Modeling effect of time delay for large network of seismic monitor

Jordi Anguera* Leevi Annala† Stefan Dimitrijevic‡
Patricia Pauli§ Liisa-Ida Sorsa¶ Dimitar Trendafilov‖
Christophe Picard**

17.7.2018

Abstract

Seismic monitoring is used to study the behavior and composition of the underground floor. For earthquake prediction and underground works precise timing and positioning information is needed. Drilling companies use equipments that are linked in a network and are generally connected to a global positioning system for synchronization. However, instruments are not continuously synchronized and may deviate in time. Hence, the periods to which the vibration of the underground floor are caught might lack accuracy. Consequently, the precise localisation of the events becomes impossible. We have delay measurements and distance data available for an example seismic network and use it to correct the timing of the data in periods without GPS reception.

1 Introduction

In operation of seismic networks high quality of data is required for accurate prediction of seismic events. Precise timing is crucial but continuous GPS

*Autonomous University of Barcelona, Spain
†University of Jyväskylä, Finland
‡University of Novi Sad, Serbia
§Technical University of Darmstadt, Germany
¶Tampere University of Technology, Finland
‖University of Sofia ”St. Kliment Ohridski”, Bulgaria
**University of Grenoble Alpes and Grenoble INP, France

synchronization of the stations' internal clocks is not possible due to high energy consumption.

The purpose of this work is to analyze noise in a large network of seismic monitors and to extract clock drift from the delay. This work suggests an alternative way to determine clock drift based on the data only.

2 Problem Statement

2.1 Assumptions

All the stations have the same orientation. The medium the wave is propagating in is homogenous.

2.2 Definitions

Δ_i	clock drift at station i
δ_{ij}	Measured delay between station i and station j
$\hat{\delta}_{ij}$	Actual delay between station i and station j
ε_{ij}	Noise between two stations i and j, including clock drift and ambient noise
r_{ij}	Distance between stations
w_{ij}	adjacency matrix element, weight between two nodes

2.3 Modelling time delay between seismic monitors

The network of monitors can be modelled as a complete graph. The signal of each monitor is cross-correlated with other monitor signals, finally yielding time-delay data $\hat{\delta}_{ij}$ between each monitor. The time delay between the monitors comprise information on the real time the signal travels between the monitors δ_{ij}, the inaccuracies of the clock (time drift), and other noise affecting the measurement. The position of each seismic monitor in the network is available that can be used to compute the distances r_{ij} between all stations.

3 Data

The data is collected from a network of 73 seismic monitor stations recording ground vibrations in the Southeast region of France. The average distance between two monitors in the network is 166 km. The GPS coordinate position of each station in the network is known (Figure 1) and they each contain a clock which is synchronized to the GPS once a month. The seismic monitor sensor records compression and decompression as discrete values -1 and 1, respectively, every second. The sensor responds to any event in the area, be it an earthquake, tremor from road traffic, airplane or any other pressure wave which travels in the ground.

