Structural Health Monitoring of Concrete Structures using Traffic Noise Interferometry

by

Lance Jermaine A. Montenegro

in partial fulfillment of the requirements for the degree of

Master of Science in Civil Engineering - Structural Engineering

at Delft University of Technology,

to be defended publicly on August 29, 2024 at 14:00 (CET).

Thesis Chairman:	Dr. Ir. Yuguang Yang	TU Delft
Thesis Committee:	Dr. Ir. Fengqiao Zhang	TU Delft
	Dr. Katrin Löer	TU Delft

TU Delft

Acknowledgements

This thesis marks the culmination of my six-month research and studies at TU Delft. I hope that this work will contribute to advancing the use of ambient traffic noise for the structural health monitoring (SHM) of new and existing civil infrastructures.

First and foremost, I would like to express my gratitude to my thesis chairman, Dr. Ir. Yuguang Yang, and my thesis committee members, Dr. Ir. Fengqiao Zhang and Dr. Katrin Löer, for their invaluable guidance and support. In the beginning, I was completely clueless about my topic, especially since my research involved a strong focus on signal processing and seismological concepts. As a civil engineering student, these were areas I had never encountered before. Despite this, they ensured I stayed on track by guiding me, encouraging me, and giving me more credit than I thought I deserved.

I would like to extend special thanks to Dr. Ir. Fengqiao Zhang, who continuously guided me throughout the process. She was always available to discuss my progress and answer any questions I had. Moreover, she generously shared her MATLAB scripts, which were instrumental in helping me understand the concept of interferometry and in developing my own Python algorithm.

I am also grateful to Yubao Zhou for his assistance in acquiring and processing the data used in my research.

My student life would not have been the same without my friends Juan Camargo, Heisuke Miyoshi, Marina Diez, Jan Bergmann, Jose Stradi, Wessel Roodenburg, and Javier Fuertes, who stuck with me through all the ups and downs of the master's program and life in the Netherlands.

Finally, I would like to thank my constant support system: my parents, Leo and Genie; my siblings, Jericho, Loraine, and Lech; and last but certainly not least, my girlfriend Dani. Their unwavering belief in me powered me throughout this academic journey.

Lance Jermaine A. Montenegro
Delft, August 2024

Abstract

With the growing concern of aging infrastructures, the need for effective and non-intrusive monitoring techniques has become increasingly important. While most current methods rely on active testing, this study explores the potential of using ambient noise interferometry as a passive method for Structural Health Monitoring (SHM) of concrete structures. It investigates whether the Green's function (GF) of concrete medium can be adequately estimated from traffic noise to assess its health condition.

Two datasets are examined: a validation dataset from a laboratory experiment simulating ambient noise on a pre-stressed concrete girder, and real-world traffic noise data from the Maastunnel in Rotterdam. For each dataset, the following aspects are analyzed: (1) signal characteristics, including amplitudes and frequency distributions; (2) the optimal pre-processing scheme, incorporating temporal and spectral normalization, along with frequency filtering; and (3) the coherence of the resulting GF estimation from interferometry, particularly time of wave arrivals.

The results from the validation dataset demonstrate that ambient noise interferometry can reliably reconstruct the GF for concrete medium, indicating its effectiveness for monitoring changes such as crack formation and strain changes. However, the analysis of actual traffic noise data did not provide sufficient evidence to support its use for SHM with the current setup. Although a coherent and usable frequency range for traffic noise was identified, the limited amount of data led to a low signal-to-noise ratio (SNR), which made it challenging to highlight relevant features.

Moving forward, future researchers are encouraged to collect sufficient amount of data for analysis to better determine the feasibility of reconstructing the GF with ambient traffic noise. Additionally, exploring alternative sampling methods like continuous recording could address one of the limitations of this research. Finally, employing decomposition methods may help in increasing the SNR.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
AE	Acoustic Emission
AIC	Akaike Information Criterion
CCF	Cross-correlation Function
EEMD	Ensemble Empirical Mode Decomposition
EMD	Empirical Mode Decomposition
GF	Green's Function
IMF	Intrinsic Mode Function
NDT	Non-destructive Testing
PII	Passive Image Interferometry
RAM	Running-absolute-mean Normalization
SHM	Structural Health Monitoring
SI	Seismic Interferometry
SNR	Signal-to-noise Ratio

1

Introduction

1.1. Context

Civil infrastructure is an important asset of every society as it enables economic activities that contribute to economic and social welfare, fostering overall development and improving the quality of life for its inhabitants. Because of this, the safety and integrity of these structures becomes a top priority. Failure due to structural degradation and hazardous events could lead not only to economic losses, but also to the loss of human lives [1].

For large-scale infrastructure, concrete is commonly used because of its exceptional versatility, durability, strength, and cost-effectiveness [12]. However, like any other material, it is susceptible to deterioration in health and performance due to several factors such as aging, fatigue, environmental effects, and lack of quality control. Natural disasters also pose a threat to the performance and life of the structure. Therefore, measures must be put in place to ensure the reliability of large-scale concrete structures until the end of their intended design lives or beyond. To address this, regular maintenance and monitoring are required.

Recently developed and increasing in popularity is the field of structural health monitoring (SHM) [1], which focuses on continuous or periodic monitoring of engineering structures to assess their structural integrity. SHM can be defined as the process of implementing a damage identification strategy to allow early detection of damage at the incipient level so that necessary precautionary measures could be applied proactively [23]. If done correctly, SHM ensures structural safety and reduces rehabilitation costs significantly [28].

In the usual practice of SHM, the predominant method is a combination of visual inspections followed by localized non-destructive testing (NDT) at typical points of failure or where possible damage is seen. However, earlier NDTs, such as ultrasonic pulse velocity (UPV), fiber optic sensing (FOS), and eddy currents, require active testing, which is intensive in terms of time, cost, and personnel.

In this study, the primary concept explored was using passive ambient noise to characterize a heterogeneous medium through interferometry, a technique adapted from seismological practices. One specific application within seismology involves utilizing the Green's Function, which is the impulse response of a medium estimated from cross-correlations of ambient noise, to reconstruct a three-dimensional image of the Earth's subsurface based on variations in wave speeds and group velocity maps [26, 36].

In evaluating the velocities derived from the GFs using ambient noise in concrete structures, the goal was to assess whether valuable information could be extracted, particularly concerning the presence of damages such as surface cracks. Research indicates that cracks can notably decrease both the amplitude and velocities of waves [33].

In this research, two datasets were used: the first from a previous acoustic emissions (AE) test of a prestressed concrete girder in a lab, and the second from passive traffic noise data collected from the Maastunnel in the Netherlands.

The first dataset was used to develop a signal-processing algorithm and validate the concept, taking advantage of the reduced uncertainties of a laboratory experiment. The second dataset was for testing the effectiveness of the concept on actual passively gathered traffic noise data.

1.2. Research Questions

- Can the Green's function between a pair of sensors installed on a concrete structure be retrieved through the interferometry of the received signals?
  - Can changes in the estimated Green's function be detected and used to draw conclusions about the state of the concrete structure?
  - Can the arrivals of waves be detected in the estimated Green's function between a sensor pair?
- How effective is traffic noise interferometry in monitoring the structural health of concrete structures in both laboratory and real-world conditions?
  - What frequency range must be extracted to obtain coherent signals that are suitable for Green's function estimation in the context of concrete structure monitoring?
  - What is the optimal pre-processing procedure to enhance the quality and coherence of signals extracted from ambient traffic noise data?
  - How do the characteristics of signals from lab-simulated ambient noise experiments compare to those obtained from real-world ambient traffic noise, and what adaptations are necessary to successfully apply these methods in real-world scenarios?

1.3. Research Approach

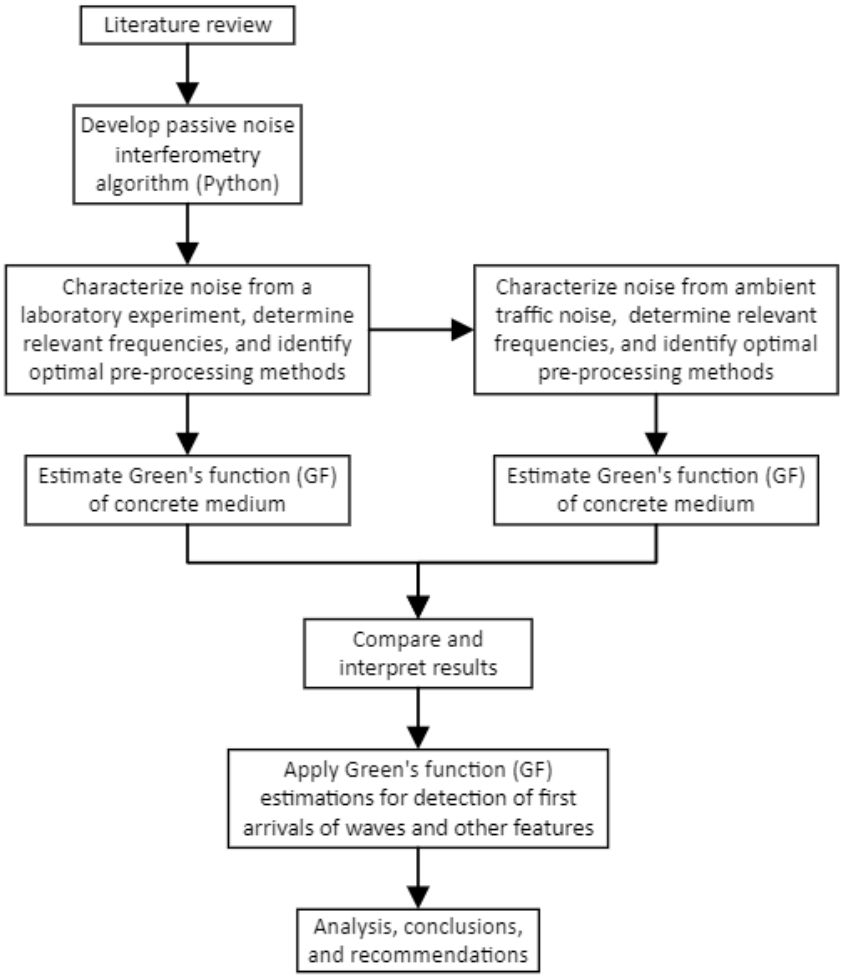

Figure 1.1: Research workflow

The main objective of this study was to explore the potential of using ambient traffic noise to estimate the Green's function of a concrete medium, with the goal of aiding the monitoring of both time- and space-dependent properties of concrete. This capability could help in detecting and localizing potential structural damage. The study began with a validation dataset derived from a controlled laboratory experiment, where the noise source was well-defined, serving as a basis and benchmark. Subsequently, the method developed and validated using the laboratory experiment data was applied to a real concrete structure (the Maastunnel) for further evaluation and comparison. Finally, conclusions and recommendations were made.

2

Review of Related Literature

2.1. Ambient Noise Data Processing

Ambient noise is the typical or persistent background noise that is present in any given environment. This includes noises generated by both natural sources, such as wind and rainfall, and man-made sources, such as traffic and other human activities. Using ambient noise recordings for interferometry (Section 2.2) eliminates the need for active sources, allowing for easier and cheaper acquisition and repeated measurements.

When collecting ambient noise data, the time series of the signal is generally expected to be flat, with occasional distinct peaks from random high-amplitude events. To emphasize the ambient noise and prevent these high-amplitude events from obscuring it, normalization and band-pass filtering are applied [4].

In this study, both analyzed datasets were from ambient noise sources, however one was obtained in a laboratory setting with minimal random noise, while the other was from actual field recordings.

Due to the variations in data collection conditions, different pre-processing techniques were used on each dataset. Once the pre-processing was done, the data was cross-correlated and then stacked to ultimately obtain an estimation of the Green's function.

In the subsequent sections, the processes that were done will be further explained.

2.2. Seismic Interferometry

In the field of seismology, seismic interferometry (SI) pertains to the method in which the cross-correlation of ambient seismic noise at two receiver locations provides an estimate of the Green's function (GF), representing the impulse response of the medium between the two stations as if one of them had been an active source (Figure 2.1) [32]. This is further explained in Section 2.2.3.

When ambient noise is influenced by transient and high-amplitude events, the formation of a broad-band diffuse wavefield is hindered, increasing the required amount of data and resulting in poorer GF estimations [10]. The ideal scenario for obtaining the GF between two receivers occurs when the ambient noise wavefield is diffuse and equipartitioned, with energy uniformly distributed across all directions and modes, facilitating faster convergence towards an accurate GF (Section 2.2.3).

A simple application of retrieving the GF using SI is the retrieval of the direct wave in a one-dimensional (1D) plane (Figure 2.1). Suppose there are two aligned sensors (s_1 and s_2) and a plane wave emitted by an impulsive source located at an extreme end of either sensor travels towards them. The wave is then received by the two sensors at different times (t_1 and t_2). By cross-correlating the signals received by the two sensors, a virtual impulse response system is formed, where the first sensor hit acts as the source (s_1) and the second one (s_2) as the receiver. The resulting system shows the time lag, both positive and negative, at which the two signals are most alike. As a result, the common paths are eliminated (the portion between the source and s_1), and only the parts of the signal between the two sensors are retained.

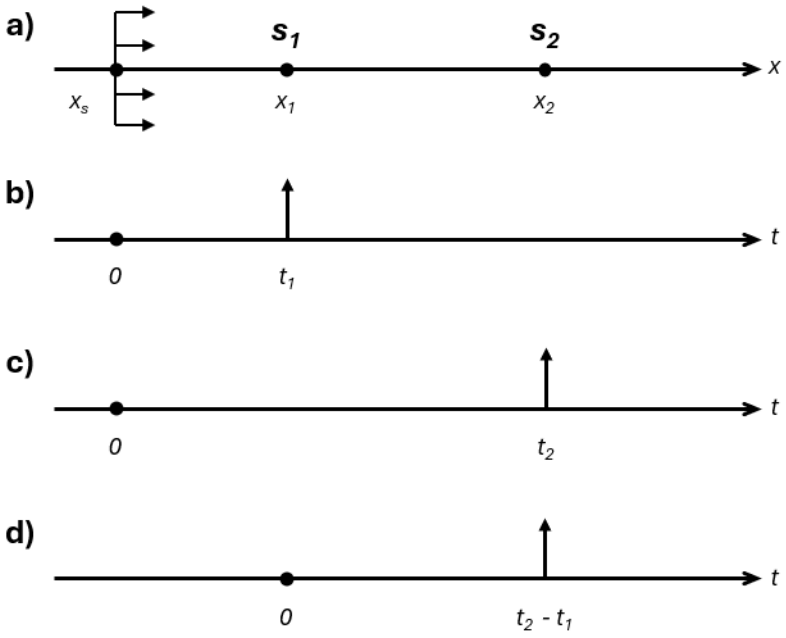

Figure 2.1: Direct wave interferometry in 1D. (a) Plane wave emitted rightwards along the x-axis from impulsive source at x_s. (b) Impulse received by s_1, which is also the Green's function $G(x_1, x_s, t)$. (c) The Green's function $G(x_2, x_s, t)$. (d) The result of the cross-correlation of b and c where s_1 has become the virtual source ($G(x_2, x_1, t)$ [32].

As mentioned in Section 1.1, this study focused on exploring the possibility of monitoring a concrete structure by estimating the GFs between different sensor pairs through the interferometry of ambient traffic noise. A similar study was conducted [25], where geophones recorded ambient traffic noise data from a concrete bridge and used Passive Image Interferometry (PII) to detect velocity fluctuations within the medium and deduce changes.

PII [26] was originally applied to obtain velocity variations within the subsurface using ambient noise from a volcano. Simply put, PII combines noise correlations to obtain GF reconstructions and Coda Wave Interferometry (CWI) to determine velocity variations (Section 2.3).

In their study, Salvermoser et al. [25] estimated GFs from noise correlations and focused on the coda, the later-arriving waves, while disregarding the first arrivals. The longer travel paths of coda waves result in more scattering, making them sensitive to small changes because they contain more information about the medium. Their findings indicated that ambient traffic noise could be used to detect damage presence by mapping velocity variations. However, the study faced limitations in spatial resolution due to the low sensitivity of the sensors used.

In contrast, this research utilized high-frequency resonant sensors and focused on the coherence of

wave arrivals to assess the structural condition of concrete medium. The consistency and coherence of the wave arrivals were used to detect changes in the medium by comparing the results between uncracked and cracked conditions. This approach was chosen to potentially enhance spatial resolution and sensitivity to changes in the medium.

2.2.1. Temporal normalization

In seismology, temporal normalization, otherwise known as time domain normalization, is a process that is used to reduce the effects of instrumental irregularities and non-stationary events on cross-correlated seismic data [4]. Additionally, this is also done to suppress the influence of transient signals [13] (Figure 2.2). This ensures that the signals are easily comparable to other signals by adjusting the amplitude scales and imposing a specific range.

In this study, two methods of temporal normalization were investigated to see which would yield more coherent signals. The investigation was deemed necessary since there is currently no literature that specifically analyzes ambient traffic noise at higher frequencies, hence the characteristics of the elastic waves produced are still unknown. It is also important to recognize that elastic waves from different sources differ, so there is not a single temporal normalization scheme that is universally effective.

The first method explored was the one-bit normalization which disregards the values of the amplitudes by replacing all positive and negative values with 1 and -1, respectively. Campillo and Paul [6] proved that the cross-correlation of one-bit normalized data yielded similar results to the cross-correlation of the same data that is truncated. Salvermoser et al. [25] used one-bit normalization to reduce the influence of direct surface waves on passive traffic noise data by decreasing their distinctive amplitudes.

The second method, running-absolute-mean normalization (RAM), involves weighting values in a time series by the average of the absolute time series values within a specific time window. The size of this window varies depending on the characteristics of each dataset. In a study by Bensen et al. [4], they determined that a window width equal to half the maximum period of the considered band-pass range was effective, suggesting it as a useful initial guideline. However, they noted that the window size could be varied significantly while still producing similar results.

The normalization weight for this method, denoted as ω_n, for a time point n is calculated using the formula [4]:

$$\omega_n = \frac{1}{2N+1} \sum_{j=n-N}^{n+N} |d_j| \tag{2.1}$$

Here, (2N + 1) is the width of the normalization window, which dictates the amount of amplitude information that is retained, and d_j is a discrete time series. The resulting weight is used to obtain the normalized datum, $\widetilde{d_n}$.

$$\widetilde{d_n} = \frac{d_n}{\omega_n} \tag{2.2}$$

Bensen et al. [4] successfully used this normalization method to obtain surface wave dispersion measurements from ambient seismic noise.

Similarly, Groos et al. [13] investigated the effectiveness of different temporal normalization schemes on seismic noise and found that RAM was a highly effective non-linear temporal normalization method that successfully suppresses transient signals.

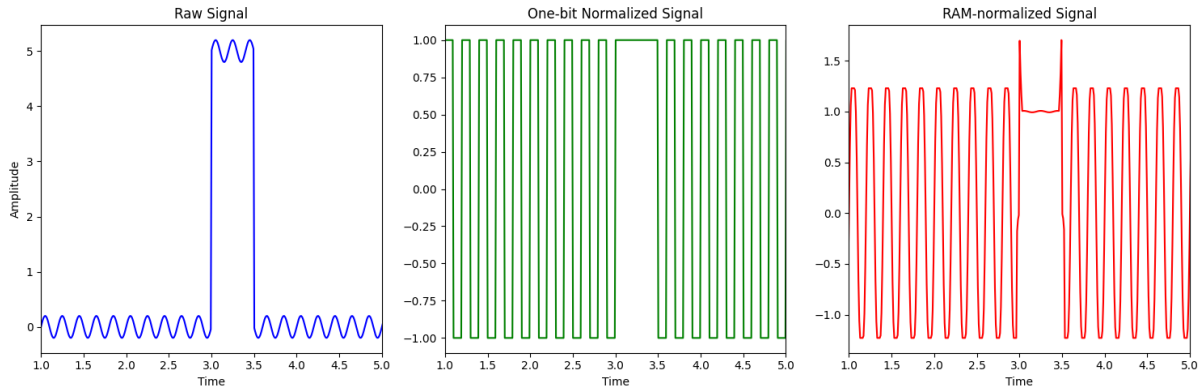

Figure 2.2: One-bit and RAM normalization of a synthetic signal with a transient event.

2.2.2. Spectral whitening

In the context of ambient noise, the goal of spectral whitening is to broaden the signal frequency band [4]. Ambient noise data often contain persistent and prominent noise sources, causing the resulting signal and cross-correlation functions to be dominated by these sources. Spectral whitening equalizes or normalizes energy in the frequency domain, reducing the overpowering effects of dominant frequencies and ensuring that other frequencies are also significantly represented. This process effectively 'whitens' the signal.

For this research, spectral whitening was performed by normalizing the signals using an envelope function. This envelope function was derived by taking the magnitude of the analytic signal, obtained through the absolute values of the Hilbert transform of the signal. Subsequently, the root-mean-square (RMS) envelope was computed by applying a sliding window of a consistent, pre-defined length to calculate the RMS at each point. Finally, the Fourier transform of the original signal was divided point-wise by the corresponding values of the RMS envelope.

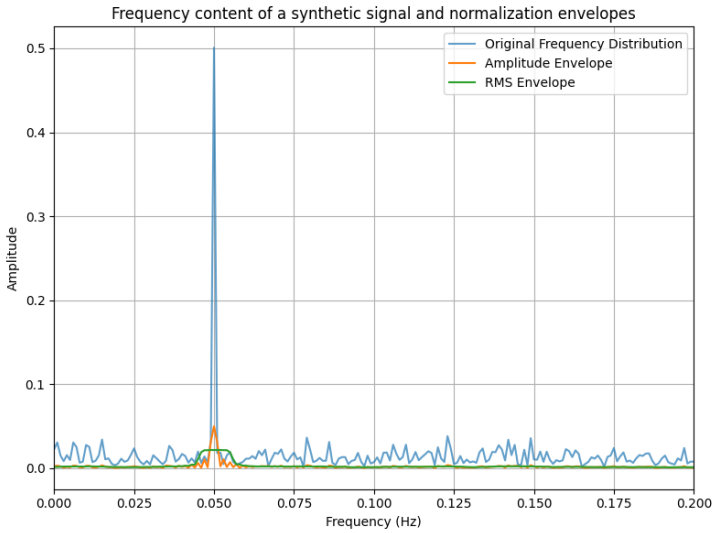

Figure 2.3: Envelopes for spectral whitening (synthetic signal).

The RMS at any point along the signal can be computed with the following formula:

$$RMS_i = \sqrt{\frac{1}{n}\sum_i x_i^2} \tag{2.3}$$

In equation 2.3, n refers to the length of the sliding window, i refers to the frequency bin, and x_i is the magnitude of the Fast-Fourier Transform (FFT) of the signal at frequency bin i.

