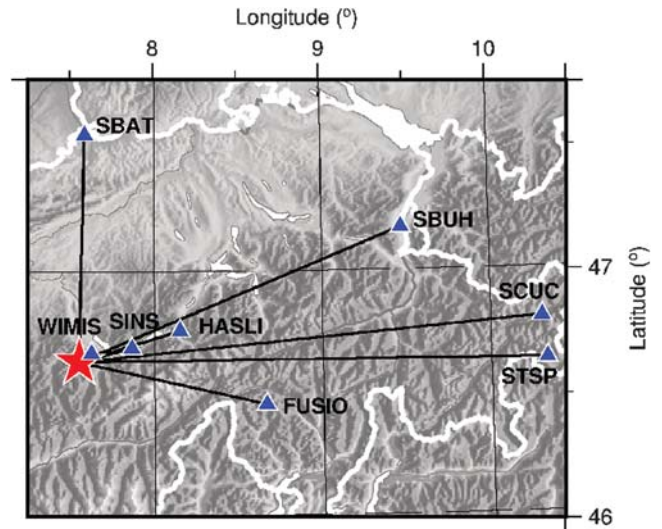


Figure 2. Map of the stations (blue triangles) used in this study and location of the epicenter (red star).

be necessary to model factors such as the geometrical spreading. For this reason, we display the relative change ΔM_w rather than the absolute value M_w . Although trade-offs are evident in all parameters, we note that trade-offs between M_w and f_c or t^* are emphasized due to the scales used: relatively small differences of 0.1–0.2 in M_w are dwarfed by differences of factors of over 2 in f_c and t^* .

Parametric Fit with Modern Seismograms

We now consider seismograms from the Diemtigen event recorded at different stations. We compare the ML method proposed in this article with an approach used in the literature, applied to individual seismograms.

The comparison method used is based on Edwards *et al.* (2011) and Edwards and Fäh (2013). We perform an inversion for Brune Fourier amplitude spectrum (FAS) model parameters t^* , Ω (long-period displacement plateau), and f_c (source corner frequency). The inversion is achieved using a combined grid-search for f_c (at 5% resolution over a range of values broadly corresponding to stress drops of 0.001–1000 MPa) and Powell's conjugate direction method for Ω and t^* . A least-squares minimization of the misfit between the Brune model to the available data above a threshold SNR of 3 is used in the log-log domain to simultaneously fit the parameters controlling both low- (Ω , f_c) and high-frequency FAS amplitudes (t^* , f_c).

Figures 4–6 depict (a) the seismogram, including both the pre-event portion used to estimate the noise covariance matrix and the earthquake portion used to compute the fit; and (b) the earthquake and noise spectra along with parametric fits of the Brune model from the proposed method and a comparison method. The comparison method uses only frequencies with an SNR greater than 3 and performs a non-linear least-squares fit in the frequency domain. The line depicting the comparison method is solid for the frequencies

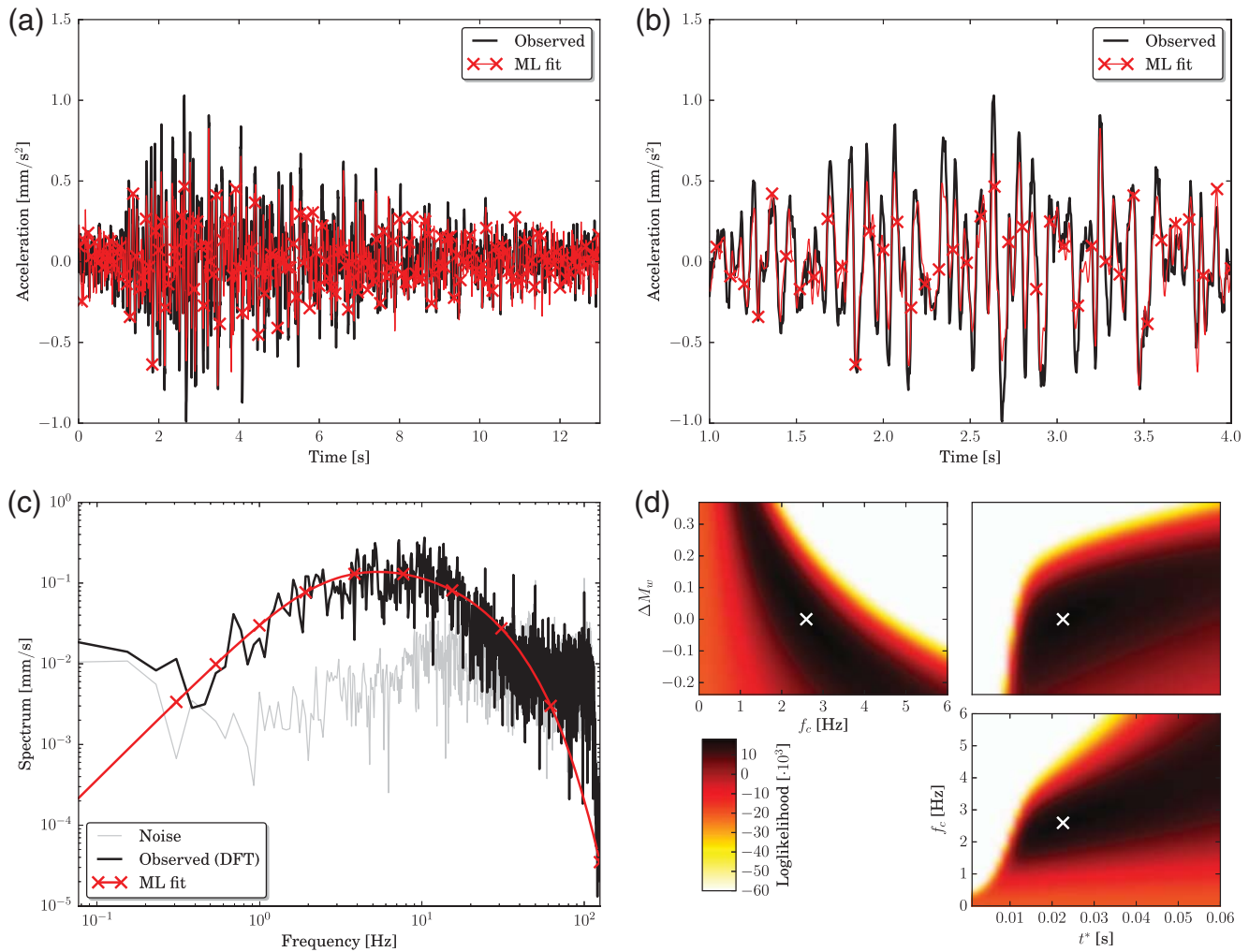


Figure 3. Brune model fit for complete data case. (a) Observed seismogram and maximum-likelihood (ML) fit. (b) Detail of 3 s of the seismogram. (c) Comparison of the fitted parametric spectrum with the discrete Fourier transform (DFT) of the seismogram. (d) Slices of the log-likelihood function, the white cross pinpoints the ML estimate.

with SNR above the threshold and thus used to compute the fit, dashed for frequencies below the threshold and not used by the algorithm. For the proposed ML approach, there is no such distinction because all the frequencies are used.