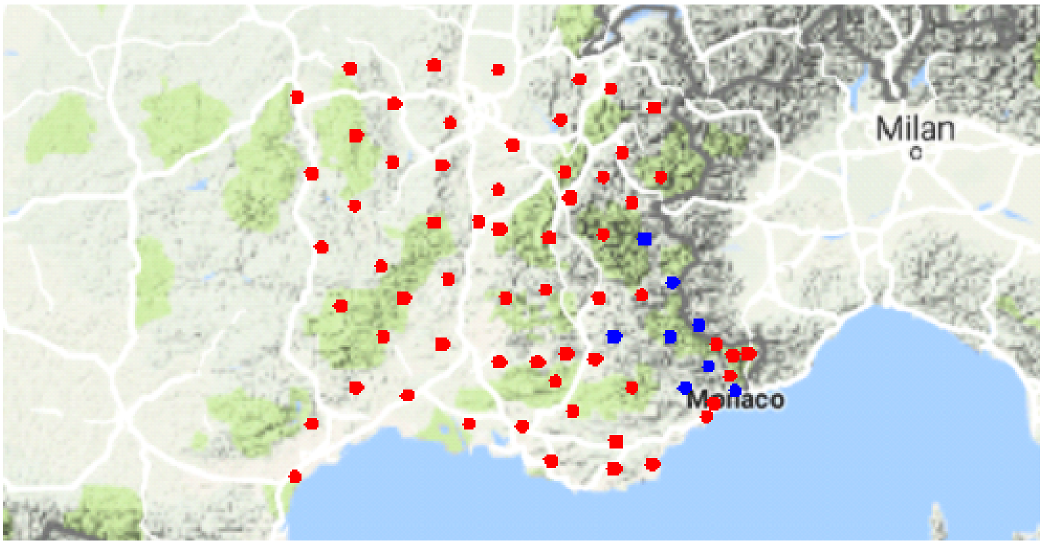


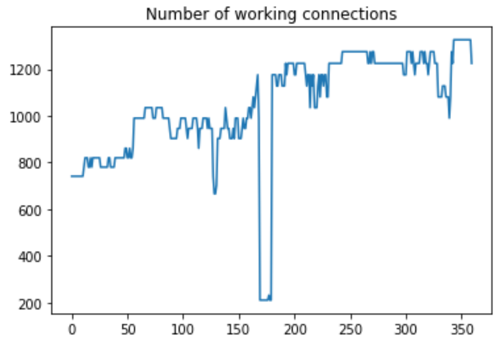
Figure 1: The network of seismic monitors are located in the Southeast France, mostly in the regions of Provence-Alpes-Côte d'Azur and Rhône-Alpes. The eight stations chosen as the small test set is shown as green in the southeast part of the area.

The stations work independent of each other and are occasionally shut down for some time frame for maintenance, repair or just random events. Just as occasionally they are brought back up to continue measuring. Therefore, the number of active stations varies over time. The Figure 2(a) shows the number of working stations, and the Figure 2(b) the number of working connections between the monitor stations over one year of measurements.

The compression and decompression data recorded by the stations is retrieved and run through initial data cleaning and filtering procedures. The data is then cross-correlated in one-hour time windows to yield time-delay data of signals between all monitors (Figure 3).

The time delays between each monitor are assumed to depend on the distance between the monitor. The further away the monitors are from each other, the larger the time delay caused by the same event recorded by the monitor. Figure 4 shows the time signal of the station pairs that are closest

(a) The number of working seismic monitor stations.



(b) The number of working connections between the stations.

Figure 2: Working stations and connections over one year time period.

and furthest apart over a time frame of 10 days supporting this assumption. One can see that the variance of the delays is much higher for the stations that are far apart.

The Figure 5 shows an example of time-delay evolution of all monitors over 24 hours. The distributions are symmetric with zero mean which implies that the delays at an instant cancel out over the whole area the 73 stations are mounted in. The interesting feature of the histograms is how the shape of the distribution changes slowly over time. However, the mean remains zero.

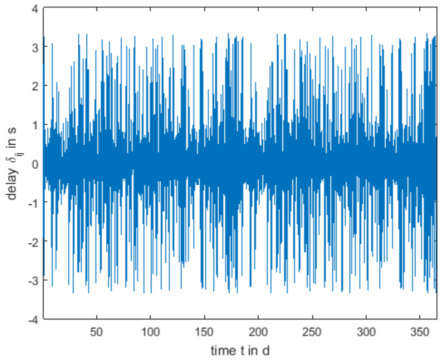
3.1 The most influential stations

The stations playing a key role in the global deviation of the total graph will be evaluated by the singular value decomposition of the global delays matrix $\hat{\boldsymbol{\delta}} \in \mathbb{R}^{n \times m}$, where n is the number of edges and m the number of time steps. !!!TODO: check the notation for the number of edges and time steps

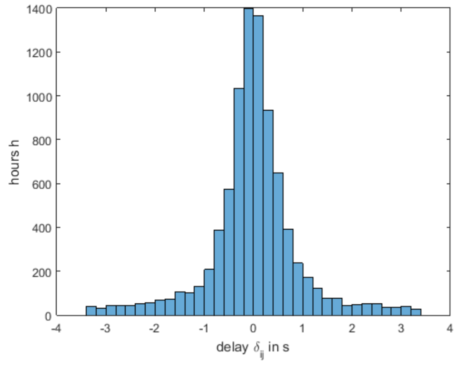
In order to find the index of the eigenvector with the most relative importance by the number of connections (regardless of the observed time delays) the time delay matrix $\hat{\boldsymbol{\delta}}$ was turned binary.