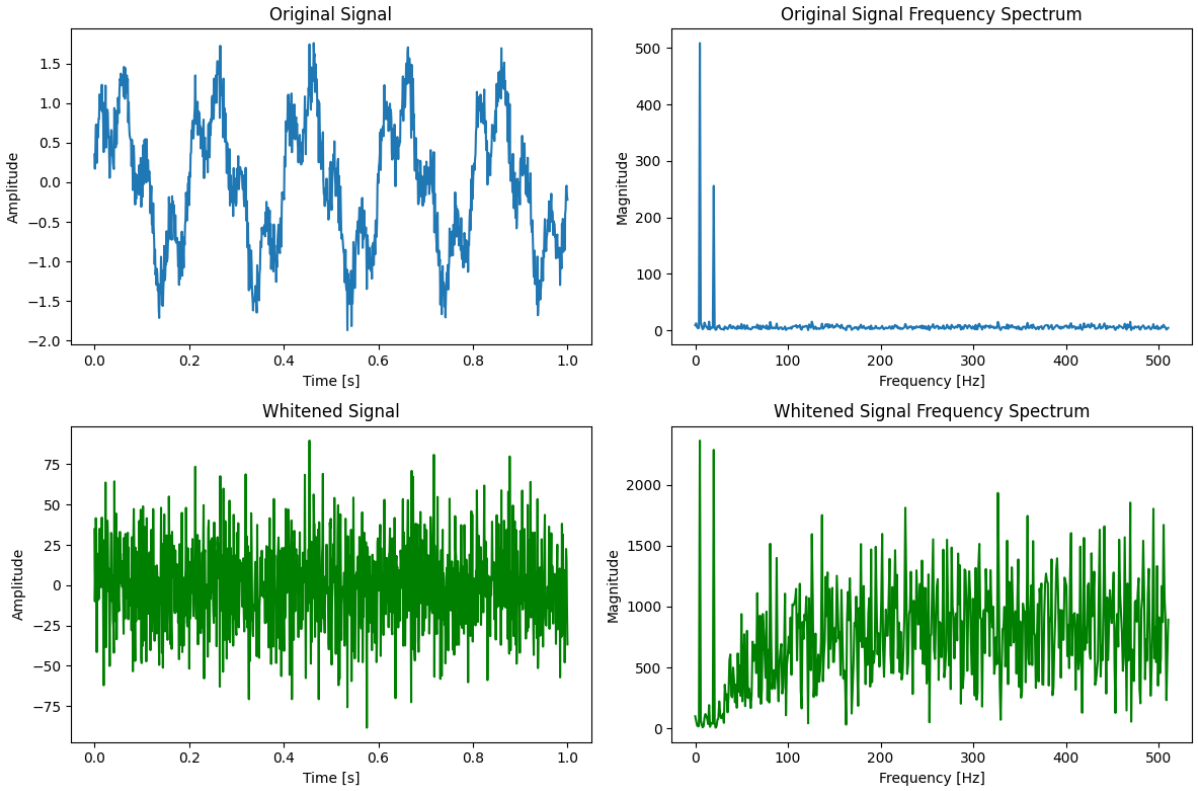

Figure 2.4: Spectral whitening of a synthetic signal using the RMS envelope

In Figure 2.4, the original synthetic signal features a periodic trend with two dominant frequencies at approximately 5 Hz and 20 Hz. After spectral whitening, the trends are greatly reduced, and the frequency spectrum becomes more normalized. While the whitened signal alone may seem less informative, it is still very useful for applications that uses other concepts such as cross-correlations.

2.2.3. Green's function retrieval

Cross-correlation is a method that quantifies the similarity between two different signals by summing the products of their values when one is shifted by a time lag t. Given two different signals observed at stations A and B, the cross-correlation function (CCF) between them can be computed as:

$$CCF_{AB}(t) = \int A(\tau) \cdot B(\tau + t) d\tau \tag{2.4}$$

After the individual CCFs for a given station pair over a specific period are obtained, they are stacked together to collectively enhance the signal-to-noise ratio (SNR). This stacking process involves grouping CCFs computed using consistent time windows, such as hourly segments for daily data or daily segments for longer periods like a month or a week. Different stacking methods exist [35], with the most common being either the summation or the averaging of CCFs from multiple windows. These methods

aim to enhance the coherence and reliability of the signals [35]. Although the underlying assumption for ambient noise is that it is free of transient signals, this is often not the case in reality. Therefore, stacking over a longer period helps average out random noise and transient events, allowing the consistent and coherent signals to emerge more clearly

The SNR value used to evaluate the stability of the stacked CCFs [4] can be solved with the following equation:

$$SNR = \frac{max(|S_{CCF}|)}{\sqrt{\frac{1}{m}\sum_{i=n}^{n+m} CCF_i^2}} \tag{2.5}$$

Here, the numerator is the maximum amplitude within the signal window, denoted as S_{CCF}. This window contains the expected or observed signal amidst the background noise. The denominator represents the root-mean-square (RMS) of the trailing noise window [4]. When selecting the trailing noise window, it is common practice to choose a region a few samples after the expected signal. This approach ensures that the noise window contains only background noise and does not include any contributions from the noise induced by the signal in question. Figure 2.5 illustrates the windowing process done for the calculation of the SNR.

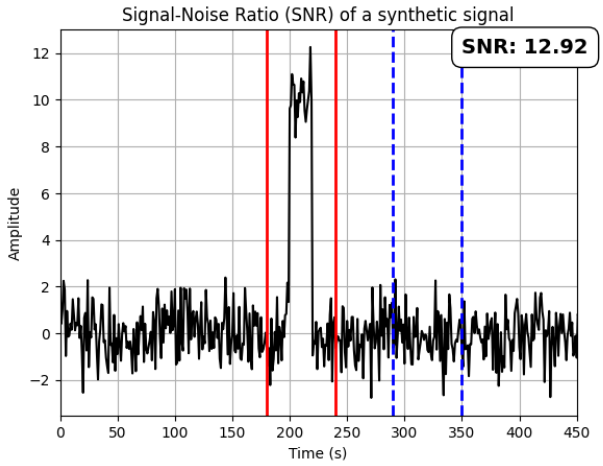

Figure 2.5: Example of getting the SNR value of a synthetic signal. Solid vertical red lines and dashed vertical blue lines represent the chosen signal window and chosen noise window, respectively.

When cross-correlations are stacked over longer periods, the resulting waveform represents an estimate of the GF or the system's impulse response. The higher the SNR value obtained from the estimate, the more desirable it is. However, its limitation is that the reliability of an estimate cannot be assessed solely based on a high SNR due to the absence of a theoretical value linking the two. Thus, a validation must be performed to determine whether the obtained Green's function is sufficiently representative of the medium in question.

To understand how a Green's function derived from cross-correlation appears, it is essential to consider the structure of the resulting CCF. A CCF consists of two components: the causal part, corresponding to the positive time lag, and the acausal part, corresponding to the negative time lag. The amplitudes of the function are directly dependent on the direction of energy flux between two stations [29]. In an ideal scenario where there is a perfectly isotropic distribution of sources around the two stations in question, the resulting CCF's causal and acausal parts will be symmetric about the time lag = 0. However, in practice, this condition is difficult to achieve, and therefore asymmetry is often observed.

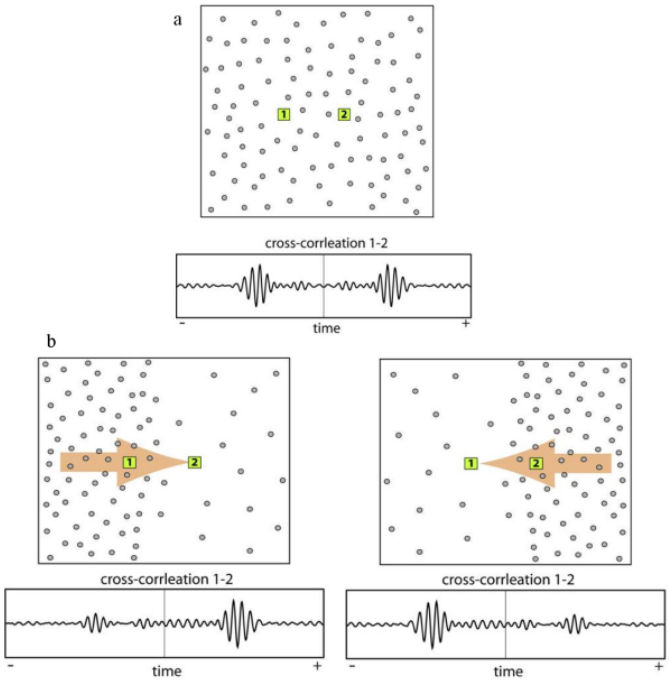

Figure 2.6: Assymmetry of CCFs with respect to the distribution of sources [29].

2.3. Coda Wave Interferometry

Coda waves refer to the 'tail' or the later-arriving signals of a seismic event which follow the direct arrival of primary waves. They are considered to be the superposition of scattered and reflected waves as a result of their interactions with heterogeneous media. The medium in which the wave propagates may be considered as an interferometer, hence the term coda wave interferometry (CWI) [31].

CWI is a method used to detect changes in a propagation medium by measuring subtle variations in travel times within the coda. To perform CWI, the stretching technique is typically used, where a reference signal is compared to a 'stretched' signal by cross-correlating them. The stretching process involves adjusting the time scale of the signal to align with the reference, and the extent of this adjustment indicates the velocity change in the coda [25].

2.4. Wave Propagation in Concrete

To assess wave propagation in concrete, an impulse is transmitted from a source to a receiver at a known distance, allowing for the measurement of velocity based on the recorded travel time. This process reveals properties of the medium, such as elasticity and material quality, and can also detect the presence of defects [12]. The underlying principle is the acoustoelastic property of elastic solids, which refers to the change in propagation of a mechanical wave due to a change in stress within the medium [11]. For concrete, since the deformations within the medium are assumed to be small, the theory of linear elasticity may be applied. Hence, Hooke's law may be used. Moreover, despite the usual non-homogeneous and anisotropic nature of concrete, it can be effectively treated as a homogeneous material if the wavelength of a propagating wave is significantly larger than the size of aggregates in the mixture.

2.4.1. Wave velocity

For a given medium, waves that propagate through it can typically be categorized as either body waves or surface waves, depending on the manner in which they propagate. Essentially, body waves travel

through the interior of the medium while surface waves travel along the surface.

Under body waves, there are two types: Primary waves (P-waves or longitudinal waves), with particle motion parallel to wave propagation, and Secondary waves (S-waves or transverse waves), with particle motion perpendicular to it. The velocities of these waves can be determined using the constitutive elastic wave propagation equations which relate the propagation velocity to the basic properties of a medium [3]:

$$v_l = \sqrt{\frac{E}{\rho} \cdot \frac{(1-\nu)}{(1+\nu)(1-2\nu)}} \tag{2.6}$$

$$v_s = \sqrt{\frac{E}{\rho} \cdot \frac{1}{2(1+\nu)}} \tag{2.7}$$

where v_l and v_s are the wave velocities in the longitudinal and transverse directions, respectively, E is the modulus of elasticity, ρ is the material density, and ν is the Poisson's ratio.

For concrete media, Rayleigh waves are the most commonly found surface waves. These are waves produced from a linear combination of longitudinal and transversal vibrations from the interaction of P-waves and S-waves [20]. Since the computation of the exact value of a Rayleigh wave's velocity can prove to be complex due to the heterogeneity and anisotropy of concrete, an approximation can be solved with the following formula [37]:

$$v_r = v_s \frac{0.87 + 1.12\nu}{1+\nu} \tag{2.8}$$

where v_r is the Rayleigh wave velocity.

For a concrete specimen with a grade of C40/50 with dynamic modulus of elasticity (E_d) = 36000MPa, Poisson's ratio (ν) of uncracked concrete = 0.2, and density (ρ) = 2500kg/m^3, the corresponding wave velocities are v_l = 4000m/s, v_s = 2450m/s, and v_r = 2234m/s.

2.4.2. Wave attenuation

As waves propagate from a source to a receiver, a loss of energy is observed as the wave moves away from the source. This phenomenon is called wave attenuation, which can be categorized as either intrinsic or extrinsic attenuation. The loss of energy is directly reflected in the decrease of the wave's amplitude.

Intrinsic attenuation refers to the loss of energy that occurs within the medium. This occurs primarily due to two mechanisms: absorption, which converts some of the energy into other forms, such as heat, and scattering, which redistributes energy as the waves encounter heterogeneities and boundaries [14]. Extrinsic attenuation, on the other hand, relates to the consequences that arise from the chosen method by which the attenuation measurements are taken [22].

During wave propagation in uncracked concrete, attenuation is primarily from geometric spreading loss and material attenuation. The geometric spreading loss in body waves follows a spherical pattern, spreading equally in all directions, while for Rayleigh waves it follows a cylindrical pattern as the energy is concentrated along the horizontal direction [5].

2.5. Acoustoelastic Effect

The acoustoelastic effect refers to the phenomenon where changes in stress within an elastic medium cause changes in wave velocity. When the medium is under compression, the wave velocity increases, whereas under tension, it decreases [17].

Hughes and Kelly [16], using Murnaghan's theory of finite deformation, derived the following equations to represent the body wave acoustoelasticity of a material when a uniaxial stress σ_{11} is applied in 1-direction [8]:

$$\rho_0 \nu_{11}^2 = \lambda + 2\mu + \frac{\sigma_{11}}{3K}\left[2l + \lambda + \frac{\lambda+\mu}{\mu}(4m + 4\lambda + 10\mu)\right] \tag{2.9}$$

$$\rho_0 \nu_{22}^2 = \rho_0 \nu_{33}^2 = \lambda + 2\mu + \frac{\sigma_{11}}{3K}\left[2l - \frac{2\lambda}{\mu}(m + \lambda + 2\mu)\right] \tag{2.10}$$

$$\rho_0 \nu_{12}^2 = \rho_0 \nu_{13}^2 = \mu + \frac{\sigma_{11}}{3K}\left[m + 4\lambda + 4\mu + \frac{\lambda n}{4\mu}\right] \tag{2.11}$$

$$\rho_0 \nu_{21}^2 = \rho_0 \nu_{31}^2 = \mu + \frac{\sigma_{11}}{3K}\left[m + \lambda + 2\mu + \frac{\lambda n}{4\mu}\right] \tag{2.12}$$

$$\rho_0 \nu_{23}^2 = \rho_0 \nu_{32}^2 = \mu + \frac{\sigma_{11}}{3K}\left[m - 2\lambda - \frac{\lambda+\mu}{2\mu}n\right] \tag{2.13}$$

where ρ_0 is the density, ν_{ij} is the wave velocity propagating in direction i and polarization direction j, λ and μ are Lamé parameters, l, m, and n are Murnaghan's constants, and $K = \lambda + 2/3\mu$ is the bulk modulus. Linearization of this system of equations can be simplified as [18]:

$$\nu_{ij}^{\sigma} = \nu_{ij}^{0}(1 + A_{ij}\sigma_{11}) \tag{2.14}$$

where ν_{ij}^{σ} is the wave velocity under uniaxial stress σ_{11} at direction i and polarization direction j, ν_{ij}^{0} is the initial wave velocity when no stress is applied, and A_{ij} represents the acoustoelastic constants based on designated Lamé parameters and Murnaghan's constants.

2.6. First Arrival Time Detection

In seismic signal processing, detecting the arrival times of seismic waves is essential for localizing seismic sources and characterizing subsurface conditions. Additionally, the arrival times allow for the identification of different wave types, which is crucial for accurately constructing seismic models and developing early warning systems.

To detect the arrival times of waves, visual inspection is an option, but it becomes impractical and inefficient when there is a large number of records to analyze. To address this, algorithms were developed for automated detection, such as the threshold method, where specific thresholds are set for parameters like amplitudes and SNR. However, a major limitation is its susceptibility to fluctuations in signal and noise levels. Setting the threshold too low may lead to false detections, while setting it too high risks missing actual first arrivals.

For this research, the arrival time detection algorithm used was the Akaike Information Criterion (AIC),

developed by Akaike in 1973 [2].

2.7. Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC), developed by Akaike in 1971, is a method for quantifying the goodness of fit of a statistical model for a specific order of an autoregressive (AR) process. The goal was to find the lowest order that would fit the data best [21]. The criterion is quantified with the following equation [2]:

$$AIC = -2\ln(L) + 2k \tag{2.15}$$

where L is the maximized value for the likelihood function for the model and k is the number of parameters in the statistical model [7]. The general idea is that the most suitable model had the lowest AIC value.

This method is convenient for detecting first arrivals as it involves dividing the time series into two parts at the phase onset, determined by the lowest Akaike Information Criterion (AIC) value, providing an estimate of the first arrival time. To locate this point of division, Carpinteri et al. [7] devised an AIC-picking algorithm that only considers a user-defined window of finite length. The reason for the window is that only the arrival of the first wave, which is usually the P-wave, is measurable as structural noise, sensor interference, and wave reflections usually affect the latter phases. The following equation is used to solve for the AIC values for each point a given window w:

$$AIC(k_w) = k_w \log\Big(\mathrm{var}\big(R_w(1,k_w)\big)\Big) + (n_w - k_w)\log\Big(\mathrm{var}\big(R_w(1+k_w,n_w)\big)\Big) \tag{2.16}$$

where R_w represents the signals in the defined window, n_w is the last sample of the windowed time series, and k_w is the data point that ranges from 1 to n_w. The terms $var(R_w(1,k_w))$ and $var(R_w(1+k_w,n_w))$ are the variances of the function from 1 to k_w and 1 + k_w to n_w, respectively. The minimum AIC value obtained within this window is detected as the first arrival.

2.8. Ensemble Empirical Mode Decomposition (EEMD)

Empirical Mode Decomposition (EMD) is a data-driven empirical analysis technique developed by Huang et al. [15] for analyzing complex, non-stationary, and non-linear time series data. This method decomposes a signal into several Intrinsic Mode Functions (IMFs), which are simple oscillatory modes characterized by zero-mean amplitude and frequency modulation. Each IMF encapsulates a specific oscillatory pattern present in the original signal, and when combined, they reconstruct the original signal itself [38]. However, one of the weaknesses of the EMD method is ”mode mixing,” which occurs when a signal’s amplitudes and frequencies exhibit monotony, causing the IMFs to become incoherent. To solve this problem, Wu and Huang [34] further developed the method by adding white noise to the signal to uncover more realizations of the IMFs, which they then called Ensemble Empirical Mode Decomposition (EEMD).

To decompose a signal into Intrinsic Mode Functions (IMFs), an iterative sifting process is used. This process involves creating an upper and lower envelope of the time series using the local maxima and minima, respectively. The mean of these envelopes is then calculated and subtracted from the original signal. If the resulting difference meets the tolerance criteria, it is accepted as an IMF; otherwise, the process is repeated with the new difference. Ultimately, if the last IMF closely resembles the mean values of the envelopes within a predefined tolerance, it is considered the ’residual’. This process is

illustrated in Figure 2.7.

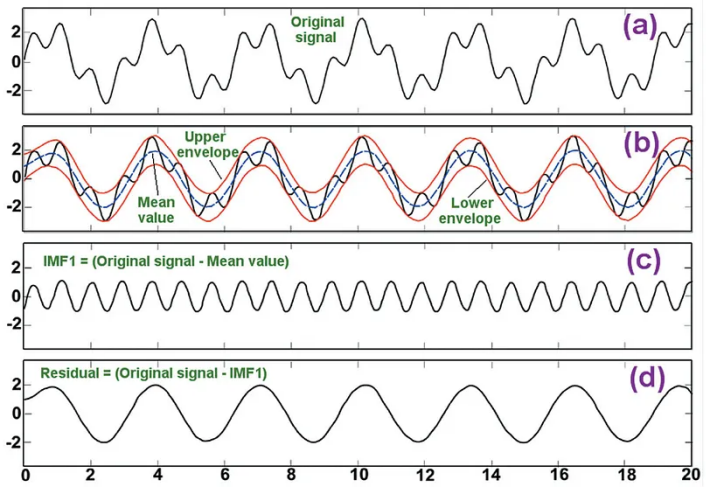

Figure 2.7: Sifting process of EMD [24].

This method was explored in the latter part of this research's analysis in an attempt to further denoise the signal and highlight more coherent features of the sound waves. However, a significant limitation of this method is its lack of a solid mathematical foundation, as its theory is primarily based on empirical techniques.

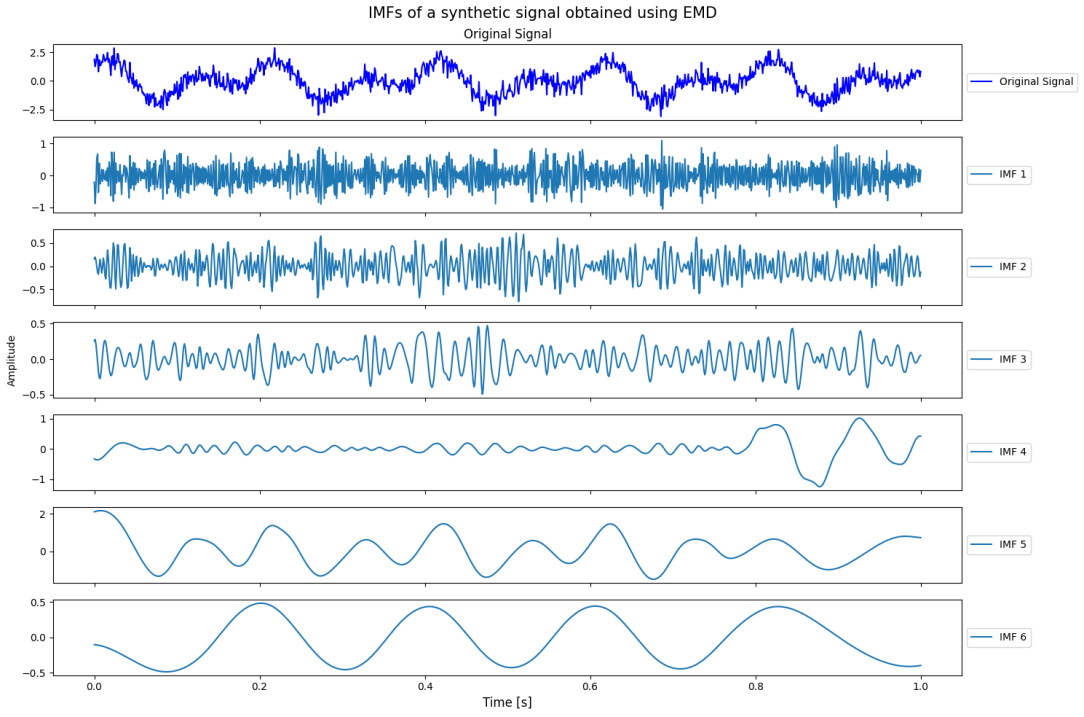

Figure 2.8: Example of EMD on a synthetic signal.

2.9. Literature Review Summary

In the table below, the significant findings found in the literature review per subject or topic and its relevance to this research is summarized.

Table 2.1: Literature review summary.

Subject/Topic	Key findings	Relevance to research
Interferometry	• In seismology, interferometry involves stacking multiple cross-correlations of recordings between a station pair to estimate the Green's function (GF) of the subsurface. This GF effectively simulates a virtual seismic signal between two sensors, allowing either sensor to act as a source or receiver [32]. • Ambient seismic noise recordings have been successfully used to produce GFs with sufficient spatial resolution to determine velocity variations [26].	• Understanding the concepts behind seismic interferometry is essential for applying it effectively in concrete structures applications. • The ideal conditions necessary to estimate a Green's function (GF) through interferometry must be understood in order to accurately assess and analyze the resulting estimations from the validation datasets and ambient traffic noise data.
Signal pre-processing	• Temporal normalization is used on raw signals to mitigate the impact of instrumental irregularities, non-stationary events [4], and transient events [13]. • Spectral whitening of ambient noise aims to broaden the signal's frequency band by equalizing or normalizing signals in the frequency domain [4].	• For any given signal dataset, determining the necessary and optimal pre-processing scheme is important for effectively extracting the desired features. Minimizing the effects of external influences is essential to achieve more coherent results.
Wave Propagation in Concrete	• Wave propagation in concrete is analyzed using the principles of acoustoelasticity in elastic solids to discern properties of the medium and detect defects [12]. This is achieved by measuring the velocities of artificially induced impulses traveling from a source to a receiver at known distances. • Assuming that deformations of concrete structures are small relative to the structure's size, the theory of linear elasticity may be applied.	• This concept is crucial for the practical application of this research, as wave speeds are directly related to wave arrivals, which is analyzed for the investigated concrete structures.