In this section, we present results obtained from the analysis of the north components. The portion of signal used to compute the fit is determined visually with the aim of including most of the energy from the *S*-wave arrival. For very low-SNR seismograms, we inspected a band-pass filtered version of the signal, but the unfiltered seismogram is used to compute the fit. In selecting the portion used to estimate the noise, we tried to avoid including any anomalous features. In Figures 4–6, event and pre-event portions of the trace are marked with solid and dashed red lines, respectively.

Figure 4 depicts the results obtained for three stations with high SNR. The station WIMIS and HASLI are located at rock sites. The station SINS lies on soft sediments known to produce broadband amplifications. The outcomes of the proposed ML method and the comparison method are very similar. In

fact, the noise level is significantly smaller than the energy from the earthquake at almost all frequencies. Therefore all the frequencies are reliable, the noise-aware scheme of the proposed method is unnecessary and brings no improvement over the comparison method. The small differences in the obtained fits of the Brune model can be explained with the different misfit functions used by the two methods. Observe that at certain stations the ground velocity is measured, whereas at others, the acceleration. Also for this reason, among the stations, the noise spectra are greatly different. In Figure 4b, the high frequencies are dominated by the internal noise of the sensor. The fit resulting from the proposed method does not appear to be affected by high values at these frequencies.

Figure 5 depicts the results obtained for three stations with high-to-intermediate SNR. The station FUSIO is located at a rock site. The station SBAT is on consolidated soil and is known to exhibit broadband amplification. We observe that, for the station SBAT, the comparison method fails to achieve a proper fit and would normally be discarded. In fact,

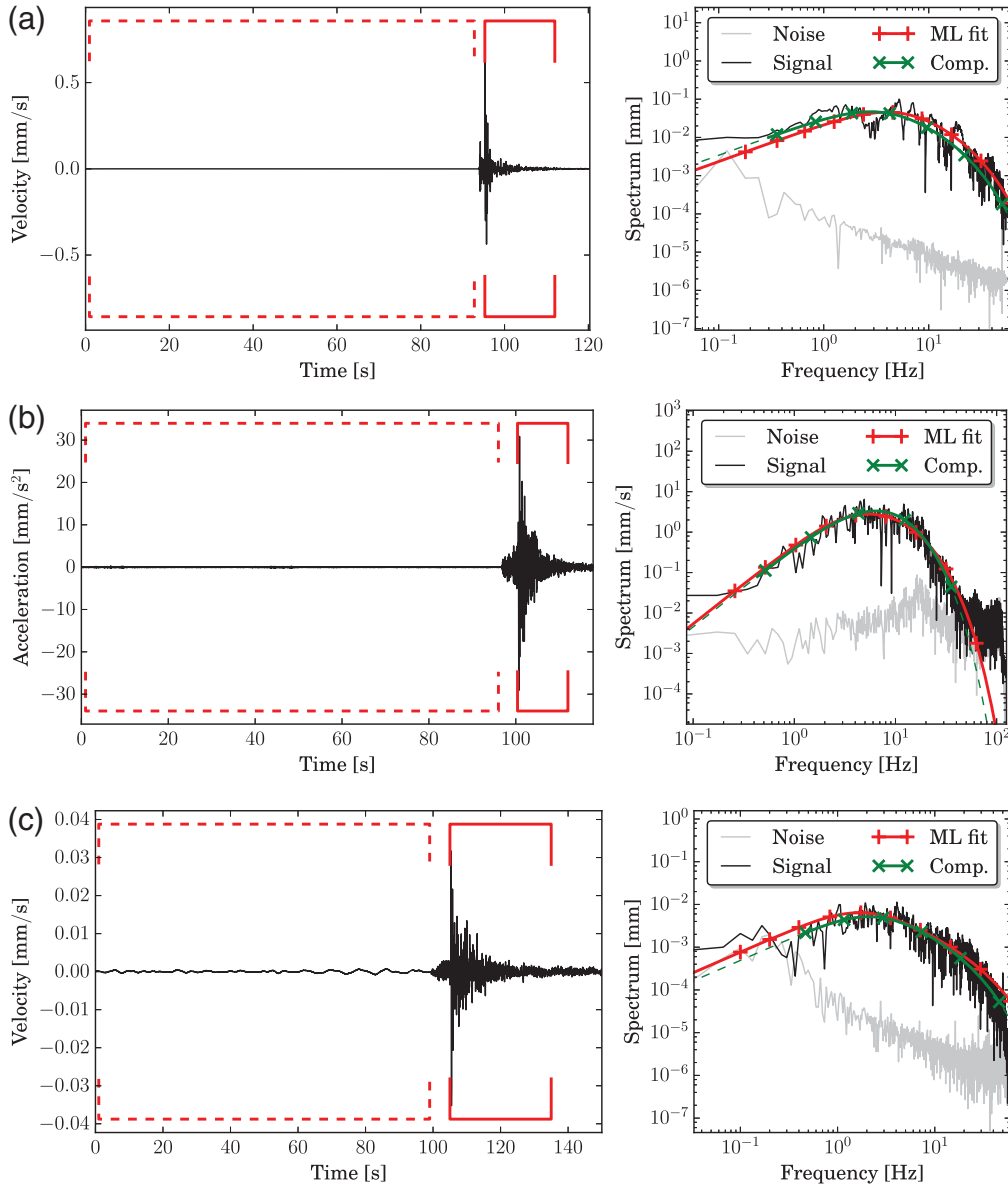


Figure 4. Seismograms and spectra from three high signal-to-noise ratio (SNR) stations, north components. (a) Station WIMIS HH, distance 4 km, SNR = 45 dB. (b) Station SINS HG, distance 23 km, SNR = 40 dB. (c) Station HASLI HH, distance 46 km, SNR = 23 dB.

the frequency band in use is too narrow to successfully determine the model parameters (compare to Fig. 5c).

Figure 6 depicts the results obtained for three stations with very low SNR. These stations are chosen because they are the farthest stations from the epicenter. The stations SBUH and STSP are located over soft soil, whereas SCUC is located over a thin layer of rather compact material. In these settings, the comparison method could not achieve a meaningful result. In fact, the SNR is extremely low at most frequencies, and the frequency band with large enough SNR is narrow or nonexistent. In contrast, our method is able to fit the spectra properly.

Table 2 summarizes parameters for both the proposed ML method and the comparison method. Estimated parameters are Ω , f_c , and t^* . The parameters f_{\min} and f_{\max} are

provided to the comparison method and correspond to the frequency range actually used for the fit (i.e., the frequency range with SNR greater than 3). Stations are sorted according to decreasing SNR as in Figures 4–6.