The Figure 6(a) shows the relative weight of the connections for each of the eigenvalue's index, the first one being the largest one. With that, the next step was to find the connections with the largest relative weight and the limit at which the connections do not have any more influence in the overall time delay, which will be shown in Figure 6(b) as a plot of the relative weight of each of the connections using the matrix $\hat{\boldsymbol{\delta}}$ with the measured time delays.

In order to select the group of connections with the largest weight a threshold of $|0.05|$ was taken, which resulted in 74 connections among 52



(a) Time Signal.



(b) Histogram.

Figure 3: The time-delay signal recorded between one station pair over a year. The time delay is computed in one-hour windows, yielding 24 time-delay values over a day.

active stations. Finally, three stations with the largest numbers of connections (34, 21, and 13) were identified. The remaining stations had only two connections or less.

These three stations were identified as the most influential stations, along with another group of five nearby these three stations stations, were taken as a small test set of stations to evaluate the methods developed in this work.

4 Methods

Let $\hat{\boldsymbol{\delta}} \in \mathbb{R}^{n\times n}$ be the measured pairwise time delays of the system, $\boldsymbol{\delta}$ the actual pairwise time delays, and $\boldsymbol{\varepsilon}$ an error term including clock drift of the station's clock and other errors

$$\hat{\boldsymbol{\delta}} = \boldsymbol{\delta} + \boldsymbol{\varepsilon}. \tag{1}$$

Each of the monitors are equipped with a clock which runs independent of others. It is synchronized via GPS system once a month and then runs independently. The inaccuracy in timing events at a station is caused by variations in the clocks' oscillators which oscillation may not be idea, they might respond to outside events and the oscillations are also affected by earthquakes.

The position of each seismic monitor in the network is known. As the monitors are spread across a large area, it is assumed that local tremors are detected by stations that are close by, and therefore the correlations found

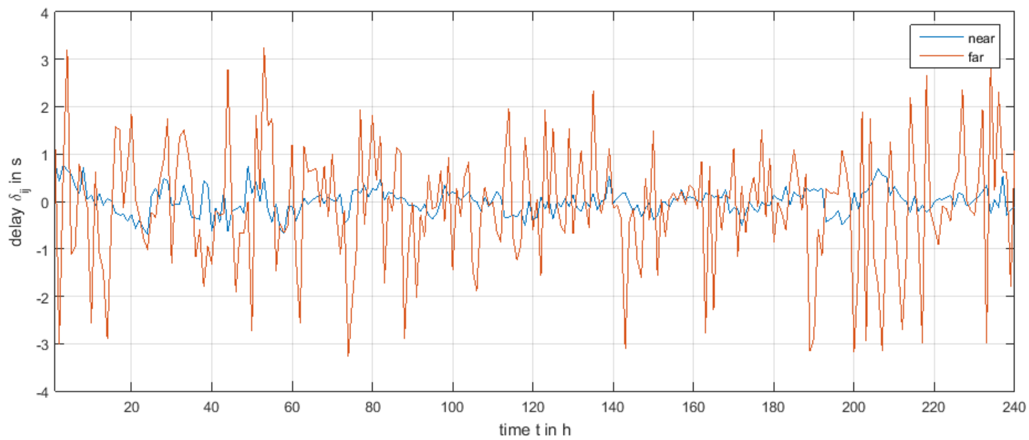


Figure 4: Comparison of the delay of the furthest and closest station pairs over 10 days.

in the pairwise cross-correlations and time delay data between them have higher likelihood to be linked.

4.1 The graph denoising model

4.1.1 Definitions

The computational denoising model developed in this work involves weighted network estimation by the use of topological graph metrics, described in detail in [2].

The monitor network constitutes a weighted graph $\mathcal{G}_\delta = (\mathcal{V}, \mathcal{E}, \hat{\boldsymbol{\delta}}_t)$ defined by a finite set of nodes $\mathcal{V}$ with $|\mathcal{V}| = n$, a set of edges $\mathcal{E} = \{(v_i, v_j) \in \mathcal{E}\}$, with $\max |\mathcal{E}| = n^2 - n$ and the weighted adjacency matrix $\hat{\boldsymbol{\delta}}$ with $\hat{\delta}_{ii} = 0$ for all i based on the measured time delays between stations. The matrix $\hat{\boldsymbol{\delta}}$ is symmetric and describes the time delays between events in the graph, and is normalized, i.e. $\hat{\delta}_{ij} \in [0, 1]$. The weighted adjacency matrix indicates the strength of connection between nodes.

