We IS 04

Systematic Recovery of Instrumental Timing Errors Using Interferometric Surface Waves Retrieved from Large N Seismic Arrays

C. Weemstra[1], J. De Laat[2], A. Verdel[3]

[1] Delft University of Technology; [2] Dublin Institute for Advanced Studies; [3] Netherlands Organization for Applied Sciences

Summary

Timing errors are a notorious problem in seismic data acquisition and processing. A technique is presented that allows such timing errors to be recovered in a systematic fashion. The methodology relies on virtual- source responses retrieved through the application of seismic interferometry (SI). In application to recordings of ambient seismic noise, SI involves temporal averaging of time-windowed crosscorrelation measurements. The retrieved interferometric responses are typically dominated by surface waves. Under favorable conditions, these interferometric responses therefore approach the surface-wave part of the medium's Green's function. Additionally, however, its time-reverse is often also retrieved. This implies time-symmetry of the time-averaged receiver-receiver crosscorrelations. In this study, this time-symmetry is exploited: by comparing the arrival time of the direct surface wave at positive time to the arrival time of the direct surface wave at negative time for a large a number of receiver- receiver pairs, relative timing errors can be determined in a weighted least-squared sense. The proposed methodology is validated using synthetic data. The results hold particular promise for large N seismic arrays.

Introduction

Seismic recordings occasionally suffer from timing errors. Ocean bottom seismometers (OBSs), for example, may have clock errors arising from the lack of a GPS connection. Various studies have recognized the potential of 'seismic interferometry' (SI), applied to recordings of ambient seismic surface-wave noise, to recover these timing errors. Sens-Schönfelder (2008), for example, exploits the (theoretical) pairwise time symmetry of the time-averaged crosscorrelations to recover timing errors. The fact that a non-uniform surface-wave illumination patterns breaks the time-symmetry of the time-averaged crosscorrelations (e.g., Weaver et al., 2009; Weemstra et al., 2013), however, will in practice adversely affect the accuracy of the recovered timing errors. In this paper, we present a method that mitigates the adverse affect of a non-uniform surface-wave illumination on the recovered timing errors. The method will be particularly useful in application to seismic arrays consisting of large numbers of sensors (so-called 'large N' arrays).

Seismic interferometry

SI refers to the principle of generating new seismic responses from existing recordings. Applied to passive seismic wave fields on Earth, simple time averaging of noise crosscorrelations often suffices to generate a new seismic response. In this study, we will restrict ourselves to the surface-wave part of the retrieved responses. Assuming the surface-wave noise illumination to be uniform (equal power from all directions) and the medium to be lossless, the time-averaged crosscorrelation $C_{i,j}(t)$ of the noise recorded at $\mathbf{x}_i$ and $\mathbf{x}_j$ will be proportional to the Green's function and its time-reversed version, convolved with the autocorrelation of the signal emitted by the (noise) sources (Wapenaar and Fokkema, 2006), i.e.,

$$C_{i,j}(t) \propto [G(\mathbf{x}_j, \mathbf{x}_i, t) + G(\mathbf{x}_j, \mathbf{x}_i, -t)] * P(t), \qquad [1]$$

where the in-line asterisk $*$ denotes temporal convolution and where $P(t)$ denotes the autocorrelation of the signal of the (noise) sources. In practice, equation [1] is often not exact, i.e., the Green's function (and its time-reversed version) are not accurately retrieved (e.g., Weemstra et al., 2013). In particular, a non-uniform illumination pattern leads to deviations of the retrieved surface-wave responses from the actual, correct surface-wave responses (Weaver et al., 2009). Nevertheless, in case ambient surface-wave energy is propagating in both directions along the line connecting $\mathbf{x}_i$ and $\mathbf{x}_j$, a direct surface-wave response will be retrieved at both negative and positive time (albeit with a potential error on their arrival time).