Subject/Topic	Key findings	Relevance to research
Wave First Arrival Detection	• In signal processing, the detection of arrival times enables the localization of sources, characterization of the medium's conditions, and identification of different wave types. • Carpinteri [7] developed an arrival time picking algorithm based on the Akaike Information Criterion (AIC), originally proposed by Akaike [2] to assess the goodness of fit of a statistical model for a given order of an autoregressive (AR) process. The AIC parameter is computed for each point within a predefined window length, and the point with the lowest value is selected as the first arrival time.	• In the practical application of this research, the focus is on detecting the arrivals observed from the estimated Green's function through ambient noise interferometry. To ensure consistency and objectivity in arrival time detection, an automatic picking algorithm is essential.
Traffic Noise Interferometry State of the Art	• Salvermoser et al. [25] used Passive Image Interferometry (PII), a method that combines noise correlation and Coda Wave Interferometry (CWI), to infer changes in a concrete bridge by examining velocity variations. They analyzed the 'coda' obtained from the GF estimates and found that traffic noise could potentially be used for monitoring purposes. However, they noted poor spatial resolution due to the low sensitivity of the instrument used.	• This study aimed to address several gaps in research, including the methodology of data collection, the instruments employed and their corresponding sensitivities, the frequency ranges considered, and the analysis approach for ambient traffic noise. Additionally, the study aimed to propose a method for localizing damages.

3

Methodology

This chapter introduces the datasets used and their corresponding collection methods, the algorithm building and optimization, the role of the algorithm in post-processing, and the methods used for analyzing the data.

As mentioned in section 1.1, this research used two datasets: one from a laboratory experiment and another from field recordings on a concrete tunnel. These datasets will be further explained in the following sections.

3.1. Validation Dataset

The dataset used to validate the proof of concept and initially form the algorithm was collected from an acoustic emissions (AE) experiment conducted on a precast concrete girder in the Stevin laboratory at TU Delft. This involved repeatedly dragging a chain back and forth along a straight line between two points as the loads applied on the concrete girder were increased incrementally.

Although the experiment (refer to Figures 3.1 and 3.2) was mainly intended to measure AE parameters focusing on amplitudes, energies, and AE event counts, the data was still ideal for assessing the concept of ambient noise interferometry on concrete structures for the following reasons: 1) there were no unexpected random noise sources, as the only source was the chain being dragged; 2) the experiment continued until cracks formed, allowing for comparisons of signal behavior between undamaged and damaged states; and 3) the locations of damage could be pinpointed since the experiment was conducted in a controlled laboratory environment, allowing for verification if a change was observed in the signal.

3.1.1. Experimental setup

The concrete member used in the experiment was a 15m-long precast concrete girder, consisting of two separate members — one 11.25m long and the other 3.5m long. These were then made continuous using a topping and a cross-beam, both cast in situ (refer to Figure 3.1). One support was positioned beneath the cross-beam, while the other was situated at the further end of the girder, away from the cross-beam, effectively creating an overhanging portion. The loading points were located 0.63 meters from the end of the overhang and 4.77 meters from the opposite end of the girder, positioned within

the middle third of the fully-supported section. In Figure 3.1, the translucent gold and red rectangles represent the sensor zones for the top layer and web sensor groups, respectively.

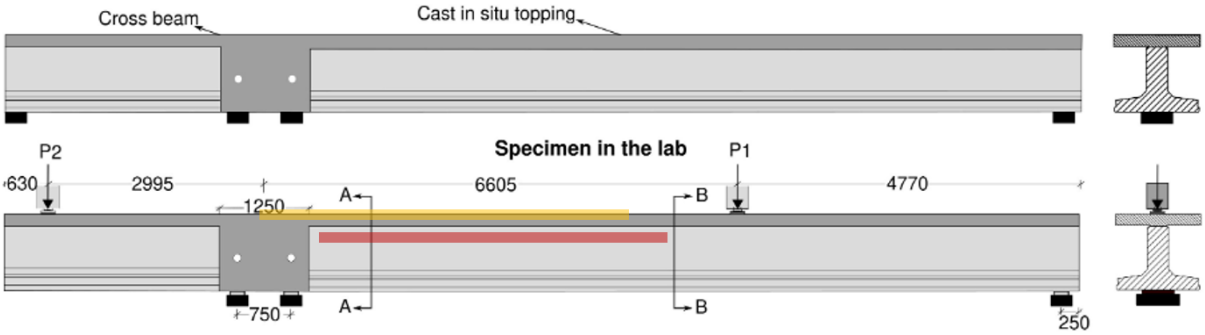

Figure 3.1: Dimensions of the precast concrete girder in millimeters (mm) [9].

To generate the ambient noise, a heavy chain was pulled across the surface of the girder at a consistent velocity for each load increment. The chain was moved along the surface starting adjacent to one of the hydraulic loading jacks (P2), then across to the other jack (P1), and back again.

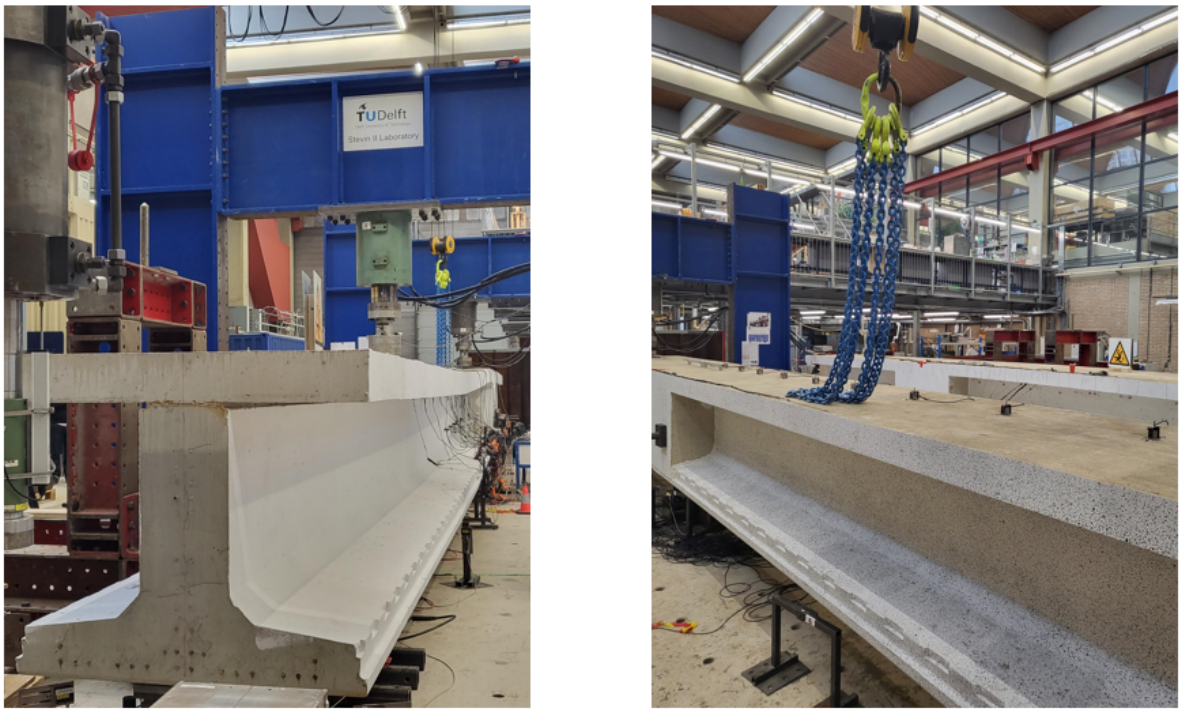

Figure 3.2: (left) Cross-section of the girder and (right) chain dragging on top of the girder [9].

3.1.2. Sensor layout and details

Piezoelectric sensors were used in the experiment, positioned on top of the girder (sensors 19 to 27) and along the web (sensors 11 to 18), spaced at various distances from each other. They were aligned in a straight line, parallel to both the girder's length and the chain's path. With an operating frequency range of 40-100 kHz, the minimum sampling frequency required to prevent aliasing 200 kHz, which was well within the sampling frequency used for the experiment which was 2.5 MHz. Despite the sensor's upper operational frequency limit, a higher sampling rate was still preferred to achieve better temporal resolution for the samples.

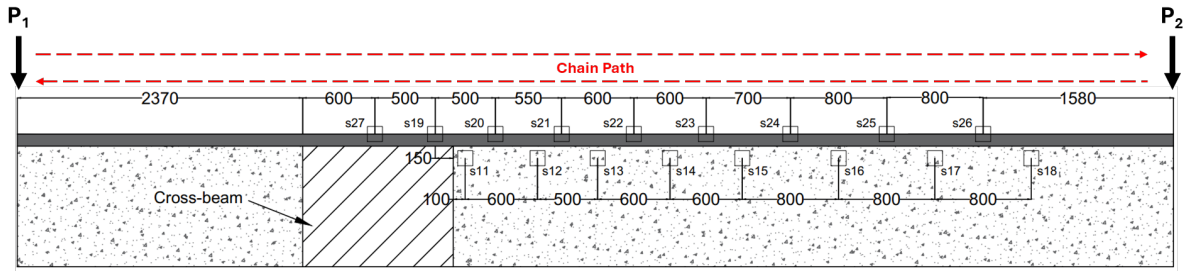

Figure 3.3: Layout of sensors attached to the concrete girder and the two loading points, P_1 and P_2.

3.1.3. Data collection

The data collection method involved recording on a ”hit” basis, resulting in a discretely sampled dataset. Initially, all sensors were activated and prepared to receive signals before the chain was dragged. Upon detection of a sound wave that exceeded the amplitude threshold of 40 decibels (dB) for any given sensor, recording started and lasted for approximately 1.64 milliseconds, constituting one hit. The time of detection relative to the very first signal recorded by any of the sensors was then also noted. If during the recording the wave's amplitude fell below the threshold, the amplitude of the recorded signal flattened (as seen in Figure 3.7), and the recording continued until the hit's duration was done. This hit-recording process continued until the chain returned to its initial position, and the procedure was repeated for each load case.

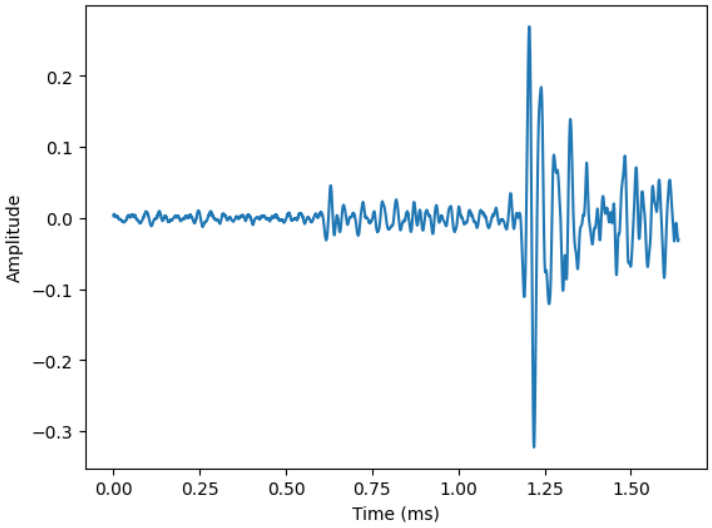

Figure 3.4: Example of one hit (AE Chain Experiment).

3.2. Ambient Traffic Noise Dataset

To test the algorithm's applicability to real-world ambient traffic noise, data was collected from the Maastunnel, an underwater tunnel in Rotterdam that accommodates pedestrians, cars, and bikes. Sensors were placed under the road slab in the middle of the tunnel, specifically beneath the portion where cars passed, capturing traffic noise. Data was collected twice within a two-hour window of typical traffic, from 11 AM to 1 PM. Both recordings had a duration of about 50 minutes each.

The challenge with this dataset, compared to the validation dataset, lay in the many potential sources of noise, such as waves hitting the exterior of the tunnel, power lines, and the presence of pedestrians and bikers, since these recordings were made in the field. Additionally, the spatial distribution, the signal

magnitudes, and the frequency spectrum were inconsistent and unknown, making this dataset more complicated.

3.2.1. Sensor layout and details

The sensors were strategically placed underneath the 350 mm-thick road slab where motorized vehicles traversed in a single direction. As shown in Figure 3.5, this location was ideal because it allowed for easy access and avoided any road closures or disturbances during installation and testing.

For the sensors, the same piezoelectric-based (PZT) sensors mentioned in the validation dataset in Section 3.1.2 were used. The sampling frequency was set to 2 MHz to ensure higher temporal resolution for the recorded signals.

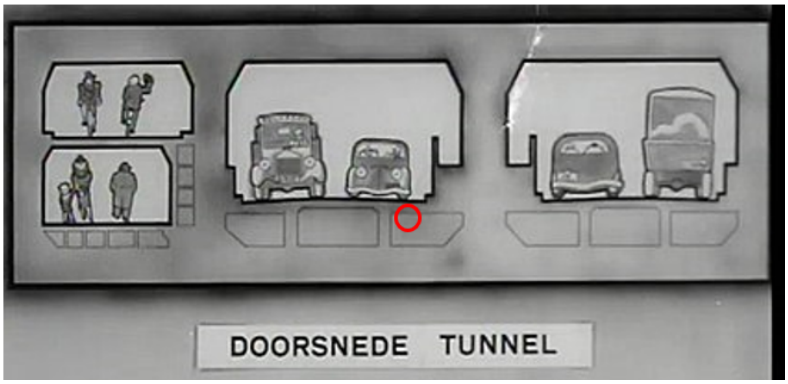

Figure 3.5: Cross-section of the Maastunnel [30].

In the road section, five sensors were installed along a straight line parallel to the flow of traffic, equally spaced at 0.5 m. Theoretically, this made the analysis simpler as the energy from the waves would travel along the same path. In essence, this dataset is similar to the validation dataset in terms of the major direction of energy and sensor placement. However, very different results were expected with the Maastunnel dataset because the spatial distribution of sources was unknown. Additionally, the road slab in the Maastunnel is part of a much larger structure, allowing the sensors to detect noise from various sources.

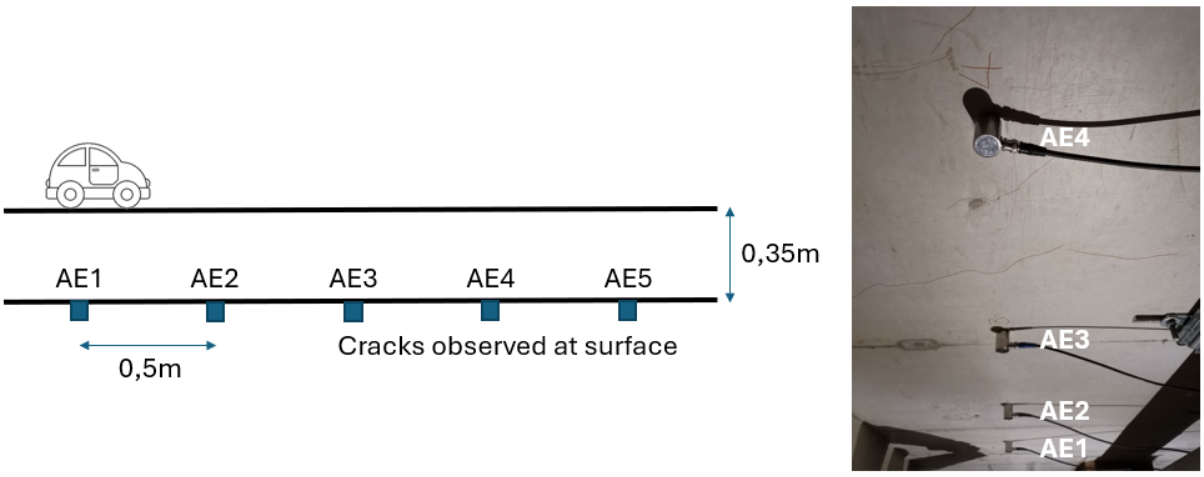

Figure 3.6: Layout of sensors underneath the Maastunnel road section.

3.2.2. Data collection

The data collection method was identical to that of the validation dataset. The sensors were activated to start recording but were configured to only record on a ”hit” basis. Sound waves were recorded as ”hits” once the sensors detected an amplitude higher than the threshold of 45 decibels (dB), and recording continued for approximately 5.12 milliseconds. When the amplitude of the hit went below the threshold, the signal flattened, and the recording persisted until the consistent duration of the hit was maintained. In addition to the amplitude content of a given hit, the absolute recording times of each hit with respect to the very first hit were also recorded.

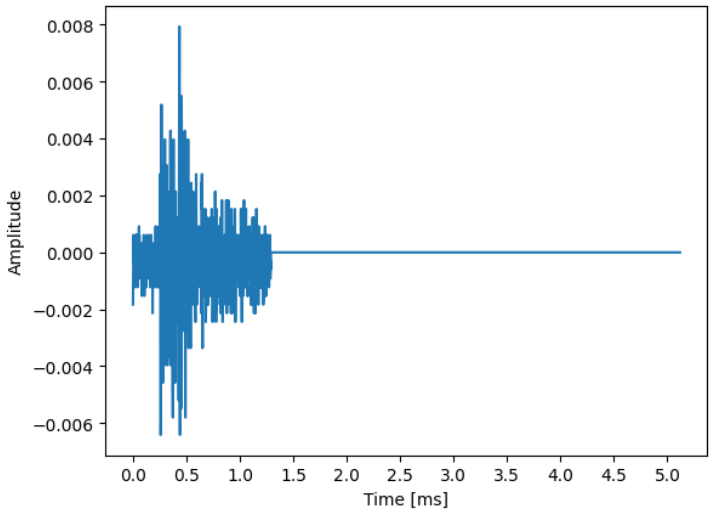

Figure 3.7: Example of one hit (Maastunnel).

3.3. Algorithm Development and Functionality

The post-processing of the data was done with a Python script, whose development and functionality are elaborated upon in the subsequent sections.

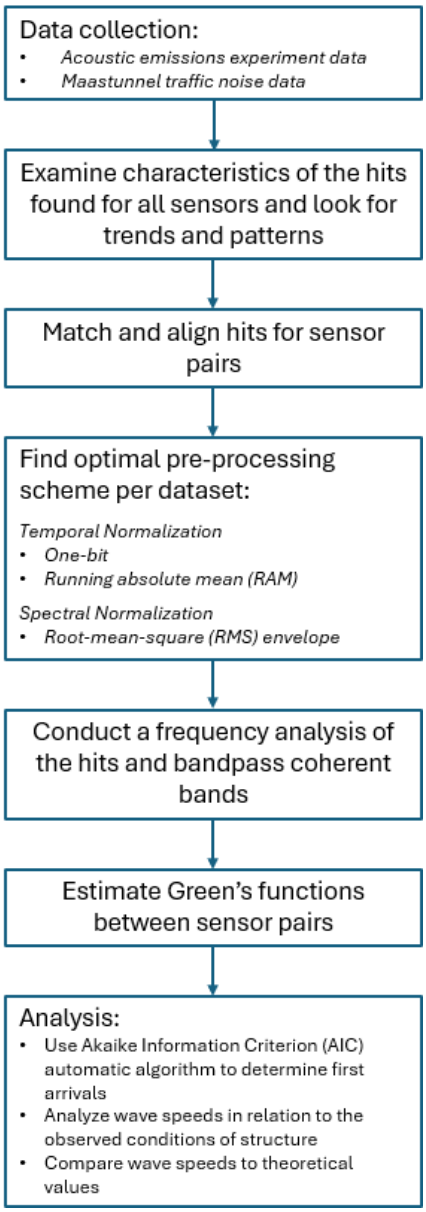

Figure 3.8: Flowchart of methodology

3.3.1. Process of matching and aligning ambient noise data

When conducting interferometry in seismology, all signals received by a given station (A) are cross-correlated with signals from another station (B), including those received by only one of the two stations, which could result from the source location, signal direction, or reflections and trajectory changes caused by subsurface heterogeneities. The issue arises when signals received by the two stations are not from the same wave but are still cross-correlated. Since cross-correlation measures the similarity between signals recorded at two different locations, if the signals do not share the same physical path, the resulting CCF will likely contain spurious arrivals. These spurious arrivals make the resulting CCF, unphysical, meaning it does not correspond to any real physical phenomenon [27]. Consequently, no meaningful

information about the medium between the stations can be derived, making it unhelpful towards the estimation of the GF. However, because a large number of signals are usually considered for each station pair and it is assumed that signals originate from all directions, the incoherent parts of the CCFs may cancel out after stacking them over a long period. Ideally, this will result in an estimation of the GF of the medium, which contains all types of waves, including body waves, surface waves, and multiply-scattered waves, between a station pair.

In this research, because both datasets are limited in terms of number of samples, duration of recording, and distribution of sources, together with the fact that a single concrete girder and a concrete tunnel are in much smaller scales compared to the earth's subsurface, it is difficult to assume that simply cross-correlating all the signals received by both station pairs would eventually normalize, though it is possible if data for longer periods of time is available.

To address these limitations, a signal-matching scheme was used to only consider direct waves that travel from one station to another, with a small tolerance for indirect waves that reflect towards the second station. This was done by comparing the arrival times of all the hits recorded by each station for a given station pair. The differences in arrival times were then compared against the expected travel time of waves for a specific distance, with reference to a conservative estimate of the speed of sound in concrete, typically around 4000m/s.

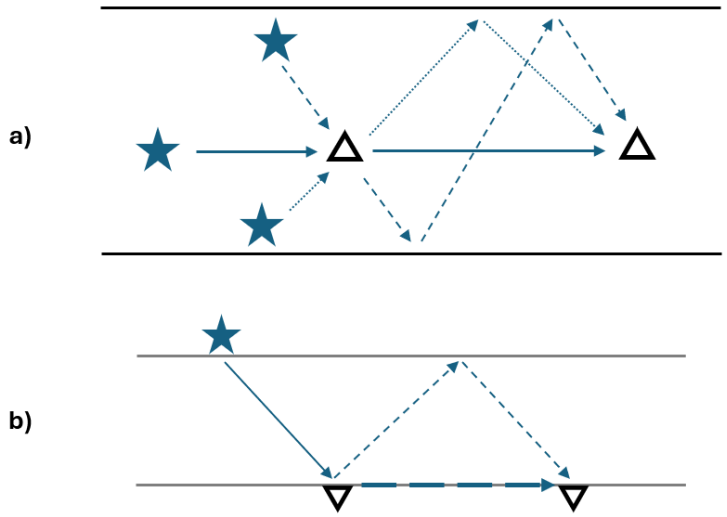

Figure 3.9: Examples of considered signal trajectories (arrows) between sources (stars) and receivers (triangles) for sensors that are a) on the same surface as the noise sources (Top View) and b) on different surfaces (Side View).