The network also specifies another weighted graph $\mathcal{G}_r = (\mathcal{V}, \mathcal{E}, \boldsymbol{W})$, which sets of nodes and edges are the same as those of $\mathcal{G}_\delta$, but the adjacency weight matrix W is based on the physical distances between the station pairs. Also these weights are normalized between $[0, 1]$.

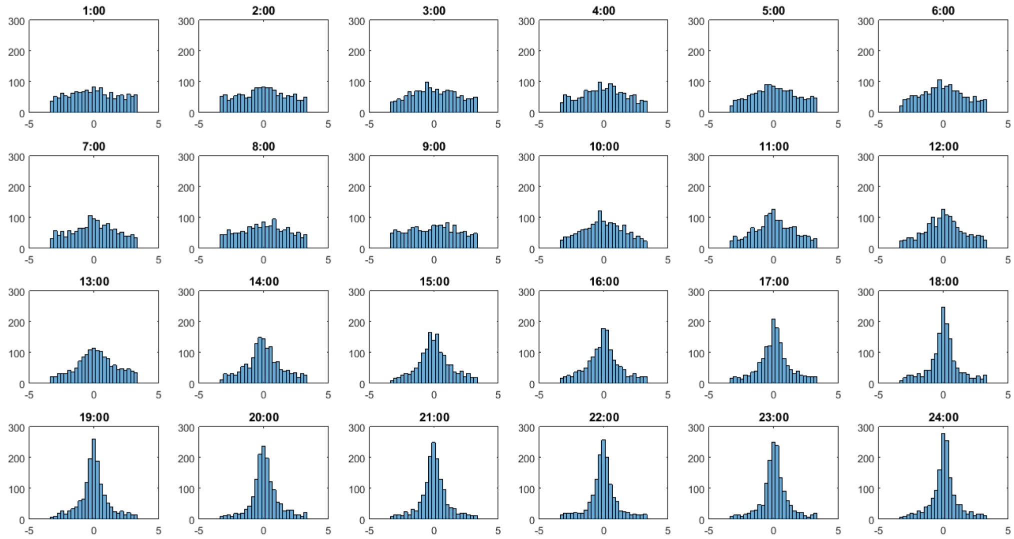


Figure 5: Histograms of time delays of all active stations recorded over 24 hours. The histogram shows how many connections have a particular delay at the instant.

4.1.2 Graph metrics

Graph metrics are scalar functions of the weight matrix X of a graph

$$f_i(\boldsymbol{X}) = K_i \tag{2}$$

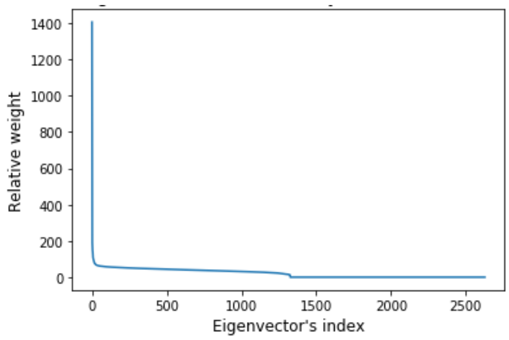
They quantify a property of the network. In this work, the metrics involve time delays (δ_{ij}) or the physical distances (r) between stations. The connection strength of a node i that is the row sum of the normalized weighted adjacency matrix:

$$f_i(\hat{\boldsymbol{\delta}}) = \sum_j \hat{\delta}_{ij} \qquad \text{for } \mathcal{G}_\delta = (\mathcal{V}, \mathcal{E}, \hat{\boldsymbol{\delta}}_t) \tag{3}$$

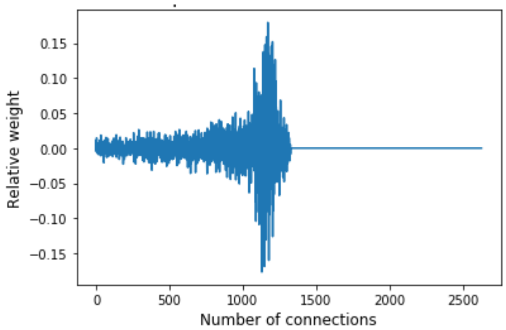
$$K_i = f_i(\boldsymbol{W}) = \sum_j w_{ij} \quad \text{for } \mathcal{G}_\delta = (\mathcal{V}, \mathcal{E}, \mathbf{W}). \tag{4}$$

The connection strength can be interpreted as the sum of all normalized delays of all of the connections to node i.