Incomplete Data

Example

We now consider fitting the Brune model of equation (28) to an incomplete seismogram as described in the [Incomplete Data](#) section. We use the recording from the station SBAT HG used in the previous section. We remove the data in correspondence of two gaps of a different duration and fit the Brune model to such a seismogram. The results are summarized in Figure 7. Figure 7a,b compares the observed seismogram

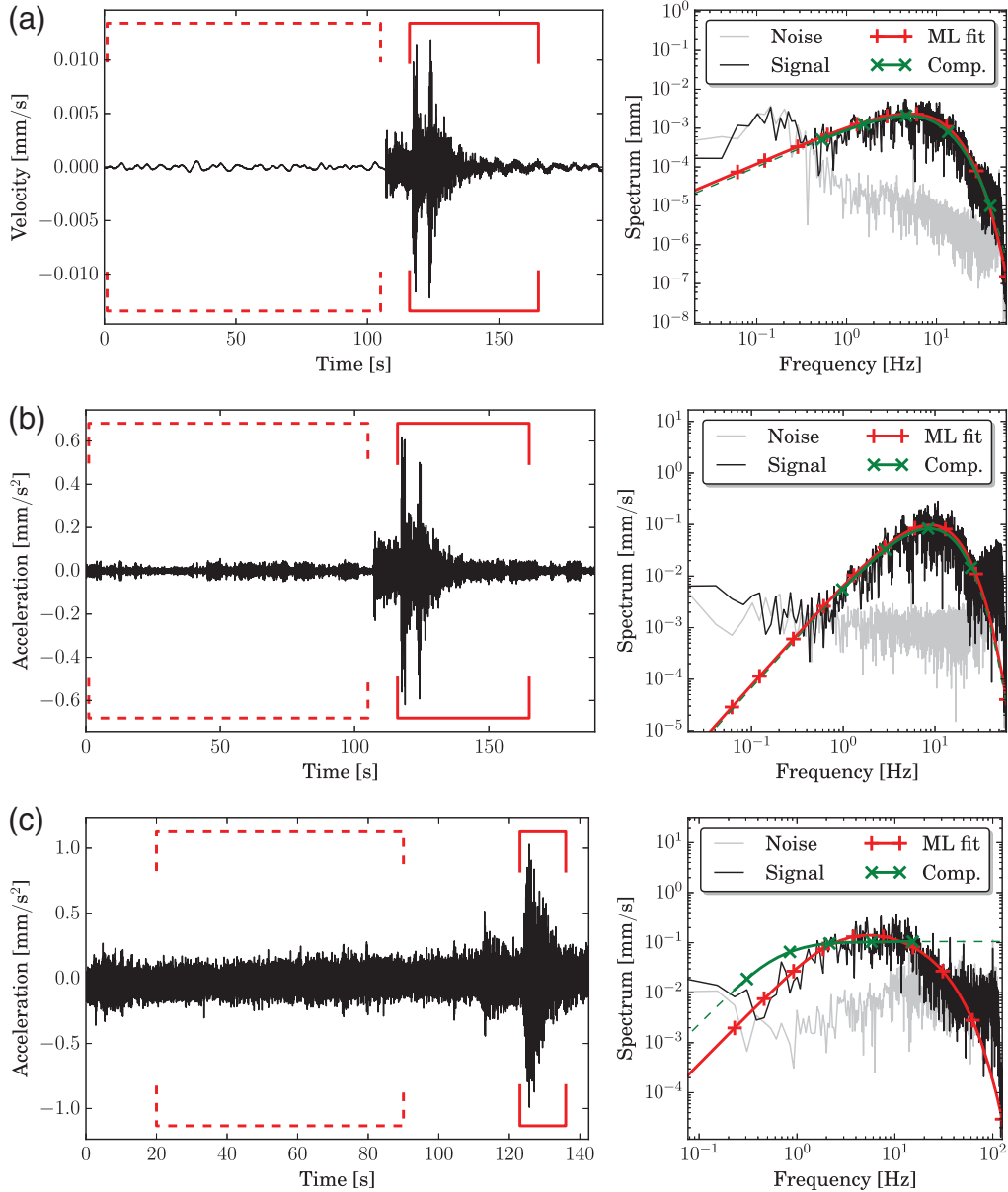


Figure 5. Seismograms and spectra from three intermediate SNR stations, north components. (a) Station FUSIO HH, distance 86 km, SNR = 19 dB. (b) Station FUSIO HG, distance 86 km, SNR = 15 dB. (c) Station SBAT HG, distance 101 km, SNR = 10 dB.

and the ML fit in the time domain. The missing data, unknown to the algorithm, are depicted with a thin gray line.

Figure 7c shows the result of the fit in the frequency domain. The nonparametric MAP estimate of the spectrum obtained from the incomplete seismogram is compared with the DFT of the complete seismogram. Some differences between the two spectra are visible and are explained with the presence of the gaps. The ML estimate of the parametric spectrum is also shown in Figure 7c, both for the complete data case (same as Fig. 3c) and for the incomplete data case. The parametric spectrum obtained in the presence of gaps is in good agreement with the parametric spectrum obtained in the complete data case. It also matches well with the non-

parametric spectrum. Again high-frequency noise does not appear to significantly affect the parametric fit. Figure 7d shows slices of the log-likelihood function from equation (27). The ML estimate for the complete data case is pinpointed with a white cross (same as Fig. 3d) and the estimate for the incomplete data case is pinpointed with a white dot.

Seismogram Reconstruction

We now show how it is possible to fill the gaps using the parameters estimated in the ML method.

Using the ML estimate of the parameter vector $\hat{\eta} = (\hat{\theta}, \hat{\phi})^\top$ (compare with equation 27), we suggest to compute an estimate \hat{y} of the complete seismogram as