In Figure 3.9a, the possible trajectories for surface waves depending on the location of the source are shown. In Figure 3.9b, trajectories are depicted for scenarios where the sources are on a different surface than the sensors. As seen in this figure, the original wave can either be fully reflected (small-dashed line), fully converted into a surface wave (large-dashed line), or partially reflected and partially converted into a surface wave.

After the hits from a given station pair were matched, they were then aligned by adjusting the later-received signal with respect to the earlier-received one. This adjustment involved shifting the samples of the later-received signal by an amount equal to the absolute difference between the recording times of the hits at the two stations, divided by the sampling rate. This synchronization was necessary because each hit was only recorded once a wave was detected by the sensor and was not continuously recorded.

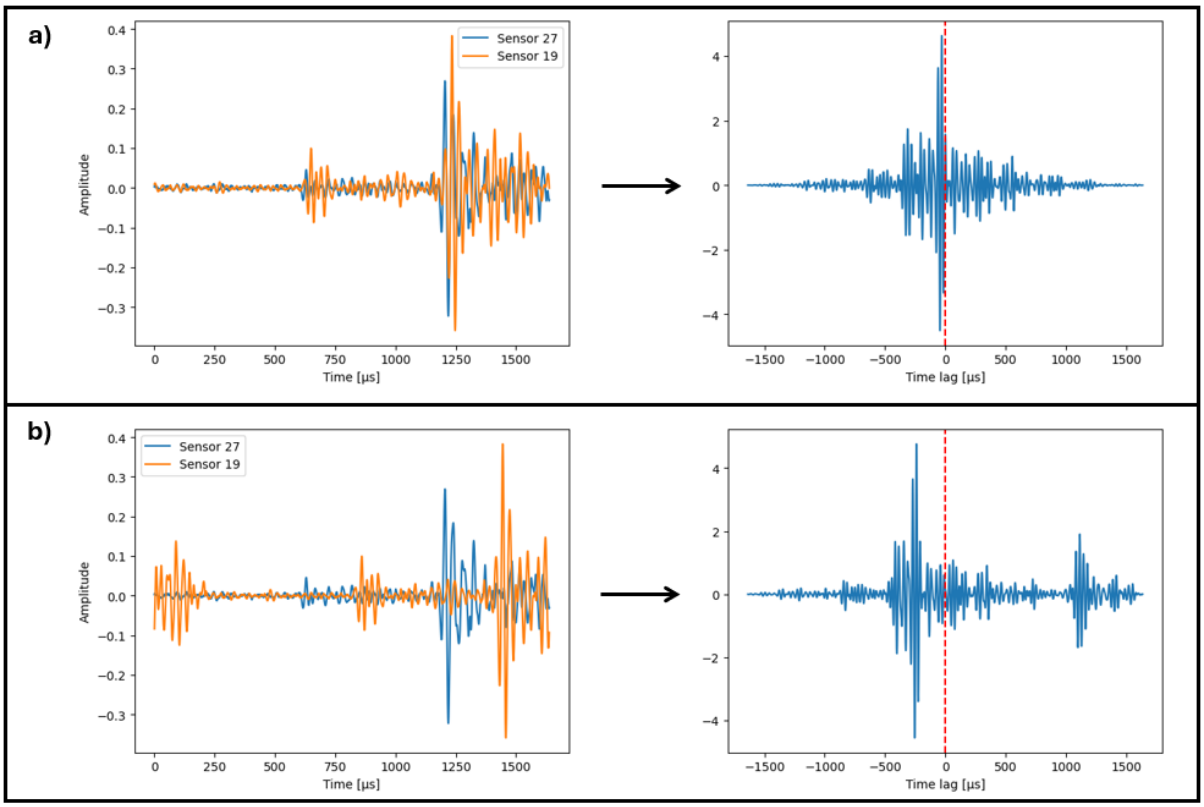

Figure 3.10: a) Unsynchronized signal pair and resulting CCF and b) synchronized signal pair and resulting CCF.

3.3.2. Data normalization and frequency analysis

Once the matched hits were synchronized, temporal and spectral normalization were done to remove outliers and to smooth out the signals. Given the limited literature available on either of the investigated datasets, two different temporal normalization methods — one-bit normalization and running-absolute-mean (RAM) normalization — were tested to determine which would yield more coherent CCFs. For spectral whitening, the only method tested was whitening of the signal in the frequency domain through the use of an envelope function. This function is expounded on in Section 2.2.2.

Afterwards, the frequency distribution of the resulting CCFs from the chosen normalization scheme were investigated to ensure that only the frequencies with coherent signals were included in the further steps. This was done by first checking the frequency spectra of the signals then subjecting them to a series of low-pass, high-pass, and band-pass Butterworth filters to find the optimal range.

3.3.3. Estimating the Green's function

After determining the most appropriate pre-processing scheme for each dataset, signals between a given station pair were cross-correlated and stacked. The stacking method used was simple averaging of all the computed CCFs for each station.

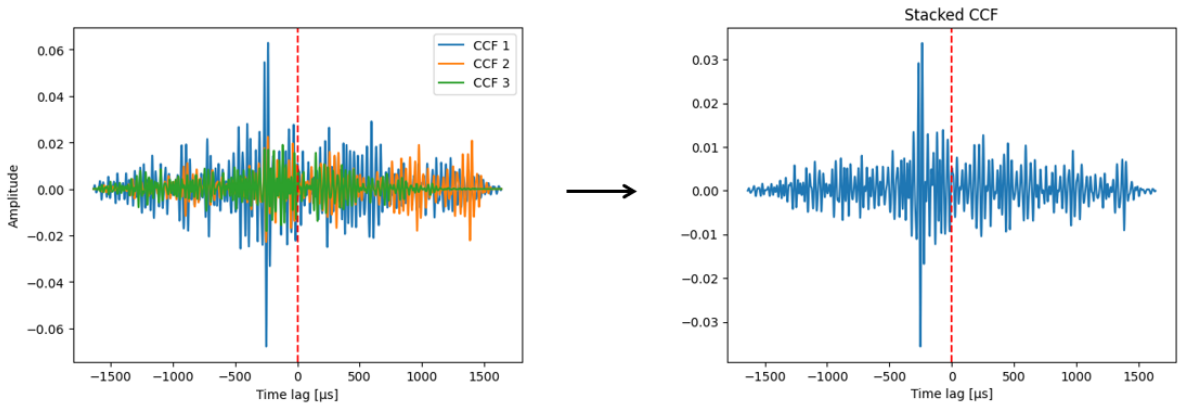

Figure 3.11: Example of stacking of 3 random CCFs from the validation dataset.

3.4. Analysis of Results

For each dataset, the analysis began after obtaining estimations of the GF for all sensor pairs. The estimated GFs were first subjected to visual inspection to evaluate their coherence and symmetry around the time lag of zero. The results were then compared with the expected behavior, considering the experimental setup, the condition of the structure, and the assumed distribution of noise sources, such as typical wave velocities in concrete (refer to Section 2.4.1) and possible asymmetry in the GF estimation if more waves were expected from one side of a sensor pair (refer to Section 2.2.3).

Upon visually evaluating whether a GF estimation was coherent enough, it was then tested to determine if valuable information regarding the waves and the medium could be derived. By observing the arrival times of waves and their phases, corresponding wave speeds could be calculated.

For GF estimations that exhibited clear arrivals, an automatic arrival time detection algorithm was necessary because visual inspection alone was unreliable due to its subjectivity. For this purpose, the AIC algorithm was employed (refer to Section 2.7).

A limitation of the AIC algorithm is that it will always detect an arrival, as it bases its detection on the AIC values found at every point within a predefined window. Therefore, it was important to use the algorithm in conjunction with critical visual inspection to ensure accuracy. Refer to Figure 3.12.

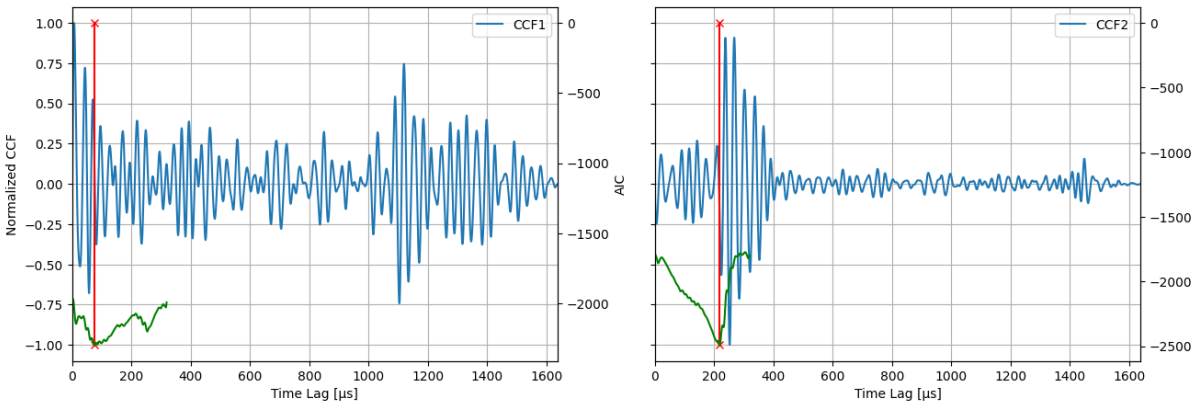

Figure 3.12: Example of AIC automatic detection algorithm finding arrivals irrelevant of a signal's coherence.

The wave speeds calculated for different station pairs were compared to each other and to the theoretical wave speeds for consistency and patterns. Additionally, they were assessed based on the prevailing

structural condition during testing. For example, in the validation dataset, wave speeds were compared between a sensor pair before and after cracks were detected. This information is crucial for localizing and detecting damage in the structure.

4

Validation Dataset Results and Discussions

Initially, the analysis conducted on the validation dataset focused on the simplest scenario, which was the 0 kN load case. This was done to try and limit the number of uncertainties while finding the optimal pre-processing scheme to use. When considering loaded structures, it is important to take into account that there could be damages within that already exist or are forming internally that still cannot be seen from the exterior because these can result to unexpected behaviors in the resulting GF estimations. Additionally, the acoustoelastic effect of waves travelling through elastic media should also be considered as this may also influence the signal. Generally, a medium under compression exhibits faster wave speeds, while one under tension exhibits slower wave speeds (Section 2.5).

After having decided on the optimal scheme and the correct frequency bandwidth, the different load cases were considered to see whether expected patterns would emerge given the damage condition and stress state at different loads.

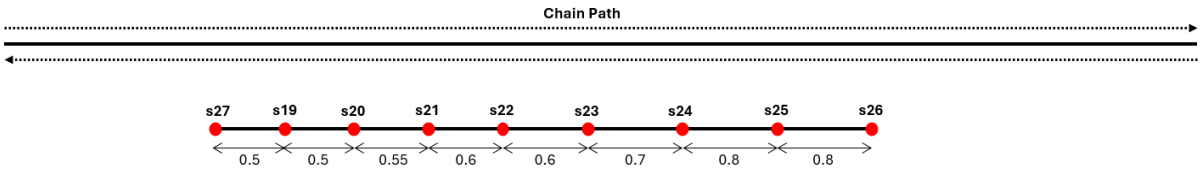

Figure 4.1: Top view showing the layout of the top sensors and the path of the chain (signal source).

4.1. Dataset Characteristics and Optimal Pre-processing Scheme

The first aspect analyzed in the dataset was the distribution of hits based on the expected arrival times, given the known path of the chain. It was anticipated that the number of hits for each sensor would peak as the chain passed by, since no other noise sources were introduced into the experiment. To visualize this, histograms were plotted for each sensor group (on top of the slab and on the web), where the y-axis represented the number of hits, while the x-axis represented the absolute recording times of each hit relative to the very first hit recorded by the leftmost sensor (sensor 27), not the moment when the chain first started moving. The plots, arranged from top to bottom starting with the sensor closest to

the chain's origin and ending with the farthest, are shown in Figure 4.2. Each histogram had a bin width equivalent to 5 seconds.

Since there was no synchronized video recording while the chain was being dragged, the exact times when the chain was adjacent to any of the sensors were unknown, as this was not considered an important parameter during the experiment. However, using the available recordings, the known total elapsed times, and the known distance traveled by the chain, an estimated chain speed of roughly 0.134 m/s was deduced.

With this chain speed, the position of the chain could be inferred. The peak in the histogram was first visually observed for the very first sensor (sensor 27). Using this observation and the estimated chain speed, the times the chain passed the succeeding sensors were then determined. These times are marked with red dashed lines in the subsequent figures.

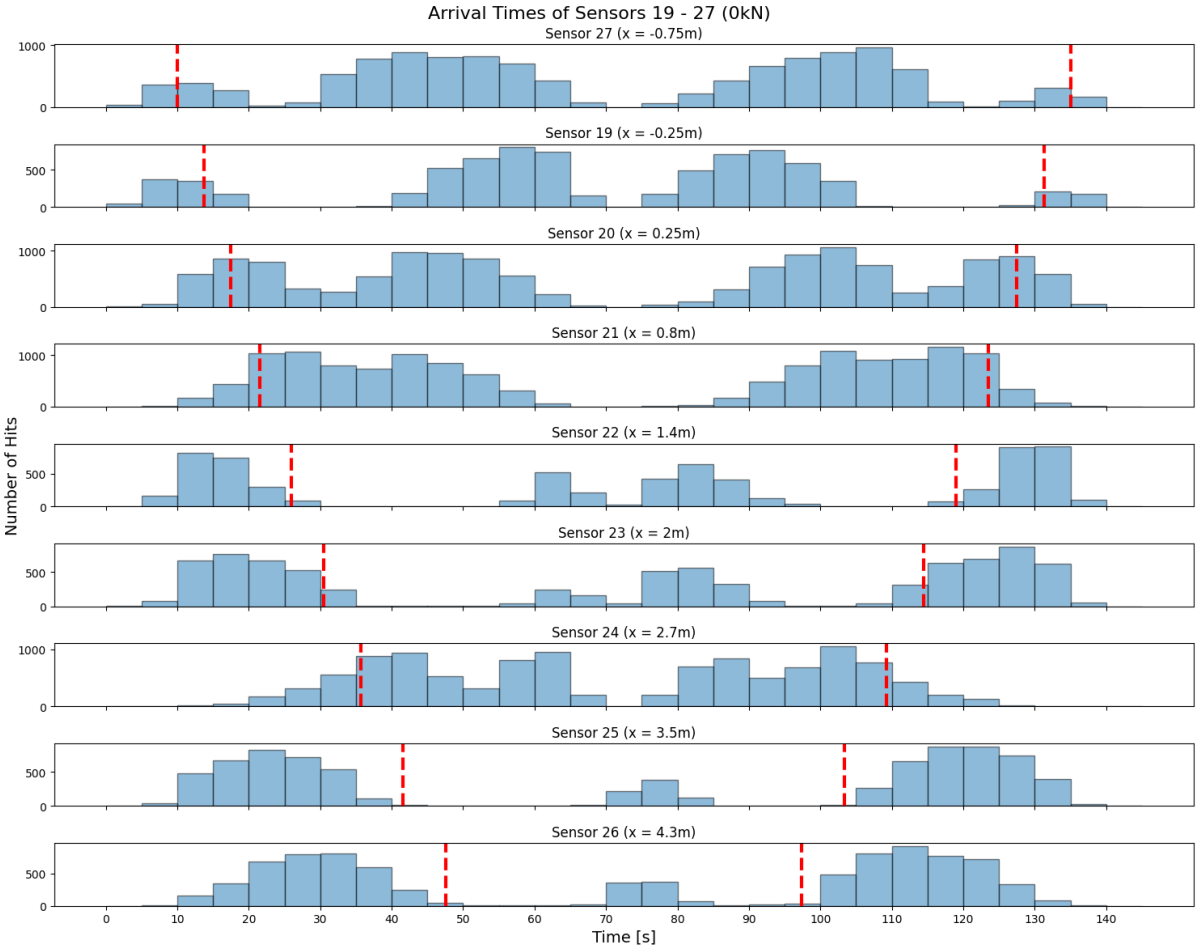

Figure 4.2: Hits for sensors 19-27 (top layer sensor group) for the 0 kN load case.

In Figure 4.2, a peak for sensor 27 is observed around the 10-15 second bin. It is assumed that the chain passes the sensor at 10 seconds within this bin. Given that the adjacent sensor (sensor 19) is 0.5 meters away, the estimated time the chain passes it is around 13.7 seconds, which still falls within the same bin. Using this approach for the subsequent sensors, the movement of the chain is traceable until sensor 21. However, the trend is not reflected for sensors 22 and 23. The trend resumes for sensor 24, but again disappears for sensors 25 and 26.

While some sensors did not show the expected trend, those that did exhibited a reverse pattern, moving from left to right and then back again, which was the anticipated behavior.

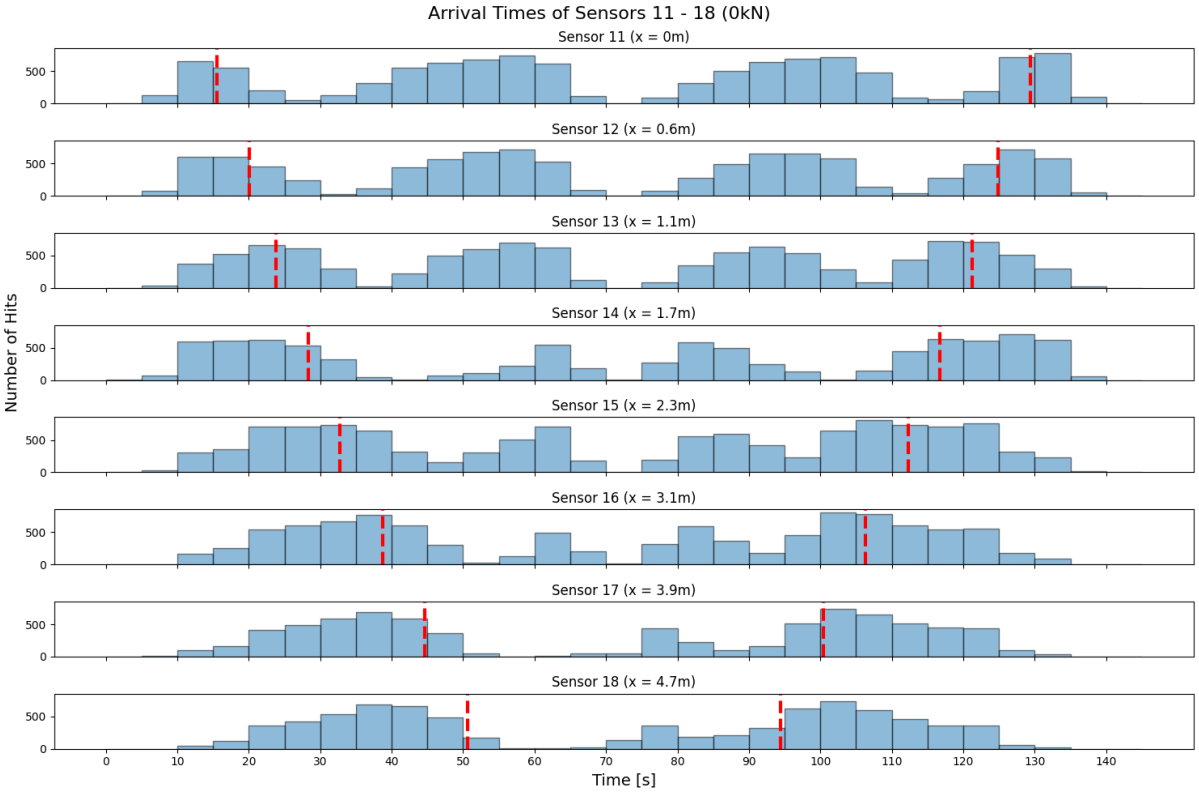

Figure 4.3: Hits for sensors 11-18 (web sensor group) for the 0 kN load case.

In Figure 4.3, the assumed time when the chain is nearest to sensor 11 is not based on the peak in the histogram but rather on the previously assumed time for sensor 27. Given that they are 0.75 meters apart along the longitudinal axis of the beam, the estimated time that the chain passes sensor 11 is around 15.6 seconds. Although this does not coincide with the peak, a consistent trend from the leftmost sensor (11) to the rightmost sensor (18) of the web sensor group is observed, effectively portraying the relative location of the chain. This pattern also reverses as the chain moves back to its origin.

This observation is significant because it demonstrates that despite the sensors on the web being much farther from the source in the vertical direction and the sound waves needing to pass through the interface between the web of the precast girder and the cast-in-situ slab, the distribution of signals remains consistent with expectations.

In both Figures 4.2 and 4.3, it is observed that after the initial peaks in the histograms caused by the chain passing next to the sensors, there are sudden drops in hits before they increase again. This pattern likely occurs because as the chain moves further away from a sensor, the signals undergo more attenuation, reaching a level undetectable by the sensor until louder sounds are generated by the chain changing directions. This change in direction results in the subsequent increase in hits near the middle of the histograms.

The next step was to evaluate the optimal normalization scheme using stacked and band-passed cross-correlations of matched signals between a given station pair. This approach was chosen because relying on the cross-correlation of only one or a few pairs might not have provided a representative sample.

Some cross-correlations might have been unphysical, especially if the signals were incorrectly paired (see Section 3.3.1). By stacking multiple CCFs, the noise produced by the unphysical cross-correlations was averaged out, and the more significant and recurrent features were emphasized.

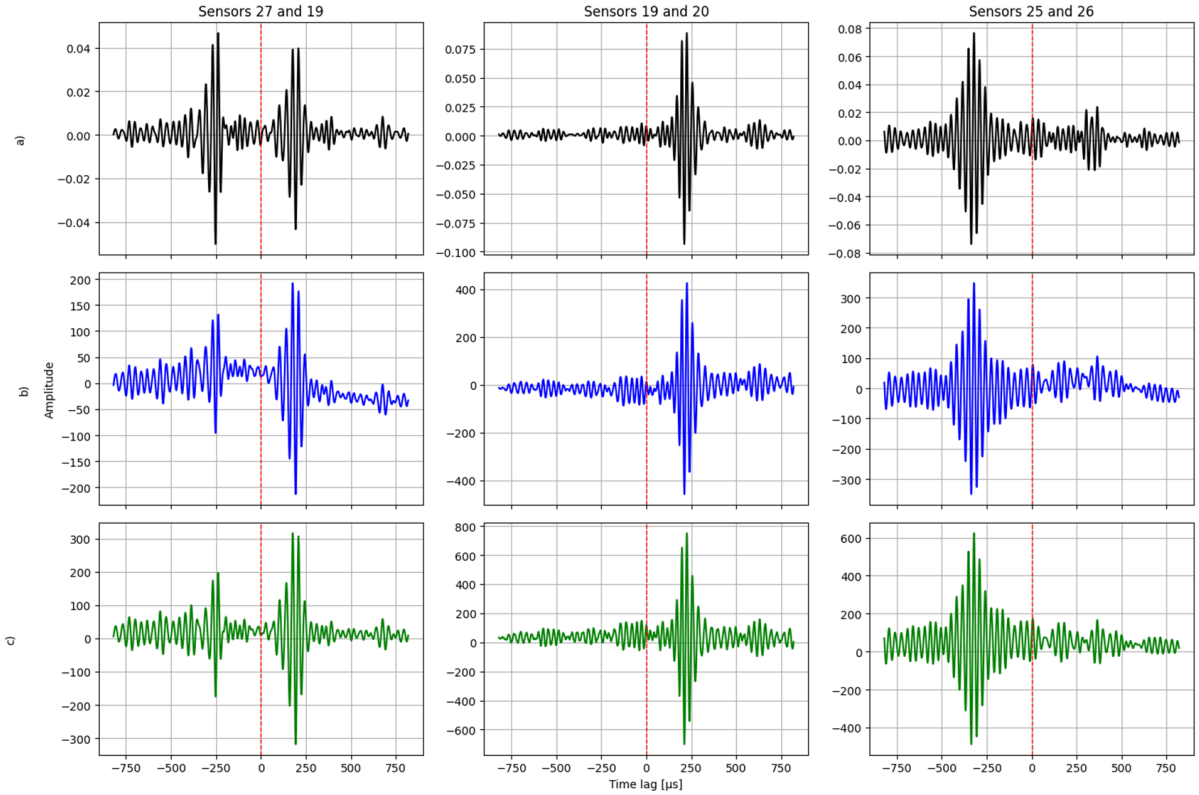

Figure 4.4: Raw (a - black), one-bit normalized (b - blue), and running-absolute-mean normalized (c - green) stacked cross-correlations of different sensor pairs.