A model that accounts for timing errors

In this section we introduce a simple model. Using this model, a set of equations is obtained that will allow us to recover the unknown timing errors. Let us denote the arrival time of the medium's direct surface-wave response at positive time, i.e., the direct surface wave in $G(\mathbf{x}_j, \mathbf{x}_i, t)$, by $t_{i,j}^{(+)}$. Similarly, we denote the arrival time of the direct surface-wave response at negative time by $t_{i,j}^{(-)}$. By definition, therefore, $t_{i,j}^{(+)} = -t_{i,j}^{(-)}$. Let us now denote a potential timing error of the station at $\mathbf{x}_i$ by $\delta t_i^{(\text{ins})}$; a negative $\delta t_i^{(\text{ins})}$ implies that the recordings at $\mathbf{x}_i$ are subject to a time delay. Since we consider direct surface-wave responses retrieved through the application of SI, i.e., we exploit relation [1], we additionally introduce a time shift in the arrival time due to a deviation of the illumination pattern from uniformity: $\delta t_{i,j}^{(\text{src})}$. Because the retrieved direct surface-wave responses at positive and negative time are associated with opposite stationary-phase regions (e.g., Boschi and Weemstra, 2015), we distinguish between $\delta t_{i,j}^{(+,\text{src})}$ and $\delta t_{i,j}^{(-,\text{src})}$, which represent (illumination related) arrival-time time shifts at positive and negative time, respectively.

Accounting for the time shifts introduced above, the apparent arrival time of the direct response at positive time $t_{i,j}^{(+,\text{app})}$ coincides with $t_{i,j}^{(+)} + \delta t_i^{(\text{ins})} - \delta t_j^{(\text{ins})} + \delta t_{i,j}^{(+,\text{src})}$. Similarly, the apparent arrival time of the direct response at negative time, denoted by $t_{i,j}^{(-,\text{app})}$, is given by $t_{i,j}^{(-)} + \delta t_i^{(\text{ins})} - \delta t_j^{(\text{ins})} + \delta t_{i,j}^{(-,\text{src})}$. Summing $t_{i,j}^{(+,\text{app})}$ and $t_{i,j}^{(-,\text{app})}$, we have

$$t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})} = 2\delta t_i^{(\text{ins})} - 2\delta t_j^{(\text{ins})} + \delta t_{i,j}^{(+,\text{src})} + \delta t_{i,j}^{(-,\text{src})} \qquad [2]$$

In the ideal case that the station couple is illuminated uniformly from all angles and the recordings are not subject to timing errors, the right-hand side of this equation evaluates to zero and hence $t_{i,j}^{(+,\text{app})} = -t_{i,j}^{(-,\text{app})} = t_{i,j}^{(+)} = -t_{i,j}^{(-)}$.

The left-hand side of equation [2] is the measurable. In case we possess synchronous noise recordings by a total of M seismic stations, a maximum of $M(M-1)/2$ time-averaged crosscorrelations can be obtained. The set of equations governing the $t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}$ can in that case be written as,

$$\mathbf{T}^{(\text{app})} = \mathbf{A}\mathbf{T}^{(\text{ins})} + \mathbf{N}^{(\text{src})}, \qquad [3]$$

where the rows and columns of $\mathbf{A}$ relate to different station pairs and stations, respectively. Each row of $\mathbf{A}$ has only two non-zero entries, which hold a 2 and a -2. The column vector $\mathbf{T}^{(\text{ins})}$ holds the M timing errors we aim to recover, and the column vectors $\mathbf{T}^{(\text{app})}$ and $\mathbf{N}^{(\text{src})}$ hold the $t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}$ and $\delta t_{i,j}^{(+,\text{src})} + \delta t_{i,j}^{(-,\text{src})}$, respectively. In application to field data, $t^{(+,\text{app})}$ and/or $t^{(-,\text{app})}$ often cannot be determined for all time-averaged crosscorrelations (i.e., all combinations of i and j). This implies that the number of rows N of the matrix $\mathbf{A}$ and the the number of elements of $\mathbf{T}^{(\text{app})}$ will often be smaller than $M(M-1)/2$.