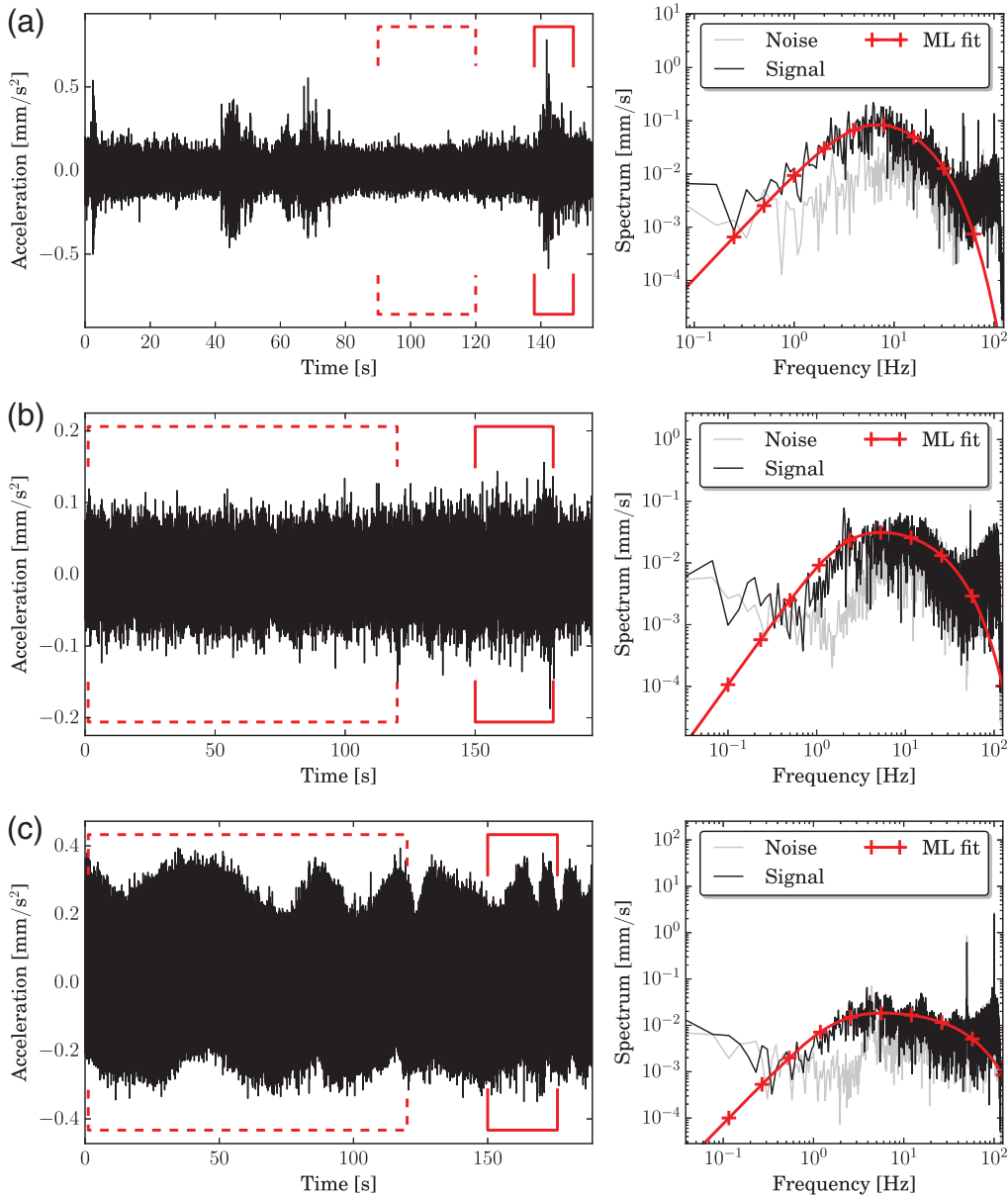


Figure 6. Seismograms and spectra from three low SNR stations, north components. (a) Station SBUH HG, distance 156 km, SNR = 7 dB. (b) Station STSP HG, distance 211 km, SNR = -4 dB. (c) Station SCUC HG, distance 209 km, SNNR = -0.19 dB. SNNR, signal plus noise-to-noise ratio.

$$\hat{y} = \mathbf{F}\tilde{u}(\hat{\eta}). \quad (30)$$

That is, we compute an inverse Fourier transform of the estimated spectrum. The estimated spectrum $\tilde{u}(\hat{\eta})$ relies on the estimate of the parameter $\hat{\theta}$ controlling the spectral amplitude and the estimate of the phases $\hat{\phi}$ best fitting the observed seismogram.

Figure 8a,b shows the reconstructed signal for the same setting of the incomplete data example presented in this section. The reconstructed signal, shown with a solid green line, does not closely match the original unobserved signal, shown with a solid gray line; the reconstructed seismogram accurately preserves the frequency content although it does not replicate isolated peak amplitudes.

Impact of Gaps on Estimated Parameters

In this section, we evaluate the impact of gaps on the estimated parameters. We consider some selected seismograms from the previous section and fit the Brune model of equation (28) to the incomplete seismogram.

We consider gaps of different duration and position, distributed uniformly along the entire duration of the signal. Gaps due to sensor clipping or to failure of a mechanical device most often appear around the strongest motion. Gaps due to the minute marks in historical data may appear at any position. The original seismogram is erased in correspondence of the gap and the procedure for fitting a parametric spectrum to incomplete data is applied. Gap durations of up to half the

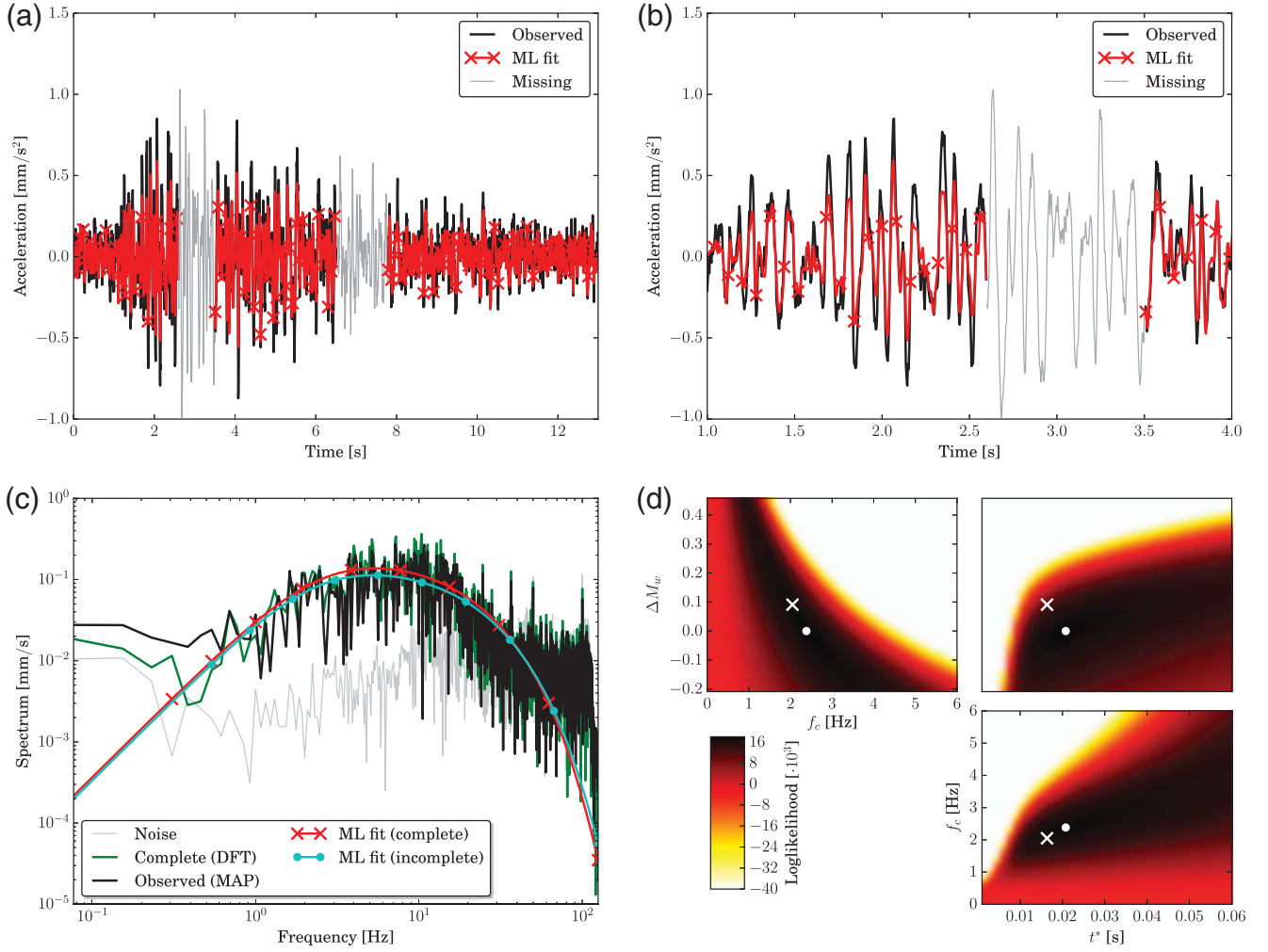


Figure 7. Brune model fit for incomplete data case. (a) Observed seismogram, unobserved seismogram, and time-domain fit. (b) Detail of three seconds of the seismogram. (c) Comparison of the fitted parametric spectrum with Fourier transform of the complete seismogram and maximum *a posteriori* (MAP) estimate of the spectrum from observed (incomplete) seismogram. (d) Slices of the log-likelihood function, the white cross and the white dot pinpoint the ML estimate for the complete and incomplete data case, respectively.