In Figure 4.4, it is observed that the stacked CCFs for the three pairs shown (27-19, 19-20, and 25-26) have asymmetric arrival times around time lag = 0, with only pair 27-19 having significant energies on both the causal and acausal parts. For sensor pairs 19-20 and 25-26, the asymmetry of their CCFs (as explained in Section 2.2.3) indicates that the distribution of sources is more concentrated on one side of the sensor pairs than the other. Given the positions of these two sensor pairs relative to the path of the chain, it is logical that their CCFs exhibit higher concentrations on opposite sides of time lag = 0: pair 19-20 has higher amplitudes and more coherent phase information in the causal part, while pair 25-26 shows this in the acausal part.

However, there is an inconsistency with pair 27-19. Despite being positioned further to the left than pair 19-20, it shows high energy distributions on both the causal and acausal parts of the CCF. A possible explanation for this anomaly is that pair 27-19 is situated directly above a cross-beam, unlike the other pairs which are just located on top of the web of the girder (refer to Figures 3.1 and 3.3).

Among the three temporal normalization schemes considered (no normalization, one-bit normalization, and RAM normalization), the CCF showing the clearest arrivals was obtained from the original, non-temporally-normalized data, followed by the RAM-normalized data. The one-bit-normalized data, on the other hand, emphasized the low-frequency component of the CCF, resulting in a wavelike appearance, as seen in the second row (b) of Figure 4.4.

Next, spectral whitening was applied to the non-temporally-normalized data, as well as in conjunction

with the two temporal normalization methods, to further explore the optimal pre-processing approach for this dataset.

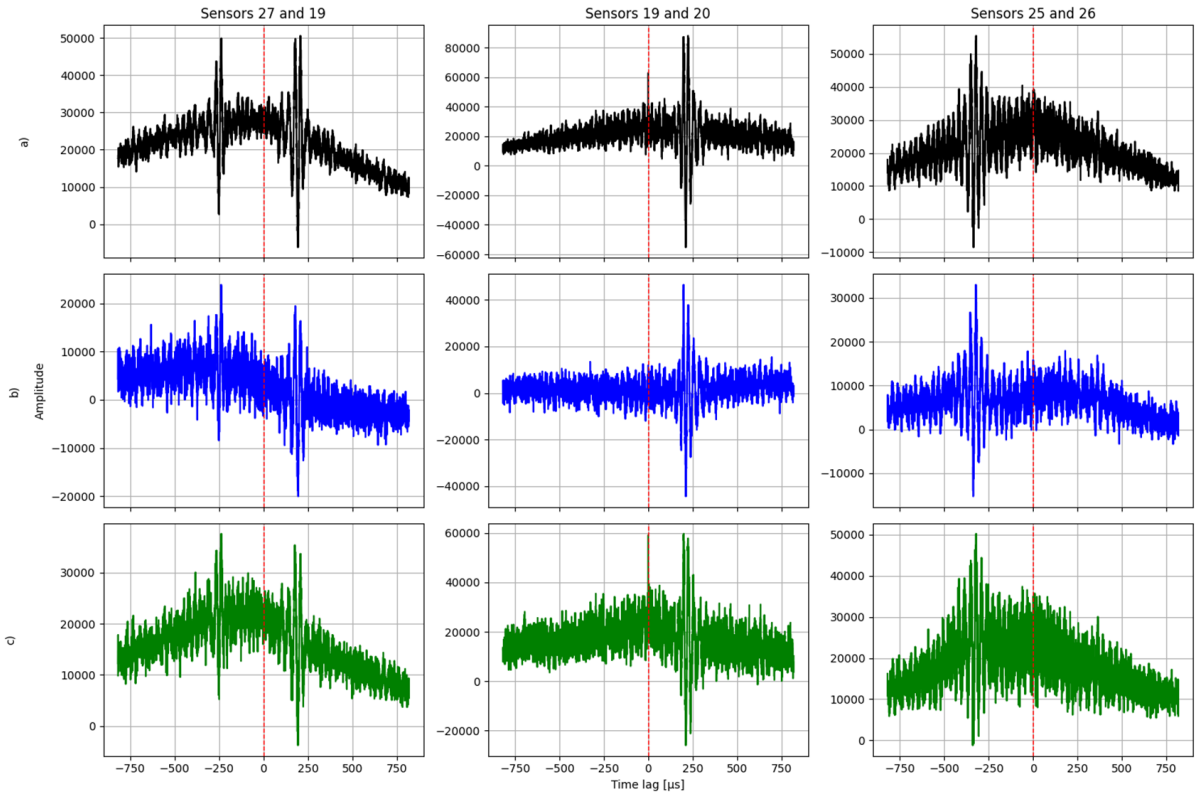

Figure 4.5: Spectrally whitened raw (a - black), one-bit normalized (b - blue), and running-absolute-mean normalized (c - green) stacked cross-correlations of different sensor pairs.

In Figure 4.5, it is observed that spectral whitening successfully emphasizes the peaks of the CCFs in the time domain, compared to the plots in Figure 4.4. However, because spectral whitening equalizes the frequency distribution of a signal, the low-frequency component observed in Figure 4.4b becomes apparent in all the plots in Figure 4.5. Additionally, more high-frequency components are introduced, which tend to dominate the time series, making it look noisier and obscuring the arrivals of the higher-amplitude events.

Upon comparing all possible combinations of normalization schemes, it was determined that the spectrally whitened raw stacked CCFs for all station pairs exhibited the highest coherence, with the clearest arrival times and the most emphasized peak amplitudes in the time domain. A plausible reason for this is that the data originated from a controlled laboratory experiment where no other noise sources were present besides the chain being dragged. There were no significant external natural influences or unexpected high-amplitude events that could obscure the signal.

After determining the optimal normalization scheme, the frequency spectra of the raw signals were then analyzed to identify the frequency range that provided the clearest arrivals.

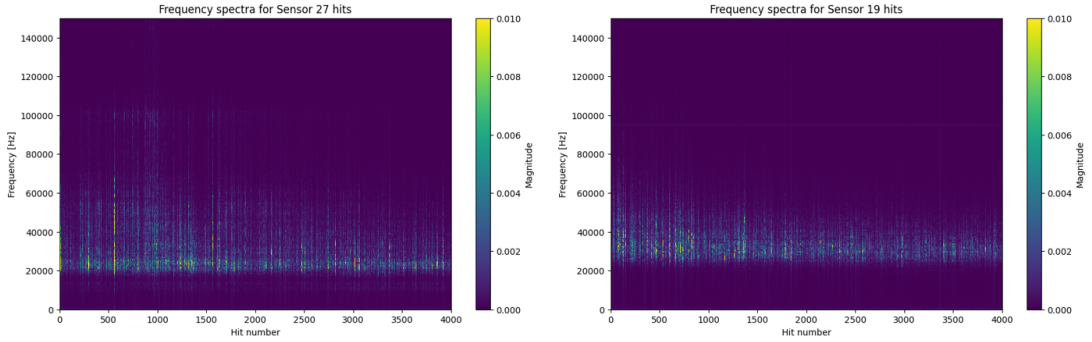

Figure 4.6: Frequency distribution of hits of the top layer sensors.

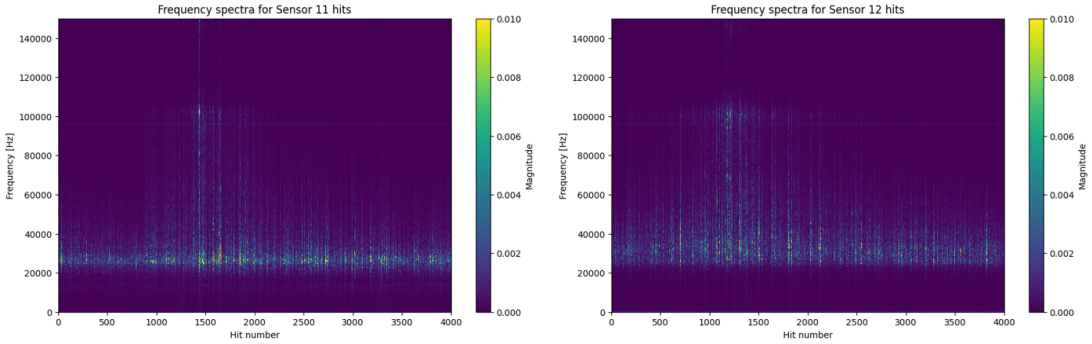

Figure 4.7: Frequency distribution of hits of the web sensors.

From Figures 4.6 and 4.7, it can be seen that the sensors on the top layer and the web mostly show signals in the 20 - 50 kHz range, with a few exceeding 100 kHz. Other frequencies are less prominent, which is addressed through spectral whitening.

To determine the frequencies to investigate based on their coherence and significance in contributing to the resulting CCFs, and to understand their overall effect on the GF estimates, the matched, aligned, and spectrally whitened hits from sensors 19 and 20 were cross-correlated and stacked. This combined signal was then subjected to a series of low-pass and high-pass filters to identify the coherent frequency range and denoise the signals. Notably, the point at which frequency filtering is performed is irrelevant, as the results are the same whether it is done before or after spectral whitening (see Figure A.3 in the Appendix). For better visual representation, the CCFs below were spectrally whitened beforehand.

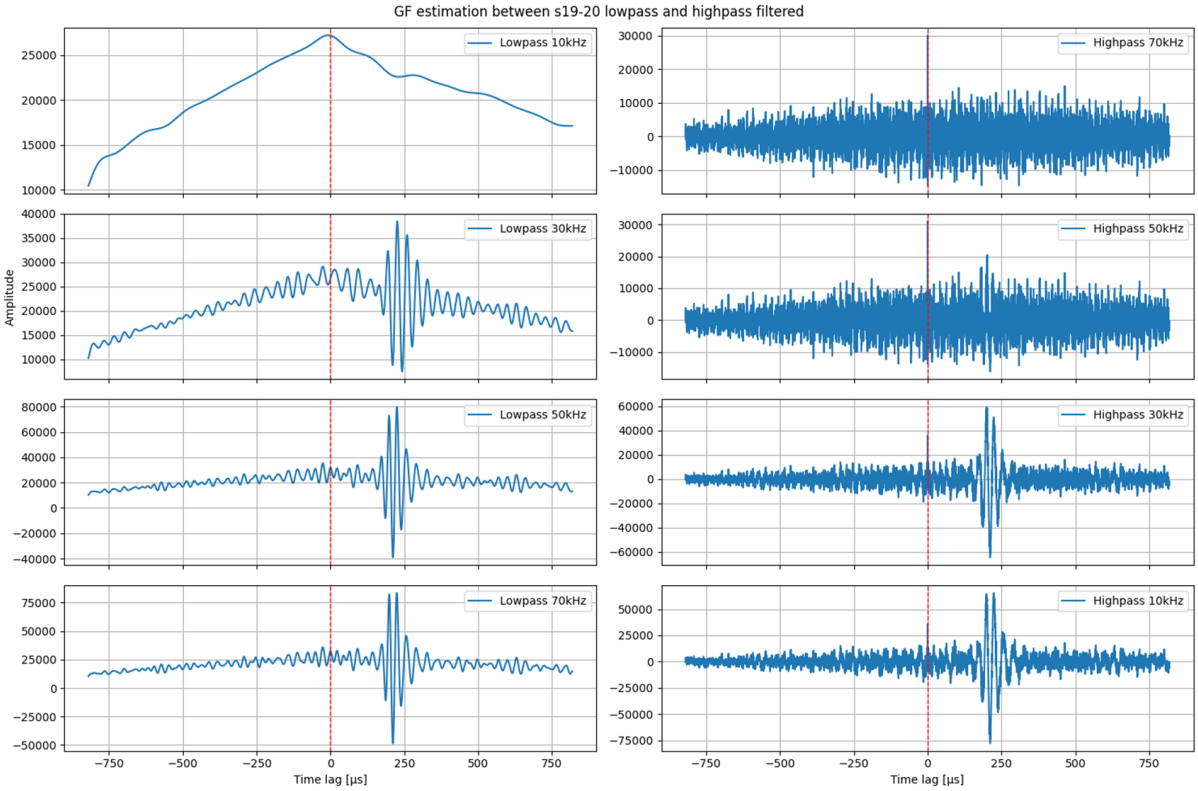

Figure 4.8: Low-passing and high-passing the Green's function estimate for sensor pair 19-20.

In Figure 4.8, it is observed that after applying a low-pass filter at 50 kHz, no significant changes occur in the signal. The results from applying low-pass filters at 50 kHz and 70 kHz are seemingly identical. Similarly, after applying a high-pass filter at 30 kHz, no significant changes are observed for a lower cutoff frequency.

Based on these observations, it was decided that the frequency range to be considered for this dataset would be from 20 kHz to 60 kHz.

4.2. Green's Function Estimation

Following the decision on the pre-processing scheme for this dataset, GF estimations were obtained for all adjacent sensor pairs within the concrete girder. Given that the sensors were placed in two different locations (top layer sensor group and web sensor group) with different conditions, they were analyzed separately. Additionally, the GF estimations were also investigated for other load cases applied in the experiment.

As explained in Section 3.1.1, the concrete girder in the validation dataset had two load application points. Although the loads applied at these points were always of equal magnitude, the effective load on the girder was not simply twice the magnitude of the applied load. This was due to the geometry of the girder and the points of application relative to it. These loading conditions are summarized in Table 4.1. For simplicity, the load cases will be referred to in the latter sections by the applied load per point of application, as shown in the left column of Table 4.1.

Table 4.1: Loads considered for further analysis of wave speeds.

Approx. applied load (kN)	Effective load (kN)
0	0
100	164.08
200	321.58
400	653.2
600	978.93
800	1306.66

4.2.1. Top layer sensor group

The initial idea tested for determining whether the estimated GF through ambient noise interferometry in concrete structures was accurate involved checking for normal moveout (NMO) by pairing a single sensor with subsequent sensors at increasing distances or offsets. The goal was to see whether the first arrival of a wave in the first sensor pair would be linearly related to the first arrival of the same wave in the subsequent sensor pairs. To test this, sensor 27 was chosen as the virtual source, and sensors 19-22 were chosen as the receivers.

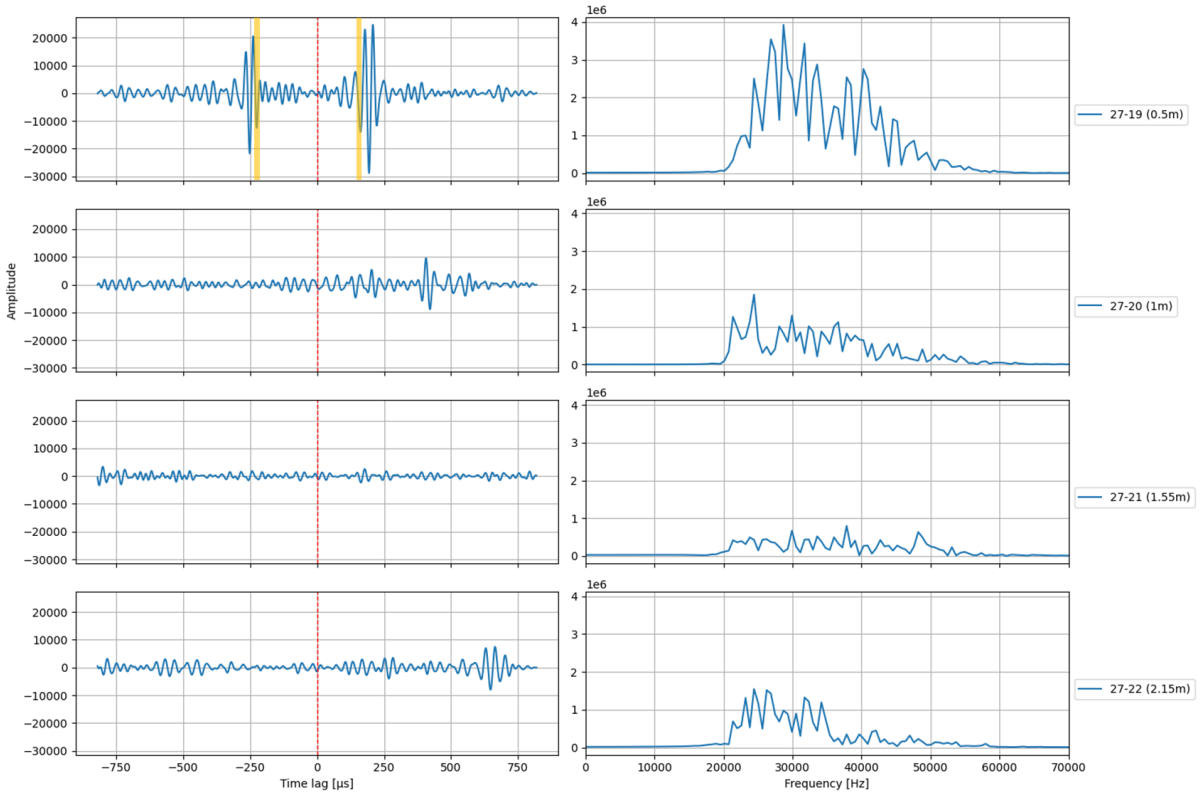

Figure 4.9: Stacked cross-correlations and corresponding frequency content of virtual source-receiver pairs with sensor 27 as the source, with yellow lines indicating the arrivals.

In Figure 4.9, the GF estimations for the four sensor pairs are plotted using the same y-scale to emphasize differences in signal amplitudes as sensor distances increase. Among these pairs, only pair 27-19 displays a coherent GF estimation, with peaks indicating apparent arrivals highlighted in yellow at approximately -200 and 150 microseconds. For the other three pairs, at distances beyond 0.5 meters, the estimations become monotonic, with only a few indistinct peaks. As a result, detecting arrivals and phase changes becomes challenging, rendering the results incoherent for identifying arrivals.

A possible reason for this is the inherent dispersivity of concrete, which pertains to the tendency of wave

velocities to vary with frequency. This dispersivity arises from the highly heterogeneous composition of concrete, which includes aggregates of different grain sizes, cement binders, and steel reinforcements. This heterogeneity causes different frequency components of a wave to travel at varying speeds. Another reason could be the matching and aligning algorithm used. Because the algorithm is based on the time a wave is detected by a sensor, a longer distance between a pair allows for a larger window of time for signals to arrive. This can lead to falsely matched signals, resulting in an unphysical CCF that corresponds to an incorrect GF estimation between the two sensors in question.

In addition, post-experiment analysis revealed that sensors 21 and 22 were not optimally coupled to the beam. Discrepancies were observed in the amplitude content of the signals recorded by these two sensors, indicating possible loose coupling.

Based on the observations from Figure 4.9, it can be suggested that estimating the GF between two stations spaced up to 0.5 meters using ambient noise may be more reliable. This is consistent with the fact that higher frequency components exhibit greater attenuation [19], particularly since the high frequencies of 20 - 60 kHz were investigated. However, further research is still necessary to confirm this. Moving forward, in the interest of minimizing the distance between sensor pairs to get more coherent GF estimations, only adjacent sensor pairs were analyzed. The distance between pairs varied from 0.5m to 0.8m.

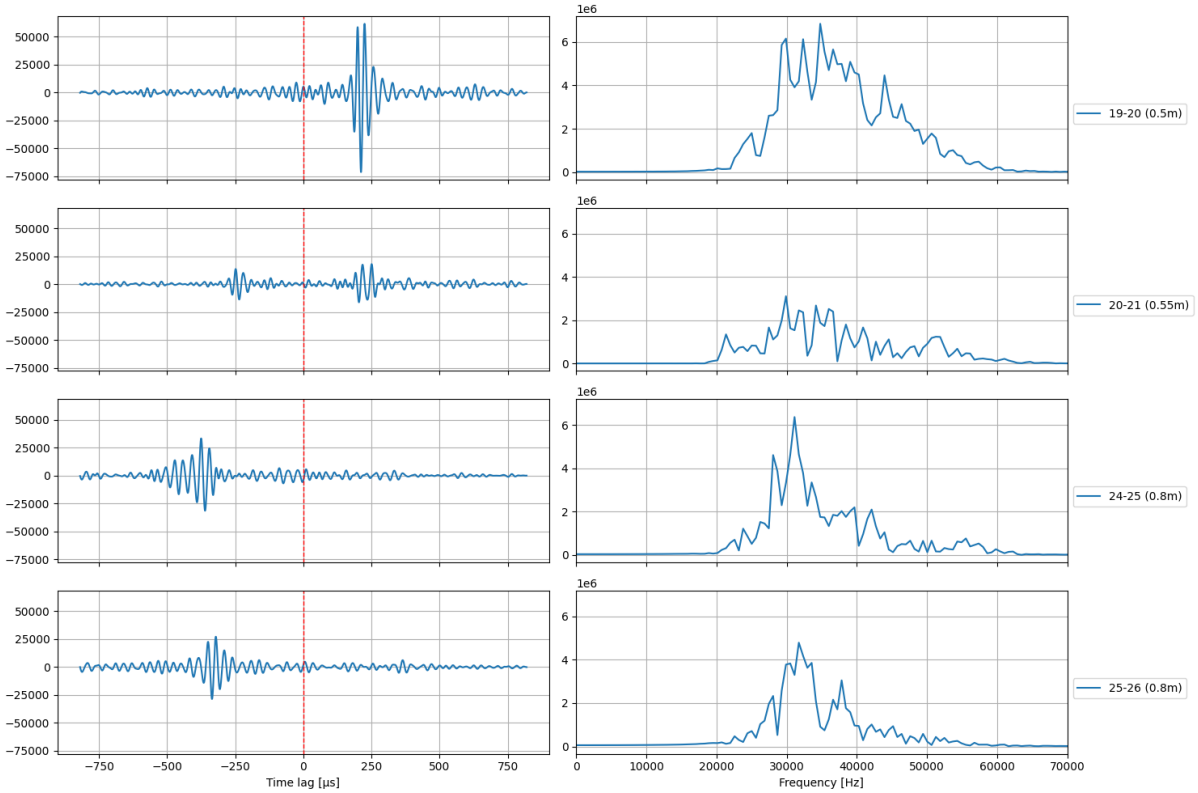

Figure 4.10: Stacked cross-correlations and corresponding frequency content of adjacent sensor pairs of the top layer sensor group.