Figure 7 shows a minimal example of how the metric is computed. The node 1 has three connections yielding a connection strength $w_{12}+w_{13}+w_{14} = 0.1 + 0.3 + 0.6 = 1.0$. The other nodes and their connection strengths are computed similarly.



(a) Singular values for the binary connection matrix.



(b) Influences of connections.

Figure 6: Determining the influences of connections.

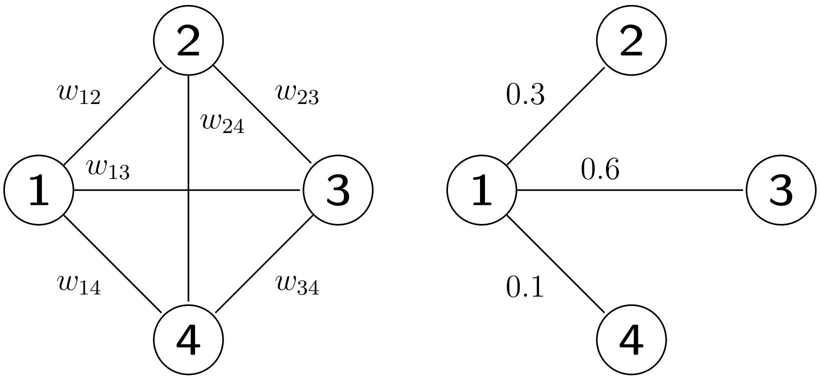


Figure 7: Example of graph and graph metrics

Next, we look at the K_i. Since we have the distance measurements available and expect a correlation between the delays and distances, we want to use this information to denoise the delay measurements and define

$$K_i = f_i(\mathbf{W}). \tag{5}$$

It is the connection strength as well. But here we plug in the distance matrix $\mathbf{W}$. The distances are constants and yield constant K_i. The weight matrix $\mathbf{W}$ is a function of the distance between node i and node j.

1. First, we choose the function to be the reciprocal of the distances of the node to all other nodes, $w_{ij} = 1/r_{ij}$, normalized between $[0, 1]$. Higher values correspond to nodes with short distances to all other nodes.

2. Second, we try a metric with weights proportional to the distances $(w_{ij} \propto r_{ij})$, normailzed between $[0, 1]$ accordingly. Here, high values correspond to long distances between two stations.

Other possible metrics are average neighbor degrees (resilience), transitivity, or a clustering coefficient. The analysis based on these metrics is beyond the scope of this work.

4.1.3 Cost function

We assume to have estimates of M differentiable graph metrics, one for each node i, of the original matrix $\hat{\boldsymbol{\delta}}$, i.e. $f_i(\hat{\boldsymbol{\delta}}) = K_i$, where $i \in \{1, \ldots, M\}$, then we can formulate a cost function that measures the deviation of the observed weight matrix's metrics $f_i(\hat{\boldsymbol{\delta}})$ to the estimates K_i as:

$$c(\hat{\boldsymbol{\delta}}) = \sum_i e_i^2(\hat{\boldsymbol{\delta}}) = \sum_i (f_i(\hat{\boldsymbol{\delta}}) - K_i)^2 \tag{6}$$

The error is minimized with gradient descent updates on $\hat{\boldsymbol{\delta}}$:

$$\hat{\boldsymbol{\delta}}^{(k+1)} = \hat{\boldsymbol{\delta}}^k - \mu \sum_i E_i(\hat{\boldsymbol{\delta}}^k) \frac{df_i(\hat{\boldsymbol{\delta}}^k)}{d\hat{\boldsymbol{\delta}}^k}, \tag{7}$$

where k is the iteration index and μ the learning rate. We find the derivative for the gradient descent update step. For $i = 2$ it is

$$\frac{df_2}{d\boldsymbol{\delta}} = \frac{\sum_j \hat{\delta}_{2j}}{d\hat{\boldsymbol{\delta}}} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 1 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}. \tag{8}$$

It is a symmetric matrix with ones in the ith row and ith column. The diagonal entries remain zero.