Table 2
Summary of the Estimated Parameters for the Brune Model

Stations	SNR (dB)	ML Method			Comparison Method				
		Ω (m s)	f_c (Hz)	t^* (s)	Ω (m s)	f_c (Hz)	t^* (s)	f_{\min} (Hz)	f_{\max} (Hz)
WIMIS HH	45	3.75×10^{-06}	5.23	0.021	5.48×10^{-06}	3.46	0.024	0.29	49.93
SINS HG	40	1.38×10^{-05}	4.07	0.043	1.17×10^{-05}	6.90	0.055	0.37	35.59
HASLI HH	23	1.22×10^{-06}	1.75	0.060	7.91×10^{-07}	2.34	0.017	0.41	46.00
FUSIO HH	19	1.95×10^{-07}	21.65	0.058	1.64×10^{-07}	10.62	0.045	0.51	39.77
FUSIO HG	15	2.01×10^{-07}	20.67	0.060	1.74×10^{-07}	17.29	0.059	0.95	24.70
SBAT HG	10	9.62×10^{-07}	2.58	0.023	6.12×10^{-06}	0.66	0.000	0.18	15.28
SBUH HG	7	2.72×10^{-07}	4.59	0.029			N/A		
STSP HG	-4	2.72×10^{-07}	2.08	0.016			N/A		
SCUC HG	N/A	1.92×10^{-07}	1.75	0.008			N/A		

Summary of the parameters for the Brune model (compare with Figs. 4–6). Estimated parameters for both the proposed maximum-likelihood (ML) method and the comparison method are Ω , f_c , and t^* . The parameters f_{\min} and f_{\max} define the frequency range used by the comparison method. N/A, not available. SNR, signal-to-noise ratio.

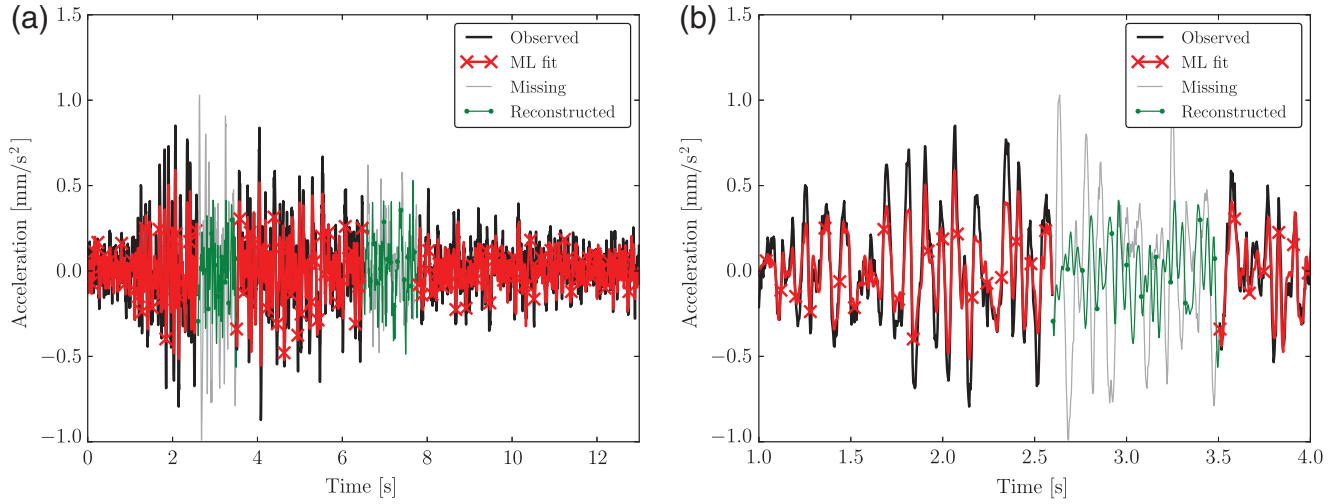


Figure 8. Reconstruction of the unobserved seismogram for the incomplete data case. (a) Observed and unobserved seismogram, time-domain fit, and reconstructed seismogram. (b) Detail of three seconds of the seismogram.

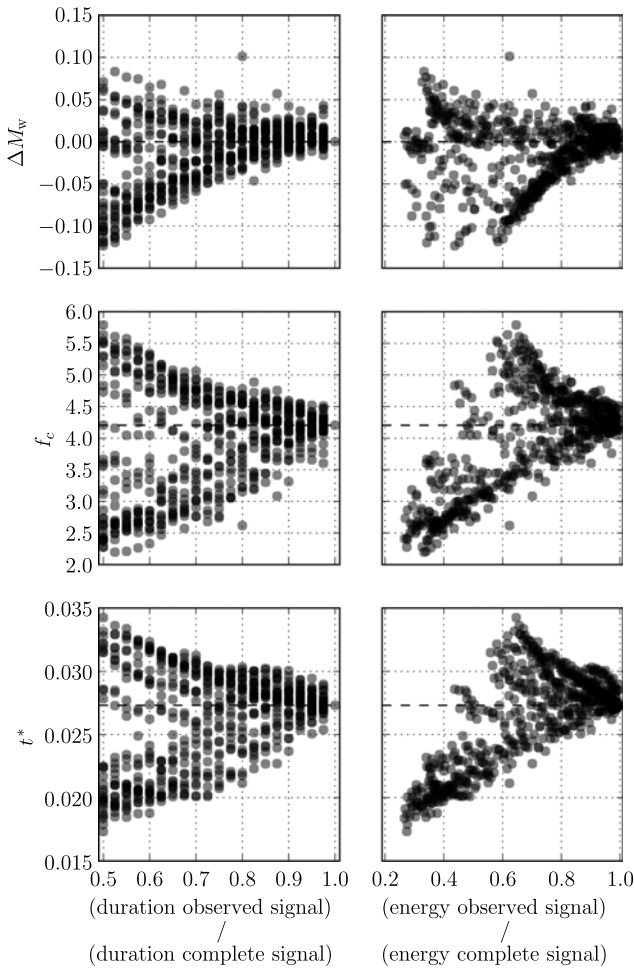


Figure 9. Estimated parameters for the incomplete data case. Station SBUH HG.

duration of the signal are considered. The portion of the seismogram used is the same as that used in the previous section and highlighted in Figures 4 and 5. The fraction

of signal removed is calculated on this selected window and not on the whole seismogram. It is evident that a gap of a given duration can remove a different amount of signal energy depending on its position. For this reason, we consider both the fraction of signal duration observed with respect to the total duration of the seismogram and the fraction of energy observed with respect to the total energy of the seismogram.