In Figure 4.10, GF estimations for the adjacent pairs following pair 27-19 are presented, with all plots using the same y-scale, similar to the approach in Figure 4.9. When comparing the results of Figures 4.9 and 4.10, it is evident that the GF estimations in Figure 4.10 are much more coherent. Clear first arrivals are observed for the sensor pairs, with pairs 19-20, 24-25, and 25-26 showing distinct first arrivals on the expected side of the time lag = 0.

Pair 19-20, situated on the left side of the sensor layout, exhibits more energy coming from the right side, resulting in a dominant causal part. The same behavior is observed for pairs 24-25 and 25-26, which are situated on the right side of the sensor layout, leading to a more dominant acausal side.

For pair 20-21, a slightly symmetric stacked CCF is observed with minor amplitude differences between the causal and acausal parts. This result was unexpected because, like the other pairs, the energy was expected to be more dominant on one side. A possible explanation is the previously mentioned coupling issues with sensors 21 and 22, which may have caused unequal amplitude contributions from sensors 20 and 21. Given the additional uncertainty with sensor 21, it was excluded from further analysis.

To further investigate and verify the consistency and accuracy of the findings in Figure 4.10, adjacent sensor pairs 27-19, 19-20, and 25-26 were examined by checking their corresponding GF estimations for all the other load cases. These particular pairs were chosen as they were deemed representative of the rest of the group in terms of location and conditions.

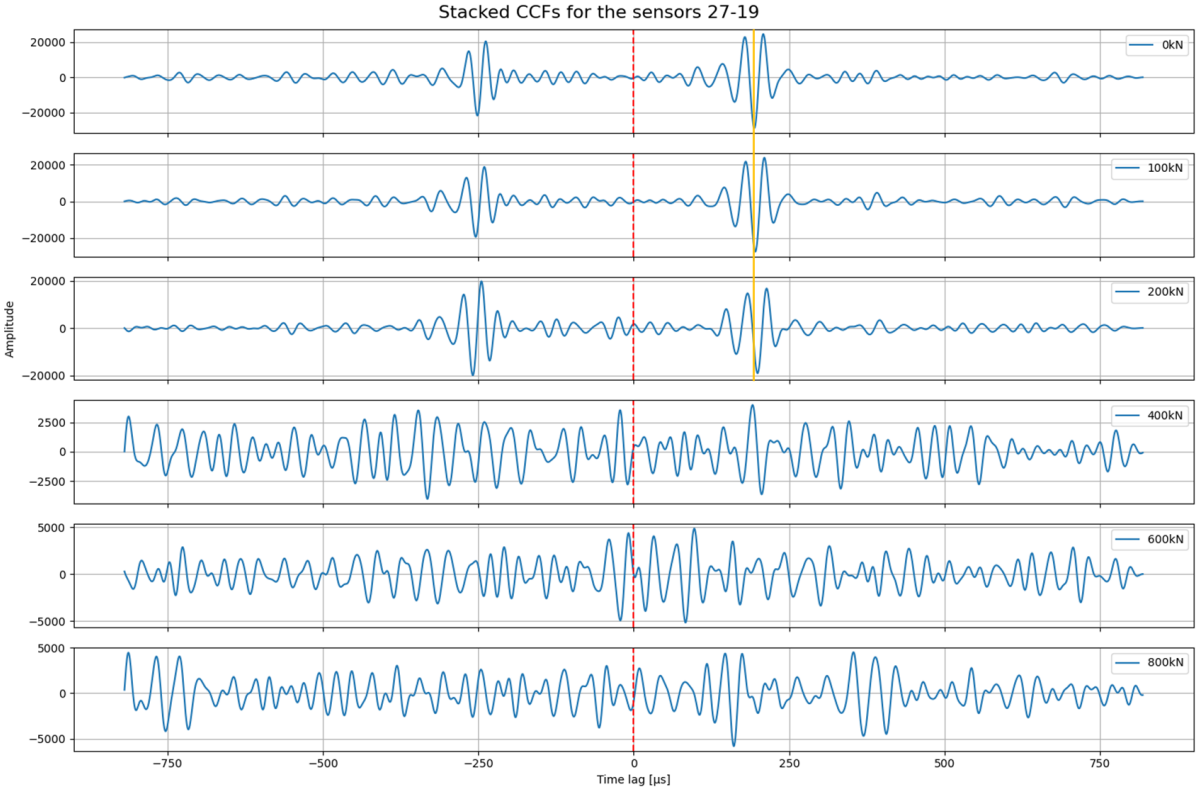

Figure 4.11: Causal and acausal parts of the Green's function estimation of pair 27-19 at different loadings.

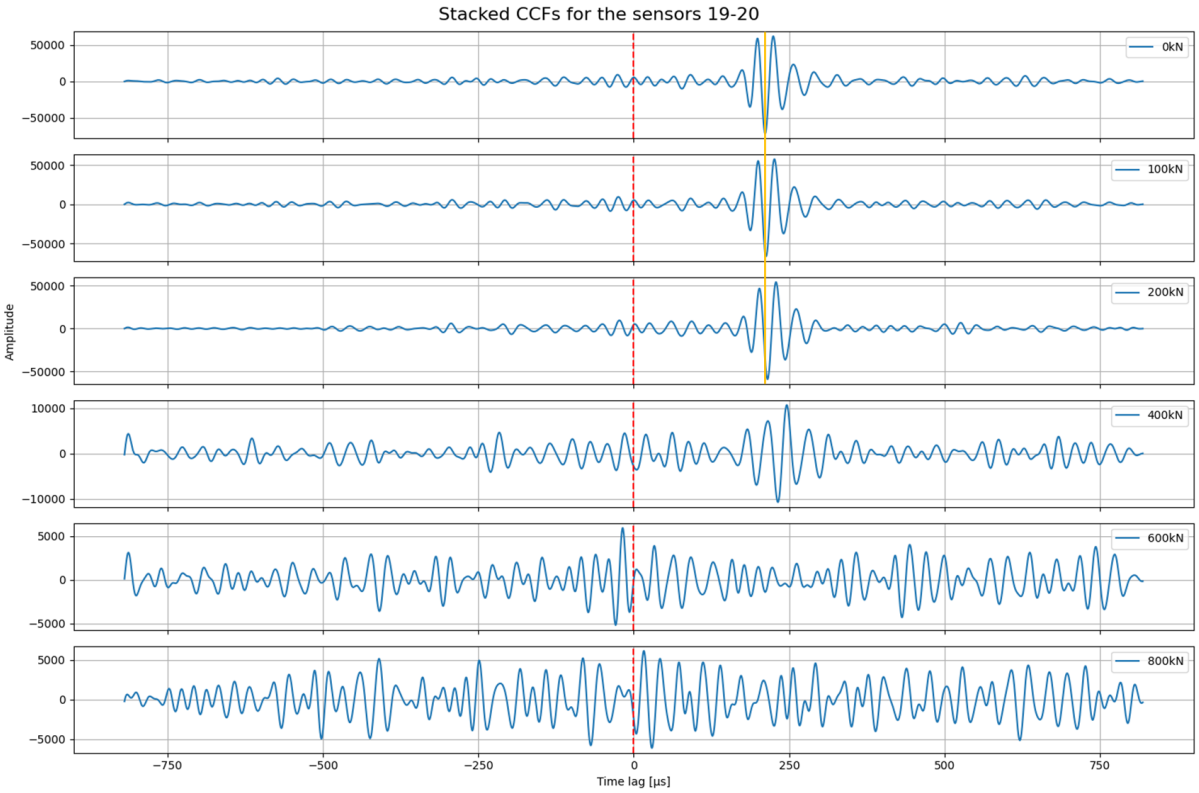

Figure 4.12: Causal and acausal parts of the Green's Function estimations of pair 19-20 at different loadings.

As seen in Figures 4.11 and 4.12, the GF estimations for pairs 27-19 and 19-20 remain consistent from load cases 0 to 200 kN. However, from 400 kN onwards, significant drops in the amplitude content occur, and noise increases in the estimations for both pairs. For the 27-19 pair, the amplitudes equalize, obscuring the wave arrivals that are clear at lower load cases. The same is observed for the 19-20 pair, although a clear peak of amplitudes amidst the noise can still be discerned until 600 kN, when it disappears. These occurrences can be attributed to the flexural cracks that develop near the supports at 400 kN (see Figure 4.13). The introduction of cracks causes sound waves to lose energy as they pass through and scatter. This energy loss and scattering are reflected in the corresponding amplitude contents of the GF estimations.

Another observation in both figures is the slight stretching in the GF estimations from 0 kN to 200 kN, marked by yellow vertical lines. The peak gradually shifts to the right as the load increases, which can be attributed to acoustoelastic effects (Section 2.5).

Since the sensor pairs were located on or near the top of the intermediate support, they were in the tension zone due to the negative moment induced in this area by the loading. Consequently, sound waves traveled more slowly, which resulted in the delay of the occurrence of the peak.

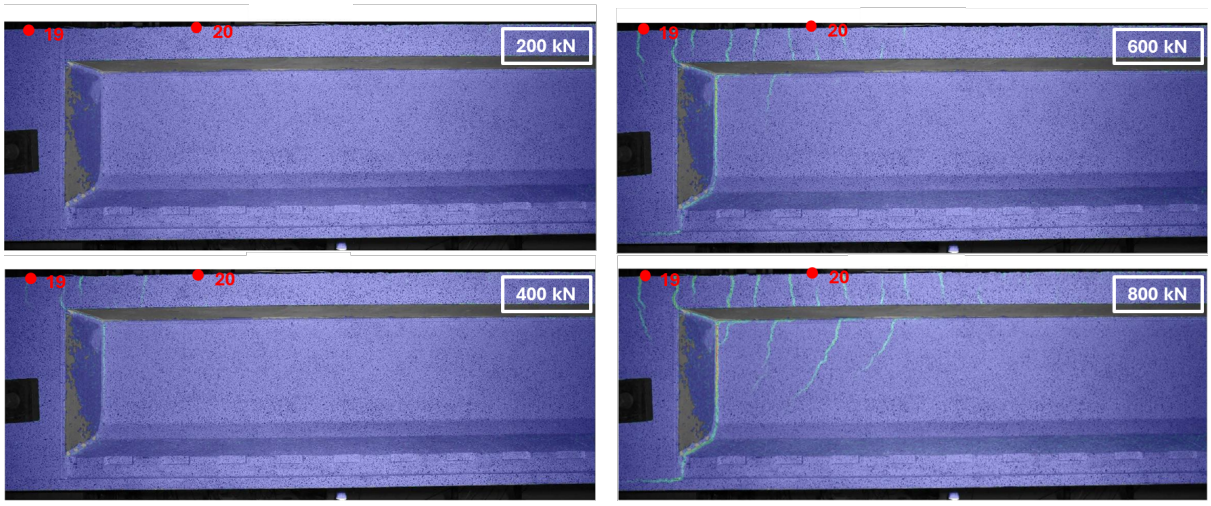

Figure 4.13: Beam cracking found near the support.

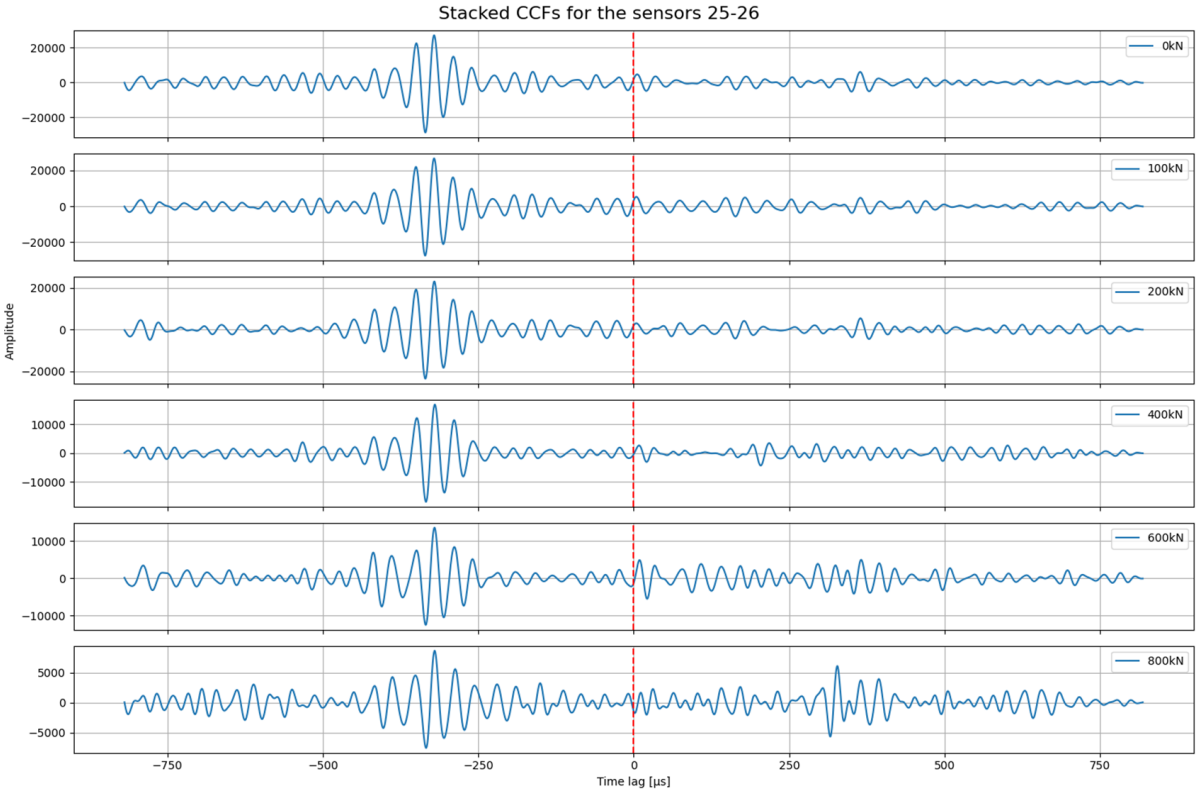

Figure 4.14: Causal and acausal parts of the Green's Function estimation of pair 25-26 at different loadings.

Further substantiating the findings in Figures 4.12 and 4.13, Figure 4.14 shows that the GF estimation for pair 25-26 remains coherent. Despite a slight equalization of amplitudes at 600 and 800 kN, the peaks remain visible.

This response was expected, as these sensors were located near the midspan of the beam, where tensile stress was lowest, and no flexural cracking occurred. Additionally, the first cracking in a vertically loaded beam typically occurs in the tensile zone at the bottom, opposite to where the sensors were positioned. However, it should be noted that there was a decrease in amplitude as the load increased. This could be due to the lower energy waves detected by the sensors, caused by cracks forming in other

parts of the girder, along with the acoustoelastic effects of concrete.

4.2.2. Web sensor group

For the hits recorded by the sensors on the web, the same pre-processing scheme used for the sensors on the top layer was applied. This decision was based on an assessment of the optimal normalization procedure, which showed that only spectral whitening had a significantly positive effect, whereas the temporal normalization methods were insignificant. This is illustrated in Figure 4.15.

Note that the differences in amplitudes in Figure 4.15 are not significant, as different normalization methods can affect the magnitude of the resulting amplitudes.

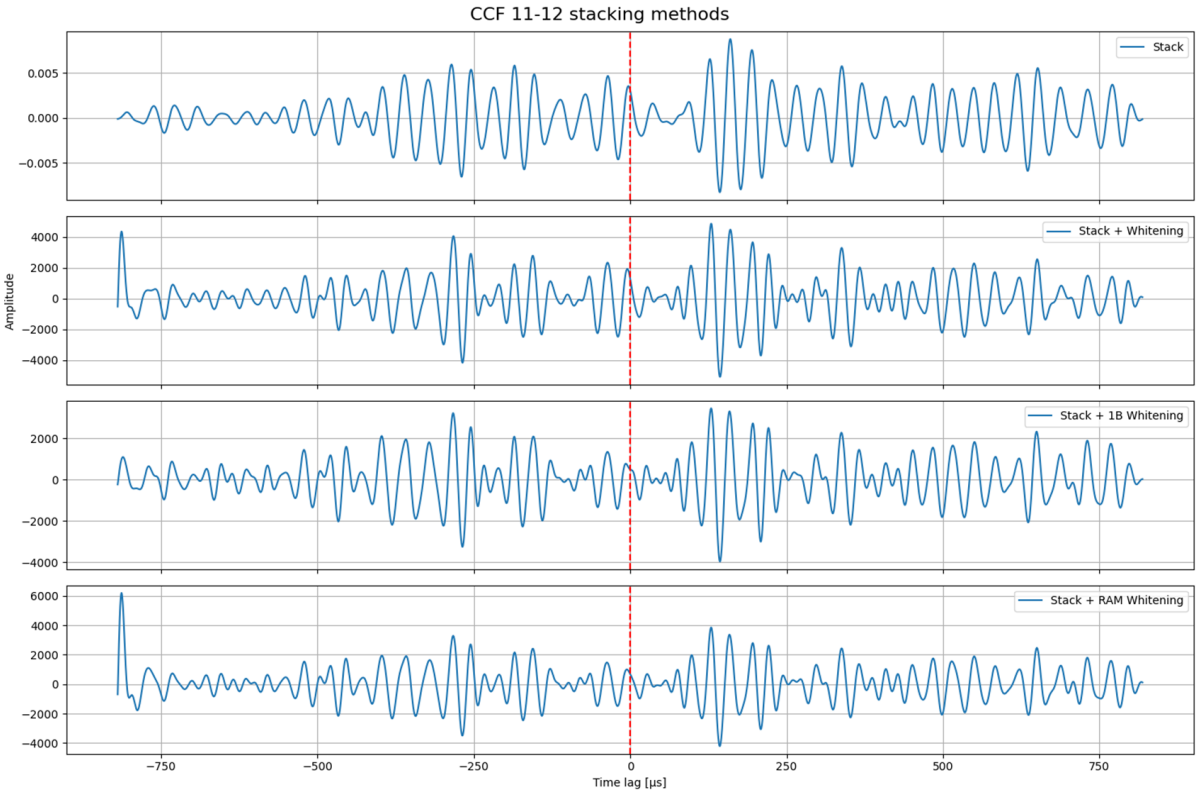

Figure 4.15: Different pre-processing methods and the corresponding stacked CCFs for pair 11-12.

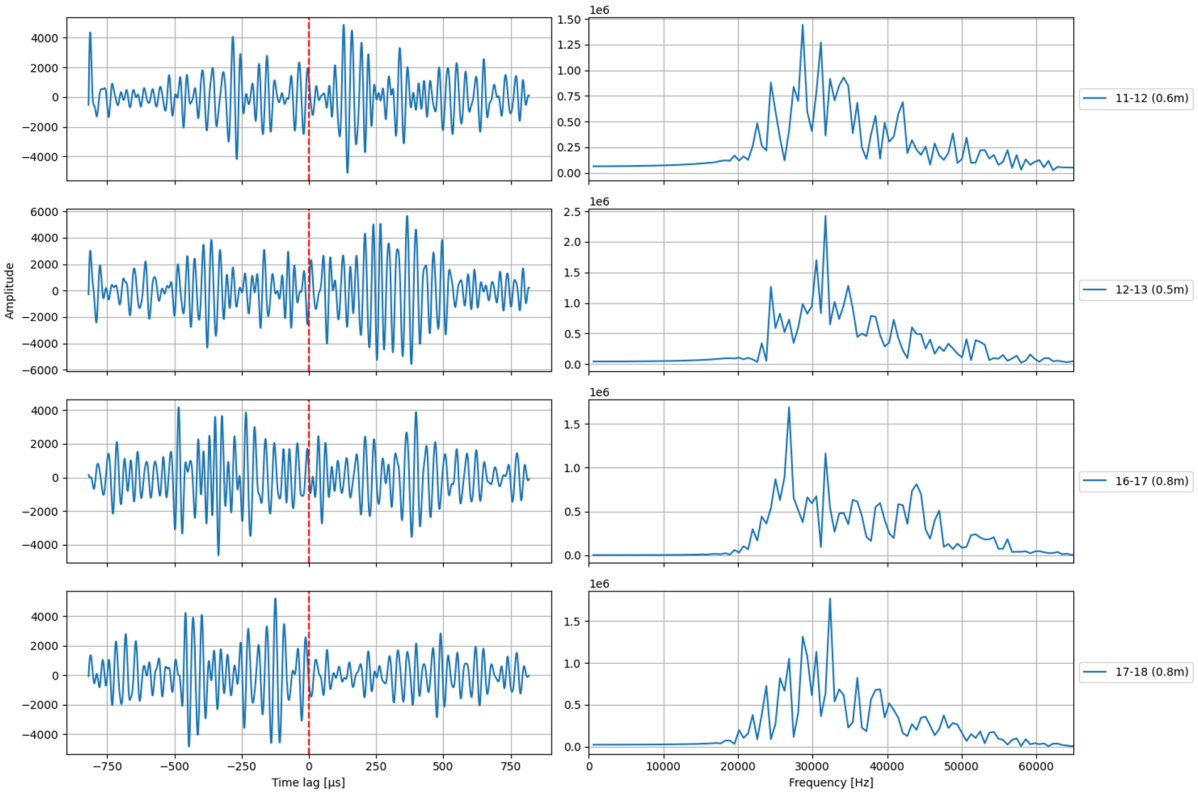

Figure 4.16: Green's function estimations and corresponding frequency content of adjacent sensor pairs of the web sensor group.

In Figure 4.16, it is observed that even though the frequency content of the signals from the web sensor group is almost identical to that of the top layer sensor group and the same pre-processing methods are used, the resulting GF estimations are significantly less coherent. Although some peaks are clear, there is too much noise in all of the estimations to reliably detect arrival times and distinct features.

This difference can be attributed to the greater distance of the web sensor group from the source compared to the top layer sensor group. Given that high frequencies were considered, attenuation was more significant, especially in a highly heterogeneous medium like concrete. Additionally, the source was located at the slab, which has a much wider width than the web (see Figure 3.2), requiring sound waves to travel specific paths to reach the web sensors. This is corroborated by the fact that the web sensors recorded fewer hits than the top sensors (see Figures 4.2 and 4.3). Furthermore, the noisy GF estimations can be attributed to the multiple reflections that some waves undergo before reaching the web sensors, causing them to arrive at similar times as other waves, possible creating unphysical signals which further obscure the arrivals.

To further investigate the estimated GFs, the behaviors of pairs 11-12, 12-13, 16-17, and 17-18 under different loads were examined. The results are illustrated in Figures 4.17 and 4.18.