Recovering the timing errors

As it stands in equation [3], the rank of $\mathbf{A}$ is one lower than the number of unknowns (M). This indicates that the system of equations is effectively underdetermined. In practice, this implies that one needs to be certain about the timing of at least one station in order to obtain a unique estimate of $\mathbf{T}^{(\text{ins})}$. Assuming that specific station (or those specific stations) to be devoid of timing errors, its (or their) $\delta t^{(\text{ins})}$ will coincide with zero. Consequently, its entry (or their entries) can be eliminated from $\mathbf{T}^{(\text{ins})}$. Removing the associated column (or columns) from $\mathbf{A}$, renders the number of unknowns and the rank of $\mathbf{A}$ equal. Denoting the transpose of $\mathbf{A}$ by $\mathbf{A}^T$, the ordinary least-squares estimator of $\mathbf{T}^{(\text{ins})}$, which we denote by $\tilde{\mathbf{T}}_{(\text{ols})}^{(\text{ins})}$, is then given by $\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{T}^{(\text{app})}$ (Snieder and Trampert, 1999).

The accuracy of recovered timing errors in $\tilde{\mathbf{T}}_{(\text{ols})}^{(\text{ins})}$ largely depends on the nature of $\mathbf{N}^{(\text{src})}$. To find an improved estimator of $\mathbf{T}^{(\text{ins})}$, we use the results by Weaver et al. (2009) for the travel time error arising from non-uniformities in the illumination pattern. These authors derive that, under conditions reasonable for the direct surface waves considered in this study, both $\delta t_{i,j}^{(+,\text{src})}$ and $\delta t_{i,j}^{(-,\text{src})}$ can be expected to be inversely proportional to the travel time $t_{i,j}$ of the direct surface wave between $\mathbf{x}_i$ and $\mathbf{x}_j$. This suggests that solving equation [3] in a weighted least-squares sense may yield better estimates of the $\delta t_i^{(\text{ins})}$ than those obtained by computing $\tilde{\mathbf{T}}_{(\text{ols})}^{(\text{ins})}$. In addition, Weaver et al. (2009) find that mean of the remaining factor (i.e., the factor with which $1/t_{i,j}$ needs to be multiplied to arrive at their predictions for $\delta t_{i,j}^{(+,\text{src})}$ and $\delta t_{i,j}^{(-,\text{src})}$), which is predominantly illumination related, can be expected to deviate from zero.

To account for the aforementioned dependencies inherent in $\mathbf{N}^{(\text{src})}$ while inverting for $\mathbf{T}^{(\text{ins})}$, we define $\delta t_{i,j}^{(\text{azi})} \equiv t_{i,j}\left(\delta t_{i,j}^{(+,\text{src})} + \delta t_{i,j}^{(-,\text{src})}\right)$, which allows us to write $\mathbf{N}^{(\text{src})}$ as,

$$\mathbf{N}^{(\text{src})} = \mathbf{F} \odot \mathbf{N}^{(\text{azi})}, \qquad [4]$$

where $\odot$ denotes Hadamard matrix multiplication (i.e., element-wise multiplication). The vectors $\mathbf{F}$ and $\mathbf{N}^{(\mathrm{azi})}$ contain the $1/t_{i,j}$ and $\delta t_{i,j}^{(\mathrm{azi})}$, respectively. To account for the fact that the mean of $\mathbf{N}^{(\mathrm{azi})}$ probably does not coincide with zero, we write

$$\mathbf{N}^{(\mathrm{azi})} = \mathbf{E}\mu + \mathbf{N}^{(\mathrm{azi})}_{(\mathrm{zm})}, \qquad [5]$$

where μ is the mean value of the elements of $\mathbf{N}^{(\mathrm{azi})}$, $\mathbf{E} \equiv (1, 1, \cdots, 1)^T$, and $\mathbf{N}^{(\mathrm{azi})}_{(\mathrm{zm})}$ is a vector with zero mean (hence the subscript 'zm'). Substituting equation [5] in equation [4] and inserting the result in equation [3], we have

$$\mathbf{T}^{(\mathrm{app})} = \mathbf{A}\mathbf{T}^{(\mathrm{ins})} + \mathbf{F}\mu + \mathbf{F} \odot \mathbf{N}^{(\mathrm{azi})}_{(\mathrm{zm})} \qquad [6]$$