In Figures 9–12, the x axis depicts the fraction of the signal duration observed and the fraction of energy observed. The y axis depicts the variation in estimated magnitude, which is given by

$$\Delta M_w = \frac{2}{3} \log \frac{\Omega}{\Omega_0}, \quad (31)$$

with Ω_0 being the value estimated using the complete seismogram, the estimated corner frequency f_c , and the high-frequency attenuation t^* . Horizontal dashed lines indicate the reference values, estimated using the complete seismogram.

With the exception of FUSIO, there is a good agreement among the reference values estimated for f_c and t^* . The recording at FUSIO is quite unusual, exhibiting two large peaks (compare with Fig. 5b). The estimated parameters are significantly different from the other stations. This example emphasizes the limitations of single station analysis.

In general, Figures 9–12 show that removing up to 10%–15% of the signal energy produces small changes in the estimated parameters. We observe that with these gaps there is no systematic over- or underestimation of the parameters. In Figure 11, it is possible to see a peculiar pattern, in which the points appear to branch out as the observed portion of the signal is reduced. This can be explained by observing the complete seismogram that is dominated by two large distinct peaks (compare with Fig. 5b).

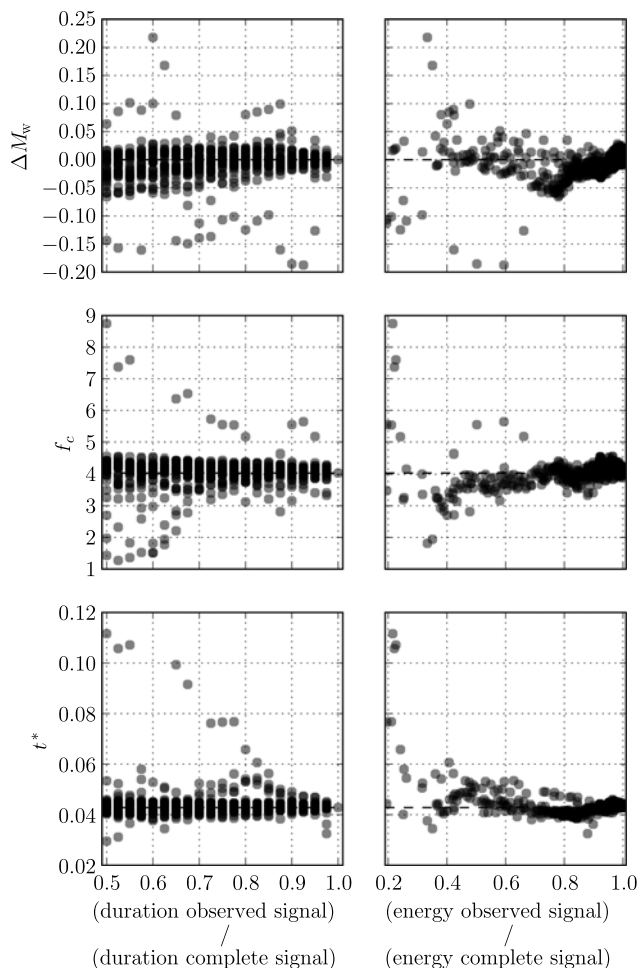


Figure 10. Estimated parameters for the incomplete data case. Station SINS HG.

Conclusions

In this article, we presented a method for the modeling of earthquake records. We presented the applicability of this method applied to fitting a parametric spectrum to earthquake recordings from modern and historical earthquake recordings. The contribution of this work is twofold: (a) we model the presence of colored noise; and (b) we account for the presence of gaps in the seismogram. In this work, we fitted the model to a single waveform from a given sensor component.

By properly modeling colored noise, we are able to weigh the contribution of the different frequencies in relation to the SNR at each frequency. The proposed method uses all the available information following the ML criterion. Instrument noise and ambient noise are naturally accounted for in the proposed framework. We validate the effectiveness of the proposed method on both modern recordings. Using digital recordings from the 15 October 2014 Diemtigen (M_w 3.0) earthquake, we have shown how the method is able to fit earthquakes with extremely low SNR. Such recordings would have been unusable by methods present in literature.

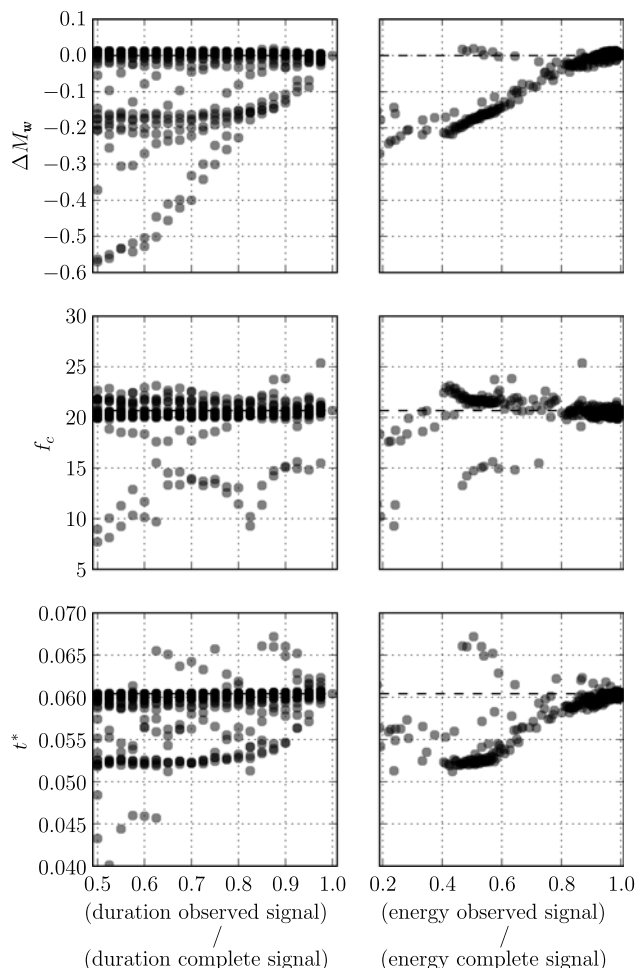


Figure 11. Estimated parameters for the incomplete data case. Station FUSIO HG.

The presence of gaps in the seismogram is a critical issue, especially in studies of historical earthquakes. In fact, the success of the analysis of a historical earthquake is hindered by the presence of seismograms with incomplete data. The proposed technique is able to fit a spectrum to a seismogram despite the presence of gaps. In the proposed method, we do not make any assumption on the missing portion of the seismogram. We tested the proposed method by erasing a part of the seismogram inserting gaps of different duration and at different positions. We observed that the estimated parameters do not greatly differ from the reference values as long as the fraction of the energy removed is less than 10%–15% of the total energy. We also presented a technique to reconstruct the unobserved seismogram preserving the frequency content of the seismogram.