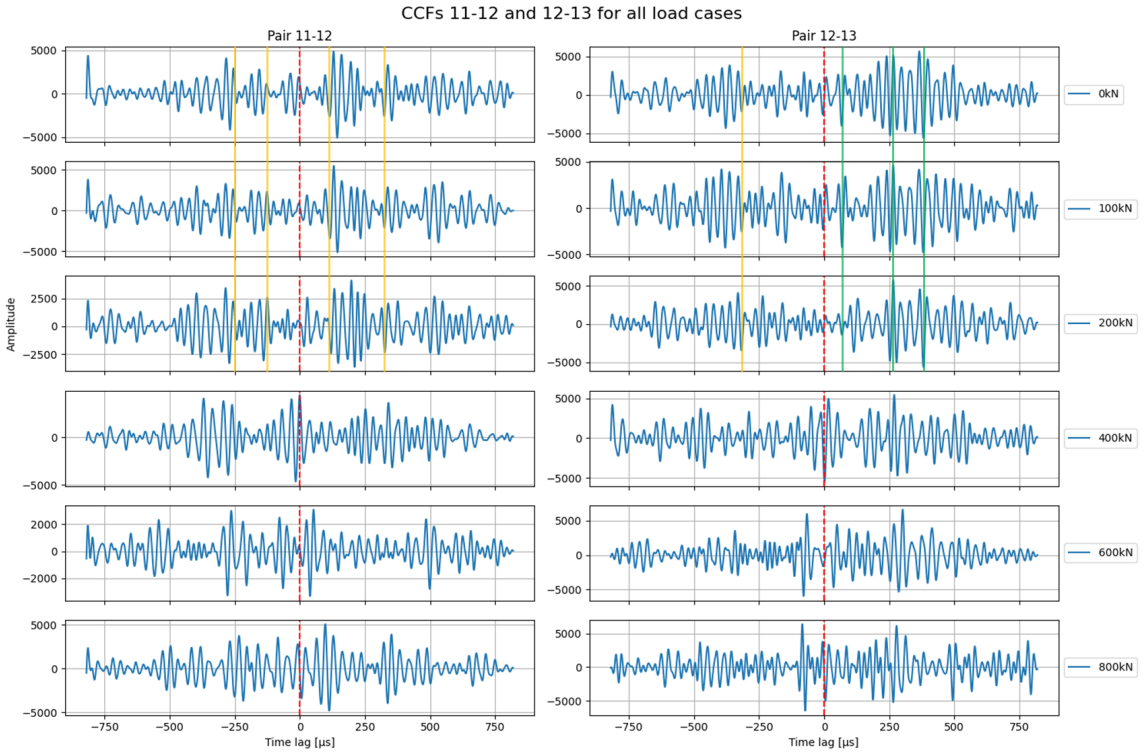

Figure 4.17: Causal and acausal parts of the Green's Function estimations of pairs 11-12 and 12-13 at different loadings, with yellow lines and green lines indicating the arrivals and peaks, respectively.

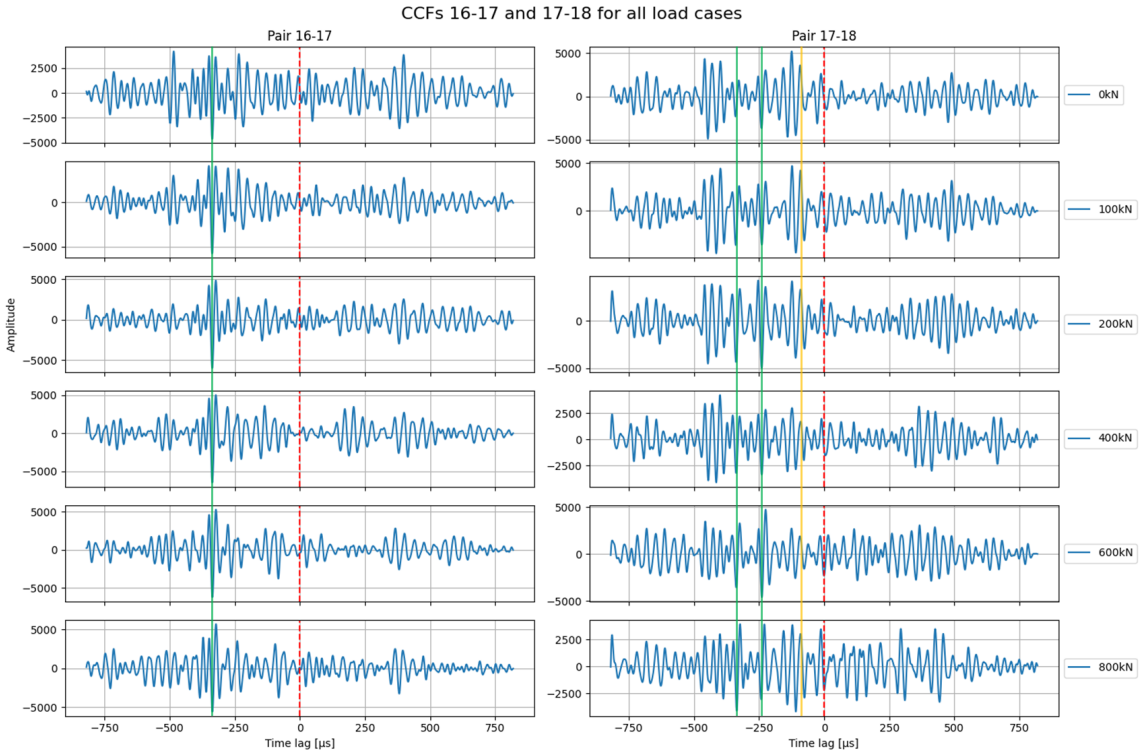

Figure 4.18: Causal and acausal parts of the Green's Function estimations of pairs 16-17 and 17-18 at different loadings, with yellow lines and green lines indicating the arrivals and peaks, respectively.

The GF estimations in Figures 4.17 and 4.18 again show very noisy signals, with possible multiple wave arrivals obscuring each other. Despite this, certain consistent patterns are observable in both figures.

In Figure 4.17, although the wave arrivals for pairs 11-12 and 12-13 are not very clear, consistent jumps and peaks in amplitude are observable across the 0-200 kN load cases, as indicated by the yellow vertical lines. For pair 11-12, amplitude peaks are consistently observed at approximately -125, -250, 100, and 300 microseconds. However, starting from the 400 kN load case, these peaks become less prominent, and more arrivals closer to the 0 time lag are visible. Similarly, for pair 12-13, despite showing significantly more noise, consistent amplitude peaks are observed around -300, 100, 260, and 350 microseconds.

These observations align with the changing girder conditions as the load increases. Similar to what was observed for the top layer sensor group, the GF estimations of sensor pairs near the support started to exhibit noisier signals beginning at the 400 kN load case, likely due to the cracks that developed on the girder at this load, as shown in Figure 4.13.

For sensor pairs 16-17 and 17-18 shown in Figure 4.18, consistent amplitude peaks are observed throughout the load cases. Unlike the other sensors, these peaks remain stable. These sensors were positioned on the right portion of the sensor array, similar to pair 25-26 on the top layer, where more energy was observed in the acausal part. Additionally, since sensors 16-18 were also at the midspan of the beam, the cracking that occurred at the supports at 400 kN had minimal effect on the GF estimations between these sensors, resulting in a consistently clear trend.

For pair 16-17, the most prominent peak in amplitude is observed around -300 microseconds and is noticeable across all load cases. Similarly, for pair 17-18, prominent peaks are consistently observed around -100, -250, and -300 microseconds across all load cases.

4.2.3. Decomposing Green's function estimations of the web sensor group using EEMD

To further enhance signal denoising, frequency isolation, and emphasize wave arrivals in the time domain, an empirical method known as Ensemble Empirical Mode Decomposition (EEMD) (see Section 2.8) was employed on the normalized and band-passed GF estimations for the web sensor group. Despite EEMD being a frequency-isolating method in itself, the band-passing prior to its application was deemed more advantageous as it allowed for the removal of unwanted frequency components, potentially leading to more meaningful IMFs related to wave arrivals.

This further step was not considered necessary for the top layer sensor group since the coherence of their GF estimations was already deemed sufficient.

As EEMD is purely empirical, the choice of which Intrinsic Mode Function (IMF) to analyze was subjective.

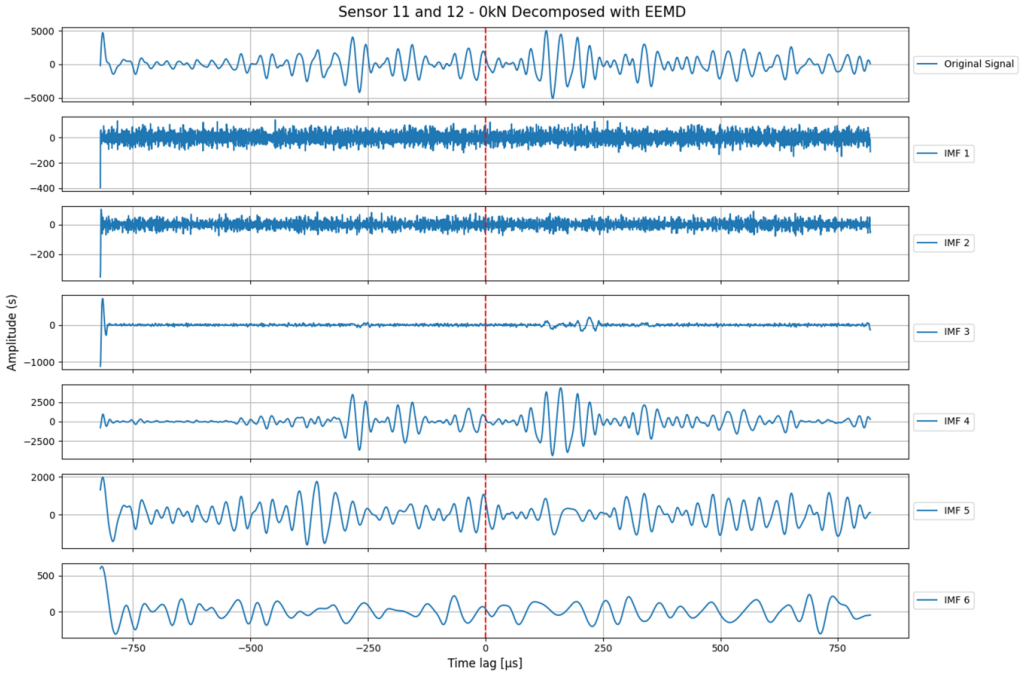

Figure 4.19: Decomposing the Green's function estimation of sensor pair 11-12 with EEMD.

In Figure 4.19, the results of EEMD are presented. The original GF estimation for sensor pair 11-12 is first augmented with white noise and then subjected to the sifting process described in Section 2.8. This process, depicted in the figure, isolates signal features based on their frequency in the time domain, with longer periods becoming more prominent as the IMF number increases.

Based on the outcomes observed, IMF4 was selected for further investigation across all other sensor pairs and corresponding load cases for the web sensors.

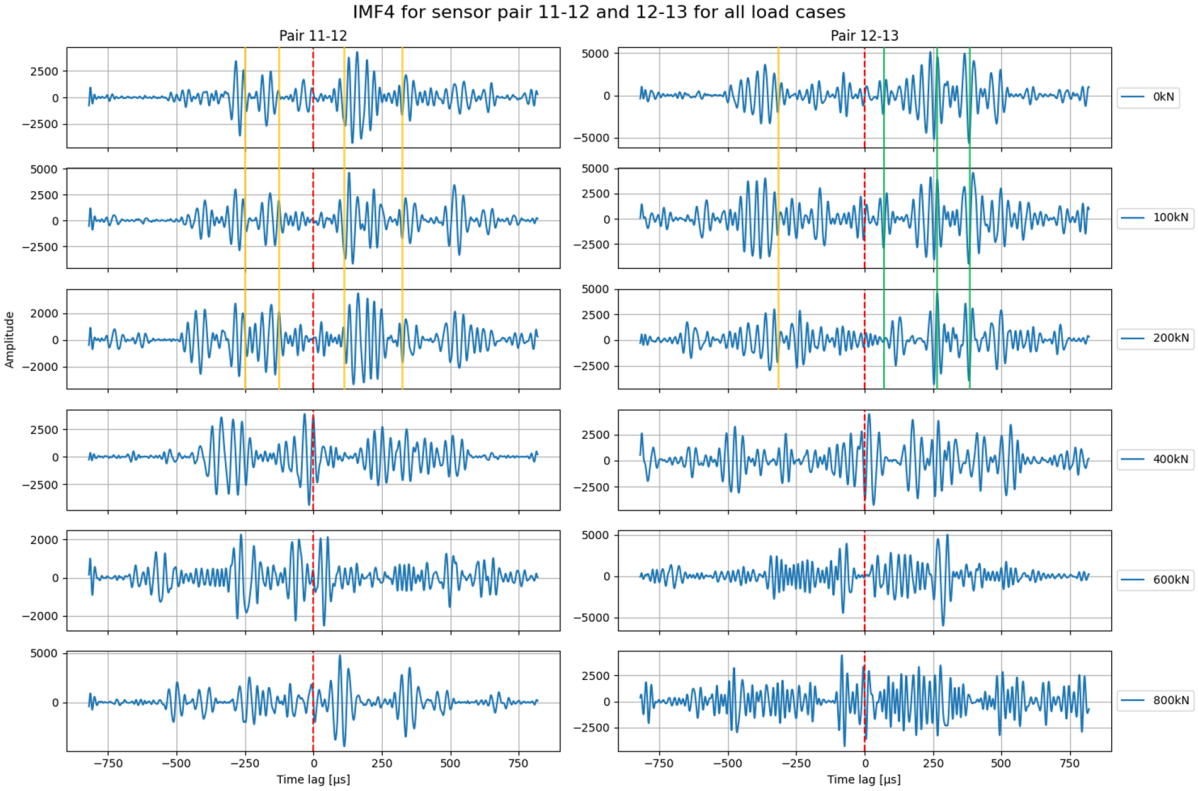

Figure 4.20: IMF4 for pairs 11-12 and 12-13 for all load cases, with yellow lines and green lines indicating the arrivals and peaks, respectively.

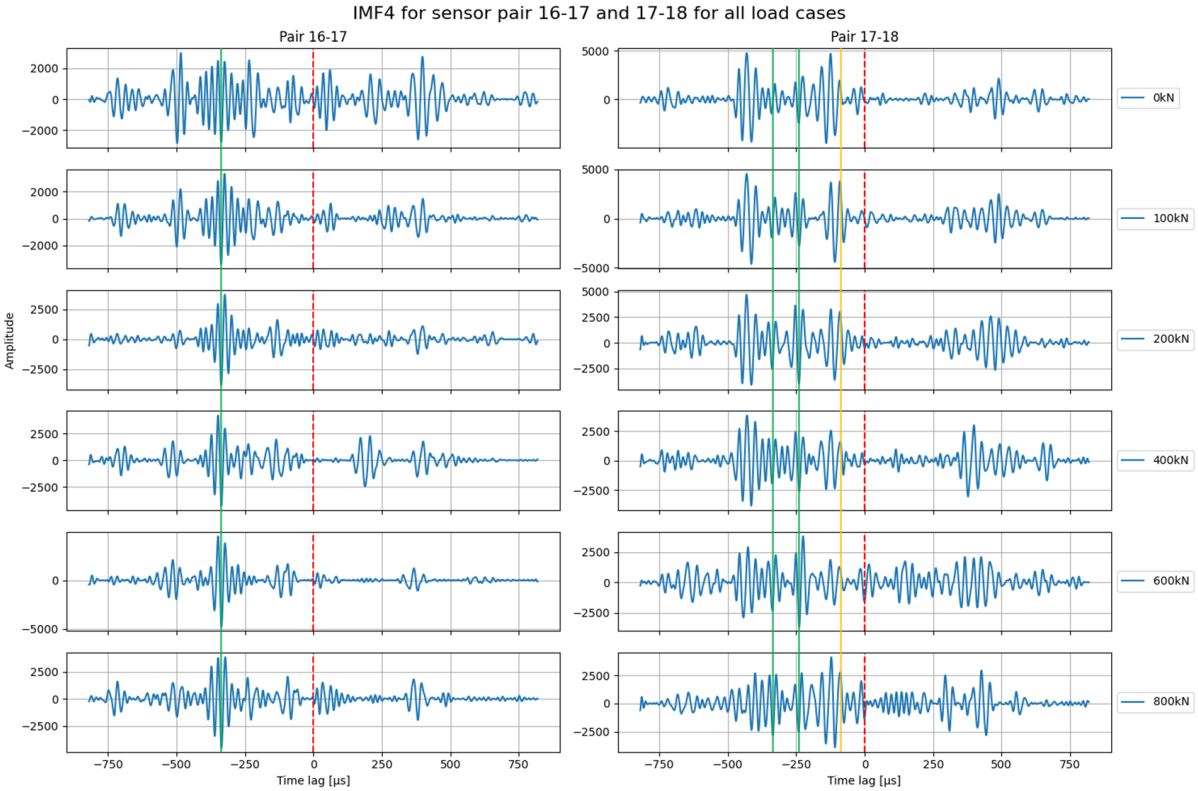

Figure 4.21: IMF4 for pairs 16-17 and 17-18 for all load cases, with yellow lines and green lines indicating the arrivals and peaks, respectively.

From Figures 4.20 and 4.21, it is evident that EEMD effectively isolates the previously observed peaks and wave arrivals depicted in Figures 4.17 and 4.18 amidst the noise.

In Figure 4.20, significant energy appears on both the causal and acausal sides of the IMF4s, as indicated by the amplitude content in the GF estimations. The consistent peaks identified in Figure 4.17 are more pronounced with this decomposition method, emphasizing both the clarity of the trends and loss of coherence observed from the 400 kN load case onwards.

Figure 4.21 demonstrates that the decomposition method effectively enhances the energy distribution from the sources, particularly emphasizing the dominance of the acausal parts for sensor pairs 16-17 and 17-18. Moreover, the consistent peaks identified in Figure 4.18 are more prominently displayed following decomposition, notably the peak around -300 microseconds for pair 16-17 and peaks around -100 and -300 microseconds for pair 17-18.

4.2.4. Comparative analysis of the sensor groups

The following observations were made:

- Although the estimated Green's functions from the top layer sensor group were much more coherent compared to those of the web sensor group, consistencies were found for all sensor pairs before and after cracking. This demonstrates their capability for local damage detection or anomaly detection.
- Despite the differences in source locations, wave travel distances, and the expected energies and trajectories of waves, the amplitude content featured by the GF estimations of both sensor groups did not vary greatly.
- The Green's function estimations from the web sensors contained more noise due to the significantly longer path of the sound waves and the need to cross an interface.
- The dominance of either the causal or acausal parts, depending on the source distribution, was consistent for both sensor groups.

4.3. First Arrival Detection using AIC Algorithm

This section aimed to employ an automatic first arrival detection algorithm based on the Akaike Information Criterion (AIC) (see Section 2.7) to determine the wave speeds of the first-arriving waves from the GF estimations of various adjacent sensor pairs. The algorithm detects the time of first arrival, which was checked on either side of the 0 time lag, depending on the position of the sensor pair relative to the chain path, influencing whether the causal or acausal part was examined.

Specifically, this analysis focused on sensor pairs 27-19, 19-20, and 25-26, as they were deemed representative of the entire top layer sensor group dataset. For pairs 27-19 and 19-20, the causal part was investigated, while for pair 25-26, the acausal part was examined. Due to the significant noise in the GF estimations of the web sensor group, which caused challenges in identifying wave arrivals, these results are not discussed further in this section. Additional findings regarding arrivals and wave speeds detected for the top layer sensor group can be found in Appendix A.

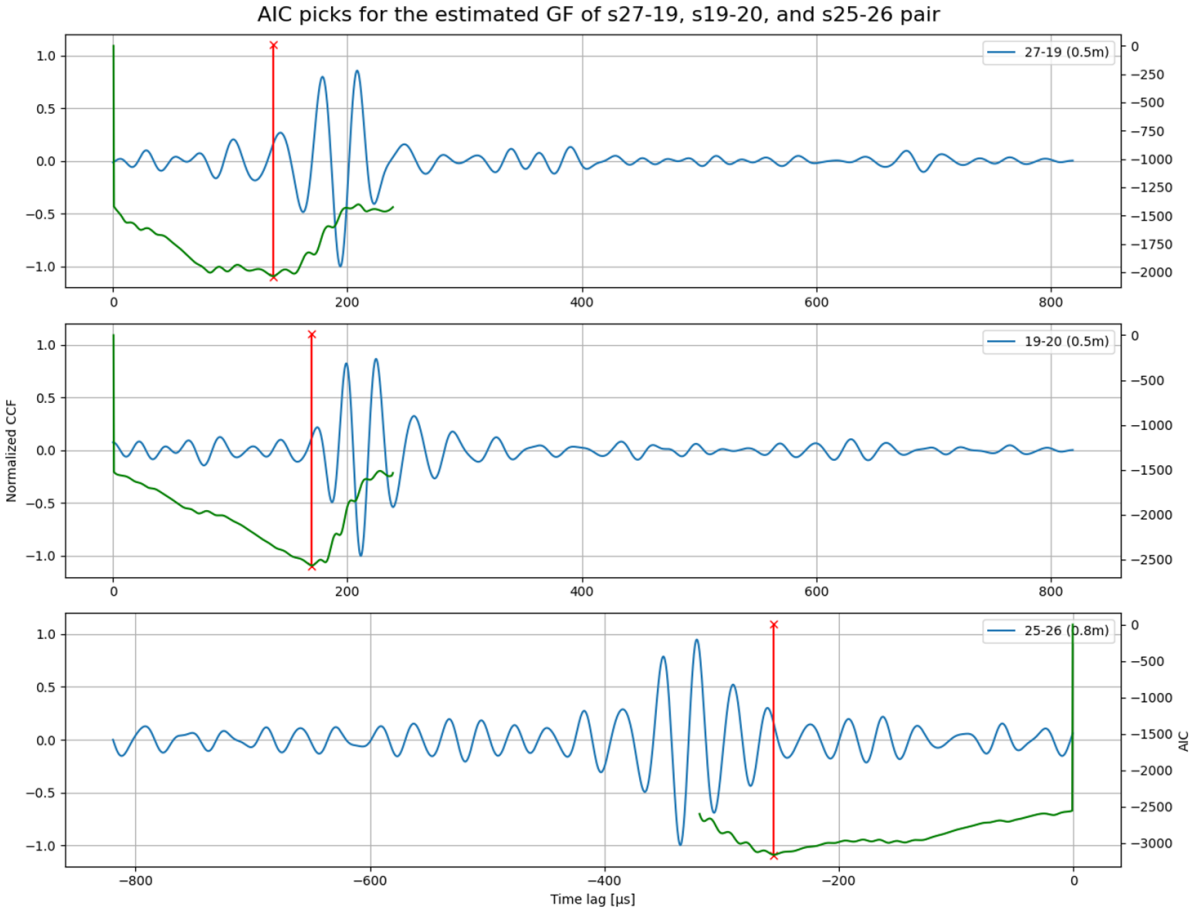

Figure 4.22: Arrival times detected with AIC picker for estimated Green's functions of 27-19, 19-20, 25-26 at $0kN$

As observed in Figure 4.22, the AIC automatic picking algorithm identifies the lowest AIC value within a given window (the length of the green line) for each sensor pair and selects that as the first arrival. The following table summarizes the arrival times and corresponding wave speeds depicted in Figure 4.22.

Table 4.2: Wave speeds from the AIC-picked first arrivals for each sensor pair of the 0 kN load case.