4.2 Clock drift estimate

The denoising model in section 4.1 computes the denoised estimate $\boldsymbol{\delta}$ of the measurement data $\hat{\boldsymbol{\delta}}$. The algorithm is run over a series of time points and an estimate for the clock drift Δ_{ij} between each station pair is obtained.

The noise, ε_{ij}, between each pair between two stations in the system is caused by clock drifts Δ_i and Δ_j, and other ambient errors E_{ij}

$$\varepsilon_{ij} = \Delta_i - \Delta_j + E_{ij}. \tag{9}$$

The clock errors of individual stations can be written in matrix form

$$
\begin{bmatrix}
1 & -1 & 0 & 0 & \dots & 0 \\
1 & 0 & -1 & 0 & \dots & 0 \\
1 & 0 & 0 & -1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & 0 & \dots & -1 \\
0 & 1 & -1 & 0 & \dots & 0 \\
0 & 1 & 0 & -1 & \dots & 0 \\
0 & 1 & 0 & 0 & \ddots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & 0 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
\Delta_1 \\ \Delta_2 \\ \Delta_3 \\ \Delta_4 \\ \vdots \\ \Delta_k \\ \vdots \\ \Delta_{n-1} \\ \Delta_n
\end{bmatrix}
=
\begin{bmatrix}
\Delta_1 - \Delta_2 \\ \Delta_1 - \Delta_3 \\ \Delta_1 - \Delta_4 \\ \vdots \\ \Delta_1 - \Delta_n \\ \Delta_2 - \Delta_3 \\ \Delta_2 - \Delta_4 \\ \vdots \\ \Delta_2 - \Delta_k \\ \vdots \\ \Delta_{n-1} - \Delta_n
\end{bmatrix}
=
\begin{bmatrix}
\Delta_{12} \\ \Delta_{13} \\ \Delta_{14} \\ \vdots \\ \Delta_{1n} \\ \Delta_{23} \\ \Delta_{24} \\ \vdots \\ \Delta_{2k} \\ \vdots \\ \Delta_{(n-1)n}
\end{bmatrix},
$$

which can be written more compactly as

$$\mathbf{Gs} = \mathbf{m} \tag{10}$$

where s is the unknown model vector and m the known data [1]. The matrix $\mathbf{G}$ has rank $n-1$, meaning that it is lacking full rank.

This represents a classical overdetermined inversion problem which can be solved by ordinary least squares regression after Tikhonov reqularization. The ordinary least squares seeks to minimize the sum of squared residuals

$$\|\mathbf{Gs} - \mathbf{m}\|_2^2. \tag{11}$$

When Tikhonov regularisation is added to the Equation 12, we obtain

$$\|\mathbf{Gs} - \mathbf{m}\|_2^2 + \|\mathbf{\Gamma s}\|_2^2, \tag{12}$$

for some suitable Tikhonov matrix $\mathbf{\Gamma}$. We choose this matrix as a multiple of the identity matrix $\mathbf{\Gamma} = \alpha\mathbf{I}$, giving preference to smaller norms.

The explicit solution is hence given by

$$\hat{\mathbf{s}} = (\mathbf{G}^T\mathbf{G} + \mathbf{\Gamma}^T\mathbf{\Gamma})^{-1}\mathbf{G}^T\mathbf{m}. \tag{13}$$

4.3 Algorithm

Leevi to plug in pseudo code of the implementation of the method

5 Results

5.1 Signal denoising

5.2 Clock deviations

6 Conclusions and Outlook

By analyzing the delay data we clearly showed that it changes over time showing patterns other than only white noise. This is due to clock drift. We used a method of weighted network estimation by the use of topological graph metrics exploiting the relation between delay and distances to denoise the data. The method is robust to temporarily disconnected stations that are only included as they are only included in the calculations when they are active. Hence, the algorithm can be run for on larger networks. The algorithm produced denoised delays we could deduce clock drifts from. Due to missing validation data we cannot make conclusions on the quality of our estimates. The algorithm could either be run on synthetic data where the results are known or on data sets that include continuous GPS synchronized data. To get more compelling results other metrics could be used, or even a combination of multiple metrics.

References

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

[2] L. Spyrou and J. Escudero. Weighted network estimation by the use of topological graph metrics. *CoRR*, abs/1705.00892, 2017.