Using a different spectral model, it is possible to account for more wave propagation effects. The proposed method can also be extended to multiple stations for the determination of the seismic moment. Assuming the noise at different stations is independent, the likelihood of the observations at multiple stations is simply the product of the likelihood at each station. Similarly, it is possible to combine the sensor components at a single station. We envision the need for further

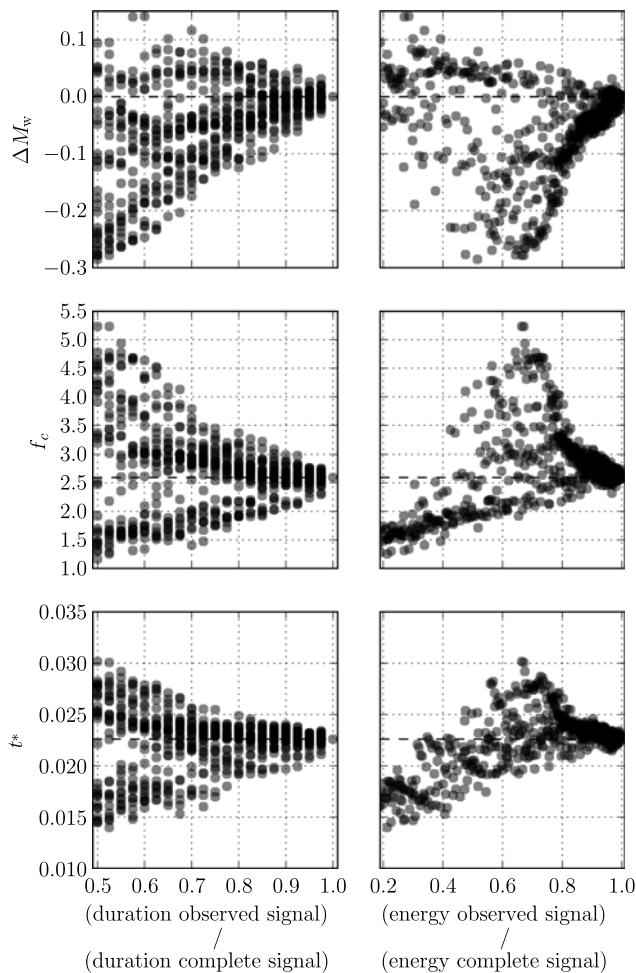


Figure 12. Estimated parameters for the incomplete data case. Station SBAT HG.

development to account for the correlation of noise among components or among different stations.

Data and Resources

Data from the recordings of the Diemtigen earthquake were provided by the Swiss Seismological Service and are publicly available at <http://arclink.ethz.ch> (last accessed October 2016). The code used to produce the results presented in this article and the code to reproduce the results will be made available online.

Acknowledgments

This work was supported by the Swiss National Science Foundation project “Advanced Single Station and Array Methods for the Analysis of Ambient Vibrations and Earthquake Recording” (200021_153633). With special thanks to Carlo Cauzzi, Jan Burjánek, and Fabrizio Bernardi for the feedback provided during the development of this work. We would also like to sincerely thank the reviewers for sharing their time and expertise in the review process. Their valuable comments helped us to improve the article.

References

- Allen, T. I., P. R. Cummins, T. Dhu, and J. F. Schneider (2007). Attenuation of ground-motion spectral amplitudes in southeastern Australia, *Bull. Seismol. Soc. Am.* **97**, no. 4, 1279–1292.
- Anderson, J. G., and S. E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high-frequencies, *Bull. Seismol. Soc. Am.* **74**, no. 5, 1969–1993.
- Atkinson, G. M. (1993). Earthquake source spectra in eastern North America, *Bull. Seismol. Soc. Am.* **83**, no. 6, 1778–1798.
- Atkinson, G. M. (2015). Ground motion prediction equation for small to moderate events at short hypocentral distances, with application to induced seismicity hazards, *Bull. Seismol. Soc. Am.* **105**, no. 2A, 981–992.
- Boatwright, J. (1982). A dynamic model for far-field acceleration, *Bull. Seismol. Soc. Am.* **72**, no. 4, 1049–1068.
- Boatwright, J., G. L. Choy, and L. C. Seekins (2002). Regional estimates of radiated seismic energy, *Bull. Seismol. Soc. Am.* **92**, no. 4, 1241–1255.
- Boatwright, J., J. B. Fletcher, and T. E. Fumal (1991). A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes, *Bull. Seismol. Soc. Am.* **81**, no. 5, 1754–1782.
- Boore, D. M. (2003). Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.* **160**, nos. 3/4, 635–676.
- Boore, D. M., W. B. Joyner, and L. Wennerberg (1992). Fitting the stochastic ω^{-2} source model to observed response spectra in western North America: Trade-offs between $\Delta\sigma$ and κ , *Bull. Seismol. Soc. Am.* **82**, no. 4, 1956–1963.
- Bora, S. S., F. Scherbaum, N. Kuehn, P. Stafford, and B. Edwards (2015). Development of a response spectral ground-motion prediction equation (GMPE) for seismic hazard analysis from empirical Fourier spectral and duration models, *Bull. Seismol. Soc. Am.* **105**, no. 4, 2192–2218.
- Brune, J. N. (1970). Tectonic stress and spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75**, no. 26, 4997–5009.
- Čadež, O. (1987). Studying earthquake ground motion in Prague from Wiechert seismograph records, *Gerland Beitr. Geophys.* **96**, no. 5, 438–447.
- Cauzzi, C., and J. Clinton (2013). A high- and low-noise model for high-quality strong-motion accelerometer stations, *Earthq. Spectra* **29**, no. 1, 85–102.
- Cotton, F., R. Archuleta, and M. Causse (2013). What is sigma of the stress drop? *Seismol. Res. Lett.* **84**, no. 1, 42–48.
- Douglas, J., B. Edwards, V. Convertito, N. Sharma, A. Tramelli, D. Kraaijpoel, B. M. Cabrera, N. Maercklin, and C. Troise (2013). Predicting ground motion from induced earthquakes in geothermal areas, *Bull. Seismol. Soc. Am.* **103**, no. 3, 1875–1897.
- Douglas, J., P. Gehl, L. F. Bonilla, and C. Gélis (2010). A kappa model for mainland France, *Pure Appl. Geophys.* **167**, no. 11, 1303–1315.
- Drouet, S., S. Chevrot, F. Cotton, and A. Souriau (2008). Simultaneous inversion of source spectra, attenuation parameters and site responses. Application to the data of the French Accelerometric Network, *Bull. Seismol. Soc. Am.* **98**, no. 1, 198–219.
- Edwards, B., and D. Fäh (2013). Measurements of stress parameter and site attenuation from recordings of moderate to large earthquakes in Europe and the Middle East, *Geophys. J. Int.* **194**, no. 2, 1190–1202.
- Edwards, B., B. Allmann, D. Fäh, and J. Clinton (2010). Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude, *Geophys. J. Int.* **183**, no. 1, 407–420.
- Edwards, B., D. Fäh, and D. Giardini (2011). Attenuation of seismic shear wave energy in Switzerland, *Geophys. J. Int.* **185**, no. 2, 967–984.
- Edwards, B., T. Kraft, C. Cauzzi, P. Kästli, and S. Wiemer (2015). Seismic monitoring and analysis of deep geothermal projects in St. Gallen and Basel, Switzerland, *Geophys. J. Int.* **201**, no. 2, 1020–1037.
- Edwards, B., C. Michel, V. Poggi, and D. Fäh (2013). Determination of site amplification from regional seismicity: Application to the Swiss National Seismic Networks, *Seismol. Res. Lett.* **84**, no. 4, 611–621.
- Edwards, B., A. Rietbrock, J. J. Bommer, and B. Baptie (2008). The acquisition of source, path, and site effects from microearthquake recordings using Q tomography: Application to the United Kingdom, *Bull. Seismol. Soc. Am.* **98**, no. 4, 1915–1935.