To be able to solve this system in a weighted least-squares sense, including the unknown parameter μ, we add this parameter to the vector of sought-for timing errors:

$$\mathbf{T}^{(\mathrm{ins})\,\prime} \equiv \begin{pmatrix} \mathbf{T}^{(\mathrm{ins})} \\ \mu \end{pmatrix}. \qquad [7]$$

Additionally defining

$$\mathbf{A}' \equiv (\mathbf{A} \quad \mathbf{F}) \qquad \text{and} \qquad \mathbf{N}' \equiv \mathbf{F} \odot \mathbf{N}^{(\mathrm{azi})}_{(\mathrm{zm})}, \qquad [8]$$

equation [6] can be written as

$$\mathbf{T}^{(\mathrm{app})} = \mathbf{A}'\mathbf{T}^{(\mathrm{ins})\,\prime} + \mathbf{N}'. \qquad [9]$$

The weighted least-squares solution to this system of equations reads (e.g., Snieder and Trampert, 1999)

$$\tilde{\mathbf{T}}^{(\mathrm{ins})\,\prime}_{(\mathrm{wls})} = \left(\mathbf{A}'^{\,T}\mathbf{W}_{\mathrm{d}}\mathbf{A}'\right)^{-1} \mathbf{A}'^{\,T}\mathbf{W}_{\mathrm{d}}\mathbf{T}^{(\mathrm{app})}, \qquad [10]$$

where for each station couple i, j the associated (diagonal) element of the diagonal matrix $\mathbf{W}_{\mathrm{d}}$ is given by $t^2_{i,j}$. As such, the inverse proportionality of the magnitude of the illumination related errors in $\mathbf{N}^{(\mathrm{src})}$ is accounted for in the inversion. In practice, however, $t_{i,j}$ is often not known. We will therefore use $|\mathbf{x}_j - \mathbf{x}_i|$ as a proxy for $t_{i,j}$ while computing $\tilde{\mathbf{T}}^{(\mathrm{ins})\,\prime}_{(\mathrm{wls})}$.

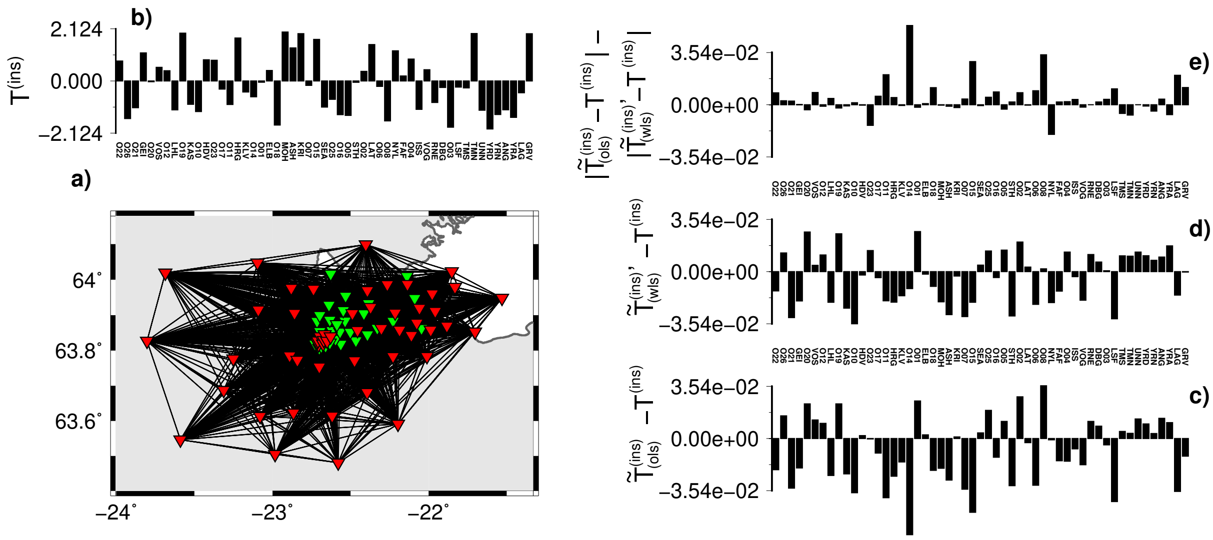

Figure 1: *(**a**) Station configuration of the RARR and eligible station couples (solid black lines). (**b**) The prescribed random timing errors. (**c**) Residual error after ordinary least-squares inversion. (**d**) Residual error after weighed least-squares inversion. (**e**) Difference in absolute residual errors. Note that the vertical axis of **b** is multiplied with a factor 60 with respect to the vertical axes of **c**, **d** and **e**.*