- Goertz-Allmann, B. P., and B. Edwards (2014). Constraints on crustal attenuation and three-dimensional spatial distribution of stress drop in Switzerland, *Geophys. J. Int.* **196**, no. 1, 493–509.
- Gohberg, I., and V. Olshevsky (1994). Fast algorithms with preprocessing for matrix-vector multiplication problems, *J. Complexity* **10**, no. 4, 411–427.
- Hanks, T. C., and R. K. McGuire (1981). The character of high-frequency strong ground motion, *Bull. Seismol. Soc. Am.* **71**, no. 6, 2071–2095.
- Harrington, R. M., and E. E. Brodsky (2009). Smooth, mature faults radiate more energy than rough, immature faults in Parkfield, CA, *Bull. Seismol. Soc. Am.* **99**, no. 4, 2323–2334.
- Hough, S. E., J. M. Lees, and F. Monastero (1999). Attenuation and source properties at the Coso Geothermal Area, California, *Bull. Seismol. Soc. Am.* **89**, no. 6, 1606–1619.
- Kschischang, F. R., B. J. Frey, and H.-A. Loeliger (2001). Factor graphs and the sum-product algorithm, *IEEE Trans. Inform. Theor.* **47**, no. 2, 498–519.
- Ktenidou, O.-J., C. Gélis, and L.-F. Bonilla (2013). A study on the variability of kappa (κ) in a borehole: Implications of the computation process, *Bull. Seismol. Soc. Am.* **103**, no. 2A, 1048–1068.
- Loeliger, H.-A., J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang (2007). The factor graph approach to model-based signal processing, *Proc. IEEE* **95**, no. 6, 1295–1322.
- Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data, *Astrophys. Space Sci.* **39**, no. 2, 447–462.
- Madariaga, R. (1976). Dynamics of an expanding circular fault, *Bull. Seismol. Soc. Am.* **66**, no. 3, 639–666.
- Monahan, J. F. (2011). *Numerical Methods of Statistics*, Cambridge University Press, New York, New York, 447 pp.
- Oth, A., D. Bindi, S. Parolai, and D. D. Giacomo (2011). Spectral analysis of K-NET and KiK-net data in Japan, Part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations, *Bull. Seismol. Soc. Am.* **101**, no. 2, 667–687.
- Ottmoller, L., and J. Havskov (2003). Moment magnitude determination for local and regional earthquakes based on source spectra, *Bull. Seismol. Soc. Am.* **93**, no. 1, 203–214.
- Raof, M., R. B. Herrmann, and L. Malagnini (1999). Attenuation and excitation of three-component ground motion in southern California mainland, *Bull. Seismol. Soc. Am.* **89**, no. 4, 888–902.
- Rietbrock, A., F. Strasser, and B. Edwards (2013). A stochastic earthquake ground motion prediction model for the United Kingdom, *Bull. Seismol. Soc. Am.* **103**, no. 1, 57–77.
- Savage, J. C. (1972). Relation of corner frequency to fault dimensions, *J. Geophys. Res.* **77**, no. 20, 3788–3795.
- Scargle, J. D. (1982). Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.* **263**, 835–853.
- Vannoli, P., G. Vannucci, F. Bernardi, B. Palombo, and G. Ferrari (2015). The source of the 30 October 1930 M_w 5.8 Senigallia (Central Italy) earthquake: A convergent solution from instrumental, macroseismic, and geological data, *Bull. Seismol. Soc. Am.* **105**, no. 3, 1548–1561.

Appendix

Here, we detail certain implementation solutions that enable an efficient implementation of the presented algorithms. For a large number of measurements N , several computations described in this article become extremely slow or prohibitive. In fact, matrix operations with N^2 elements may require too much time and even the storage in memory of such matrices may not be possible.

The matrix $\mathbf{F} \in \mathbb{R}^{N \times N}$ does not need to be defined explicitly. Indeed it is readily implemented using the fast Fourier transform (FFT) algorithm. In particular, several scientific libraries implement a function to compute the discrete

Fourier transform (DFT) assuming a real input signal and a Hermitian symmetric spectrum.

The matrix–vector multiplication with the inverse of a Toeplitz matrix is used when evaluating multivariate Gaussian (e.g., equations 11 and 27). In Gohberg and Olshevsky (1994), a matrix decomposition enabling a fast implementation of this multiplication is proposed. Using such an approach, the computational cost is greatly reduced exploiting the FFT algorithm. In addition, it is not necessary to store the entire $N \times N$ matrix in memory, but solely a vector of length N defining the Toeplitz matrix.

Computation of the determinant of a Toeplitz matrix is also necessary when evaluating a multivariate Gaussian. We used a recursive approach for the computation of the determinant of a Toeplitz matrix that is proposed in Monahan (2011). Also in this case it is not necessary to store the entire $N \times N$ matrix in memory.

In equation (14) and in the corresponding maximization for the incomplete data, the optimization vector $\boldsymbol{\eta}$ may be quite large because it contains M_i phases. We computed analytically the gradient of the objective function $\partial/\partial\boldsymbol{\phi}$ and provided this information to the numerical optimization routine. This greatly reduces the time needed for such optimization.

In the incomplete data case, the covariance matrix $\mathbf{V}_Z = \mathbf{P}\mathbf{G}\mathbf{P}^T$ is no longer a Toeplitz matrix. One simple approximation is to consider a block diagonal matrix having Toeplitz blocks and set to zero the matrix elements outside the diagonal blocks. This approximation corresponds to considering the noise independent in the different portions of contiguous data. Each diagonal block is Toeplitz and the fast method mentioned earlier can be used.

Swiss Seismological Service
ETH Zurich, NO FO 69.4
Sonneggstrasse 5
CH-8092 Zurich, Switzerland
marano@sed.ethz.ch
(S.M.)

Department of Earth, Ocean and Ecological Sciences
University of Liverpool
Liverpool L69 3GP, UK
(B.E.)

Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna
Via Donato Creti 12
40128 Bologna, Italy
(G.F.)

Swiss Seismological Service
ETH Zurich, NO H 53.1
Sonneggstrasse 5
CH-8092 Zurich, Switzerland
(D.F.)

Manuscript received 31 January 2016;
Published Online 3 January 2017