Application to synthetic data

The advantage of using the weighted least-squares estimator defined in [10] over the ordinary least-squares estimator is demonstrated using synthetic recordings of ambient seismic noise. The noise is "recorded" at the locations of the seismic stations constituting the Reykjanes Array (RARR), which was a dense seismic deployment on and around the Reykjanes peninsula, SW Iceland (the method proposed here was recently also applied to the field data recorded by this array, but the page limit didn't allow us to include those results in this paper). At each station location (83 in total), four months of synthetic seismic noise is "recorded". The stations were illuminated by uncorrelated single-mode dispersive surface waves generated in the far field. The uncorrelated effective plane waves were associated with an arbitrary (but smooth) non-uniform illumination pattern.

Time-averaged crosscorrelations were obtained by summing individual hourly crosscorrelations. Although the synthetic data contained energy between 0.05 and 0.5 Hz, the time-averaged crosscorrelations were filtered between 0.125 and 0.275 Hz. For each station couple, we estimated $t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}$ by crosscorrelating the (spline interpolated) direct surface wave arrival at positive time with the direct surface wave arrival at negative time. Figure 1**a** presents the configuration of the RARR as well as the ray paths associated with the eligible station couples: both the direct surface wave at positive time and the direct surface wave at negative time needed to exceed an a-priori specified signal-to-noise ratio threshold in order to qualify for estimation of $t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}$. Similar to the field data, the synthetic recordings by 30 station were free of timing errors (green triangles in Figure 1**a**); the other 53 stations (red triangles in Figure 1**a**) were given arbitrary timing errors between -2 and 2 seconds (Figure 1**b**) (the station names are explicitly indicated, but irrelevant in the context of this paper). The residual timing errors after the ordinary and weighted least squares inversions are depicted in Figure 1**c** and Figure 1**d**, respectively. Explicitly computing the difference between the absolute residual errors obtained with both inversions (Figure 1**e**), shows that the residual error decreases for most seismic stations when the proposed weighted-least squares approach is adopted. The average residual timing error decreases from 18.8 ms for the ordinary least sqaures solution to 14.9 ms for the weighted least-squares solution defined in equation [10].

Conclusions

We have derived a formulation that allows instrumental timing errors to be recovered using interferometric surface-wave responses. We validated the method using synthetic surface-wave (noise) recordings on and around the Reykjanes peninsula, SW Iceland. The technique could be useful in a variety of seismic (ocean bottom) contexts and will be particularly useful in application to large N seismic arrays.

References

Boschi, L. and Weemstra, C. [2015] Stationary-phase integrals in the cross-correlation of ambient noise. *Reviews of Geophysics*, **53**, 411–451.

Sens-Schönfelder, C. [2008] Synchronizing seismic networks with ambient noise. *Geophysical Journal International*, **174**(3), 966–970.

Snieder, R. and Trampert, J. [1999] Inverse problems in geophysics. In: Wirgin, A. (Ed.) *Wavefield Inversion*, Springer-Verlag, Vienna, 119–190.

Wapenaar, K. and Fokkema, J. [2006] Green's function representations for seismic interferometry. *Geophysics*, **71**(4), SI33–SI46.

Weaver, R., Froment, B. and Campillo, M. [2009] On the correlation of non-isotropically distributed ballistic scalar diffuse waves. *The Journal of the Acoustical Society of America*, **126**(4), 1817–1826.

Weemstra, C., Boschi, L., Goertz, A. and Artman, B. [2013] Seismic attenuation from recordings of ambient noise. *Geophysics*, **78**(1), Q1–Q14.