Sensor pair	**Distance (m)**	**Arrival time (μs)**	**Wave speed (m/s)**
27-19	0.5	136.4	3665.7
19-20	0.5	169.6	2948.1
25-26	0.8	255.6	3129.9

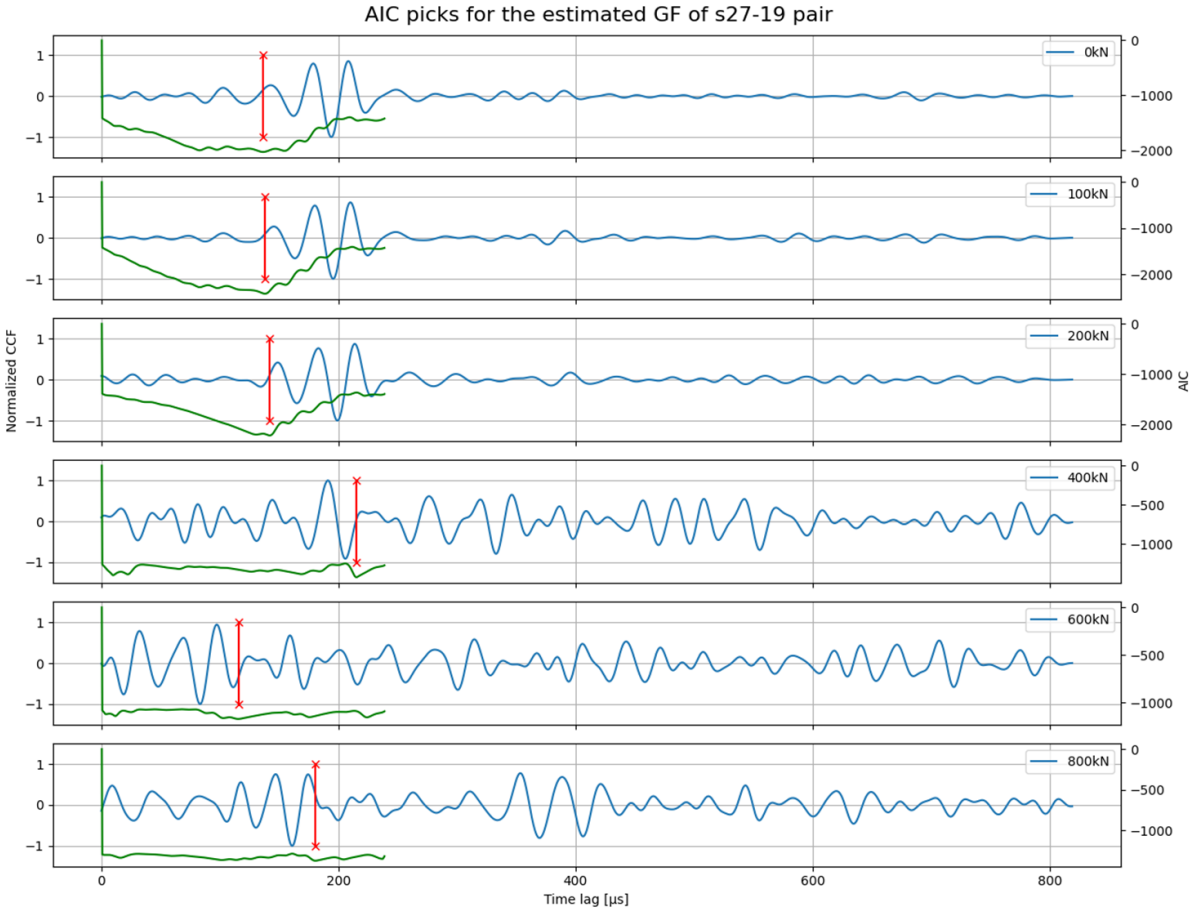

Figure 4.23: AIC-picked arrival times for pair 27-19 at different load cases.

Table 4.3: Arrival times and wave speeds for pair 27-19 at different load cases.

$d = 0.5m$	**Arrival time** (μs)	**Filtered Speed** (m/s)
0kN	136.4	3665.7
100kN	138.0	3623.2
200kN	142.0	3521.1
400kN	215.2	2323.4
600kN	115.6	4325.3
800kN	180.4	2771.6

For sensor pair 27-19, as seen in Figure 4.23, a consistent arrival time is found across the 0-200 kN load cases, corresponding to consistent wave speeds. As mentioned in Section 4.2.1, cracks developed in the beam near the supports at 400 kN, leading to the drastic change in arrival time. This arrival time is unreliable due to the increased noise in the signal and the equalization of amplitudes, possibly caused by the presence of cracks. This is a non-ideal scenario for using the arrival time algorithm, as explained in Section 2.7.

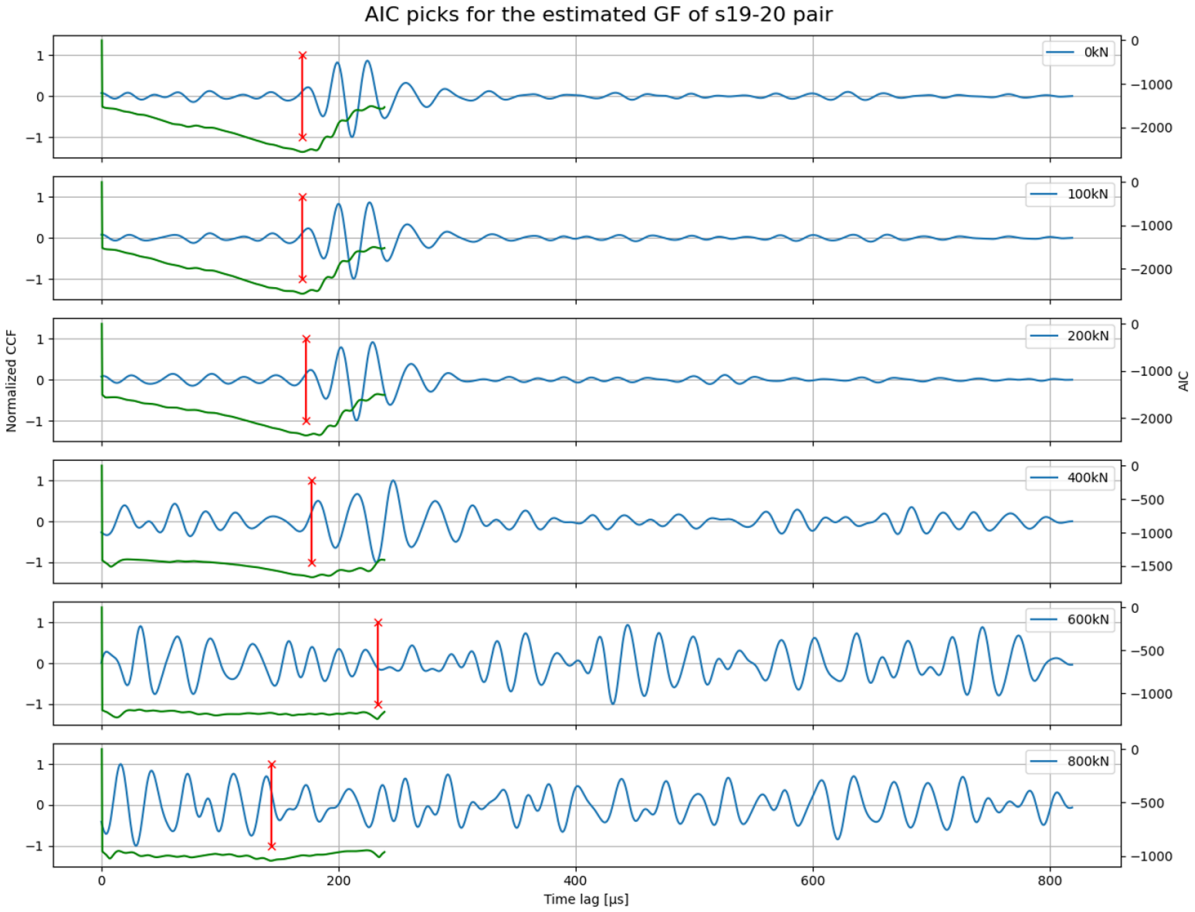

Figure 4.24: AIC-picked arrival times for pair 19-20 at different load cases.

Table 4.4: Arrival times and wave speeds for pair 19-20 at different load cases.

$d = 0.5m$	**Arrival time** (μs)	**Filtered Speed** (m/s)
0kN	169.6	2948.1
100kN	169.6	2948.1
200kN	172.8	2893.5
400kN	177.6	2815.3
600kN	232.8	2147.8
800kN	143.2	3491.6

In Figure 4.24, similar to pair 27-19, pair 19-20 exhibits consistent arrival times and corresponding wave speeds from 0 to 200 kN load cases. However, at 400 kN, the coherence in detected arrival times become unreliable. Despite both pairs being equidistant, notable differences in arrival times and wave speeds were observed.

A noticeable observation for both sensor pairs 27-19 and 19-20 is the slight increase in arrival times, leading to a decrease in wave speeds as the loading increases just before cracking occurs. This aligns with the earlier observation in Section 4.2.1, where a slight stretching was noted in the GF estimations from 0 to 200 kN for both sensor pairs.

Additionally, between the two pairs, the increase in arrival times for the 0 to 200 kN load cases is more apparent for pair 27-19, as it was positioned closer to the support and thus experienced higher tensile stresses. This observation is consistent with the concept of acoustoelasticity (see Section 2.5).

For pair 27-19, before crack development, wave speeds consistently ranged between 3500-3600 m/s, whereas for pair 19-20, the speed was approximately 2900 m/s. This discrepancy can be attributed to the inherent spatial variation of wave velocities in concrete, influenced by factors such as density, porosity, and aggregate size distribution. The heterogeneous nature of concrete, coupled with construction practices, curing conditions, and reinforcement levels, contributes significantly to these variations.

Despite these uncontrollable differences, Both sensor pairs fall within the expected range of sound wave speeds in concrete, estimated roughly in Section 2.4.1. For concrete of grade C40/50, typical wave speeds range from 2234 m/s (Rayleigh wave) to 4000 m/s (primary wave), considering standard values of modulus of elasticity and Poisson's ratio.

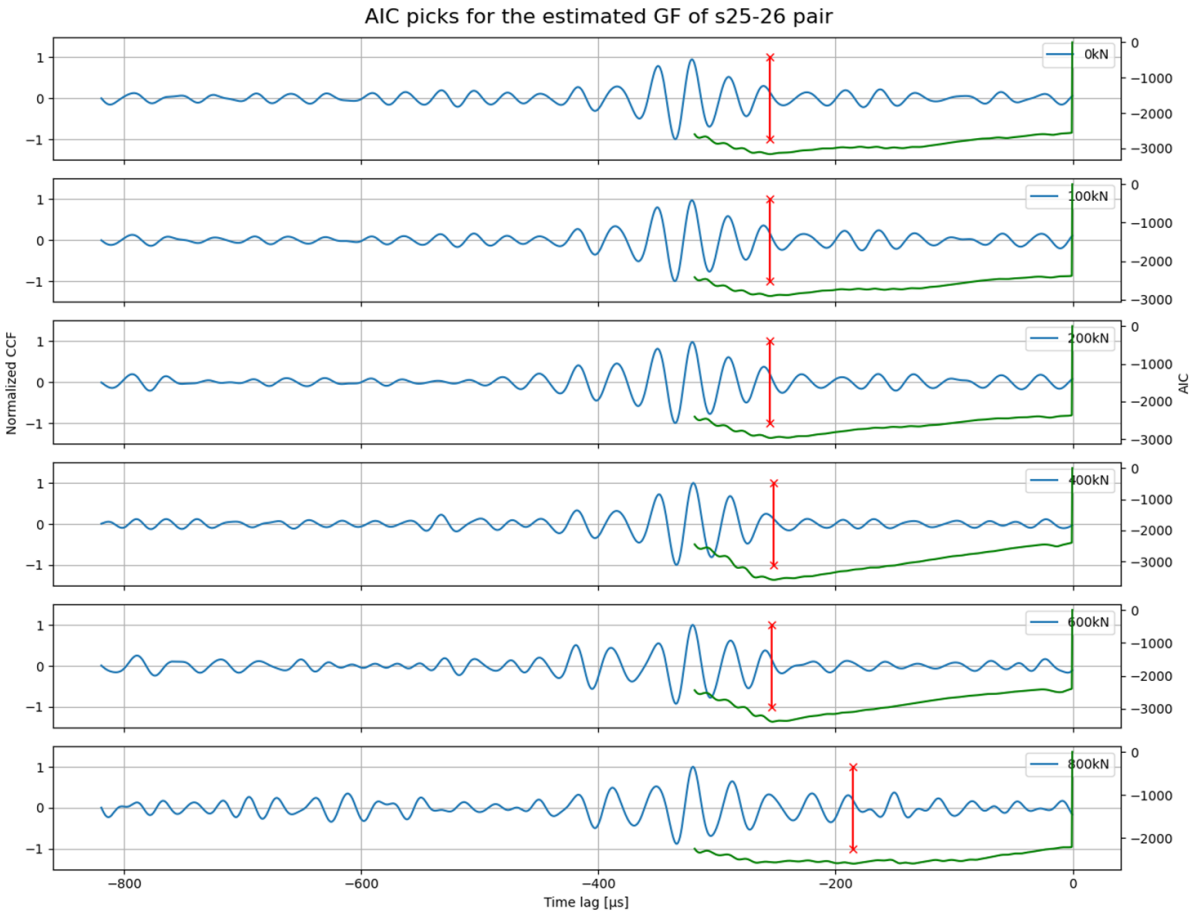

Figure 4.25: AIC-picked arrival times for pair 25-26 at different load cases.

Table 4.5: Arrival times and wave speeds for pair 25-26 at different load cases.

$d = 0.8m$	**Arrival time (μs)**	**Filtered Speed (m/s)**
0kN	255.6	3129.9
100kN	255.6	3129.9
200kN	255.2	3134.8
400kN	252.4	3169.6
600kN	253.6	3154.6
800kN	185.2	4319.7

For sensor pair 25-26, as shown in Figure 4.25, the arrival times and wave speeds remain consistent across almost all load cases, except for the 800 kN case. Despite the earlier arrival detected by the algorithm at 800 kN, it cannot be concluded that a significant change occurs in the medium between

the two sensors. The trend of the amplitudes remains consistent throughout the load cases, with the GF estimation amplitude decreasing at 800 kN and becoming more obscured by random background noise, as explained in Section 4.2.1. The flatness of the green line (AIC values) in the 800 kN load case indicates that the arrival time could be picked up at any of those points. Therefore, it is still possible that the arrival time remains unchanged but is obscured by the added noise or the equalization of the amplitudes. A clearer view of the peak prominence at around -300 microseconds is shown in Figure 4.14.

As discussed in Section 4.2.1, this stability was expected given the location of sensor pair 25-26 with respect to the beam, explaining the minimal observed change.

From the analysis of the three sensor pairs, it can be inferred that detecting arrivals and wave speeds derived from the GF estimates between sensor pairs on concrete structures may be an effective method for monitoring damages and strain changes that develop in the medium.

5

Ambient Traffic Noise Results and Discussions

This chapter addresses an ambient traffic noise dataset from the Maastunnel in Rotterdam. Due to the limited available literature on this type of noise dataset, there is little knowledge of what to expect, and no established guidelines exist for data management and pre-processing.

The data processing followed these steps: first, the data characteristics were examined; next, the optimal pre-processing method was selected; and finally, the GF estimations were analyzed.

Since the Maastunnel is an old structure and lacks a reference baseline, interpreting the results from the dataset was challenging. However, because cracks were observed between sensors 3, 4, and 5 (see Figure 3.6), the GF estimations involving these sensors were used as the primary point of comparison. These estimations were then compared with those from the uncracked areas between sensors 1, 2, and 3.

5.1. Dataset Characteristics and Optimal Pre-processing Scheme

The first aspect investigated with this dataset was the distribution of hits. As mentioned in Section 3.2.1, the data for the Maastunnel consisted of two recordings, both taken on the 16th of December 2022. The first recording began at 11:56:55 and lasted around 52 minutes, while the second recording started at 12:53:23 and lasted approximately 47 minutes. To represent each recording, two separate histograms are presented. For analyzing the distribution of hits across all five sensors, the bin sizes for the histograms were set to 100 seconds each.

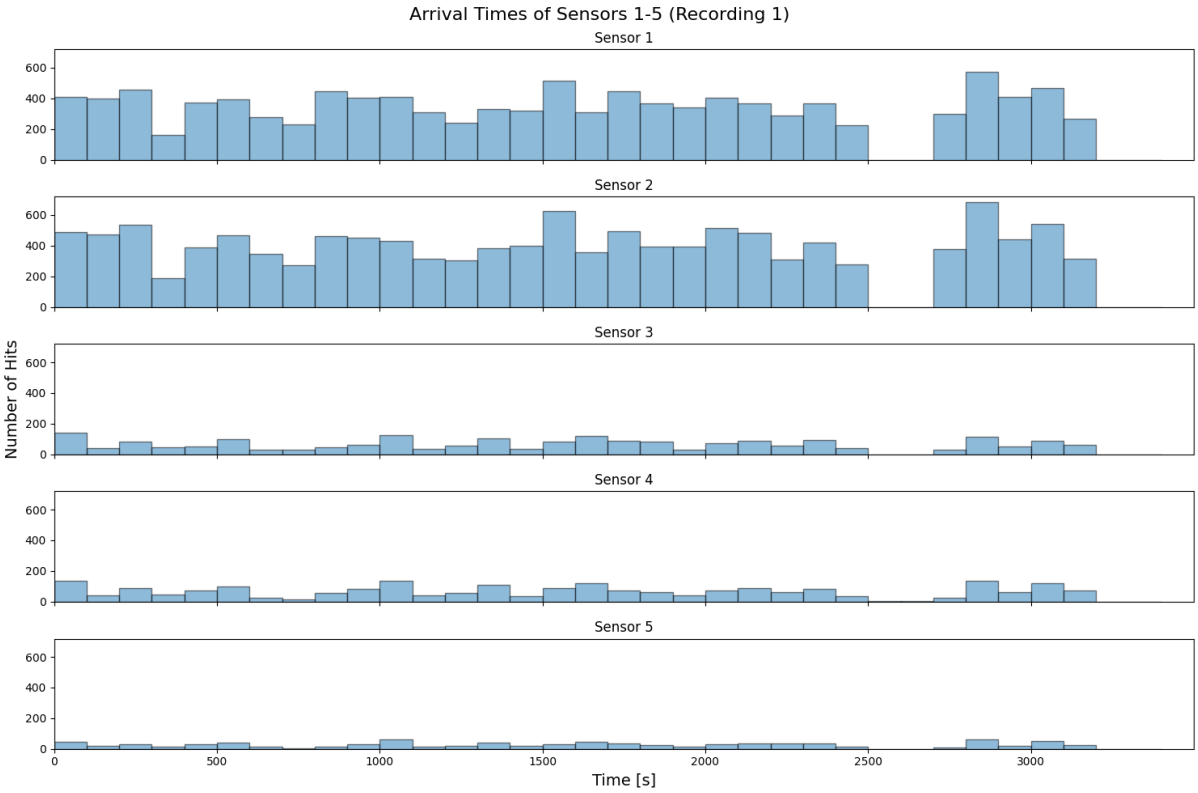

Figure 5.1: Hits for sensors 1-5 in the Maastunnel (Recording 1).

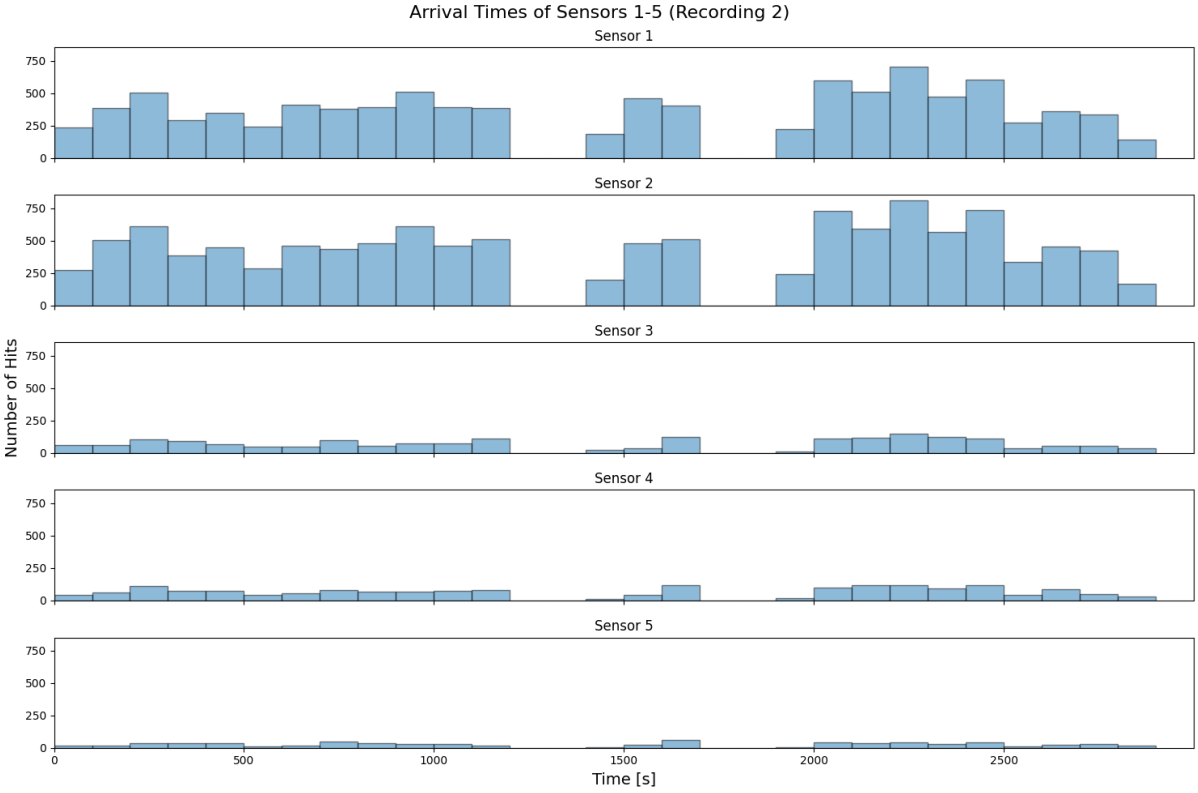

Figure 5.2: Hits for sensors 1-5 in the Maastunnel (Recording 2).

In Figures 5.1 and 5.2, a notable observation is the disparity in the number of hits recorded by sensors

1 and 2 compared to sensors 3, 4, and 5. This significant difference was later attributed to the presence of cracks in the medium between sensors 3, 4, and 5.

Examining the features, the pattern of hits and prominent peaks is quite similar for the first two sensors and the last three sensors, respectively. However, unlike the validation dataset, no clear patterns can be discerned across all sensors. The peaks in the histograms might be due to an influx of cars passing through the sensors within the 100-second time window represented by each bin. Despite the sensors being oriented parallel to the road section, a moveout of hits similar to what was observed in the validation dataset is unlikely, given the speed of the cars. With an assumed maximum speed in the Maastunnel of approximately 13.9 m/s (50 km/h), it would only take a car 0.04 seconds to pass a sensor spacing of 0.5 meters. Therefore, detecting the passage of an individual car using the recorded hit times is challenging. Since the traffic flow is nearly continuous, a uniformly distributed random noise field can be assumed. This is further supported by the lack of distinct patterns even after adjusting the bin sizes.

The next step was to determine the optimal normalization scheme for the dataset. The schemes tested were identical to those used for the validation dataset, including one-bit and RAM normalization in the time domain and spectral whitening using an envelope function. The running window length for RAM normalization in the time domain and the window used for the envelope function in the frequency domain were the same for both the validation and Maastunnel datasets. The results for sensor pair 1-2 are presented in the figure below, with results for the remaining pairs available in Appendix A.

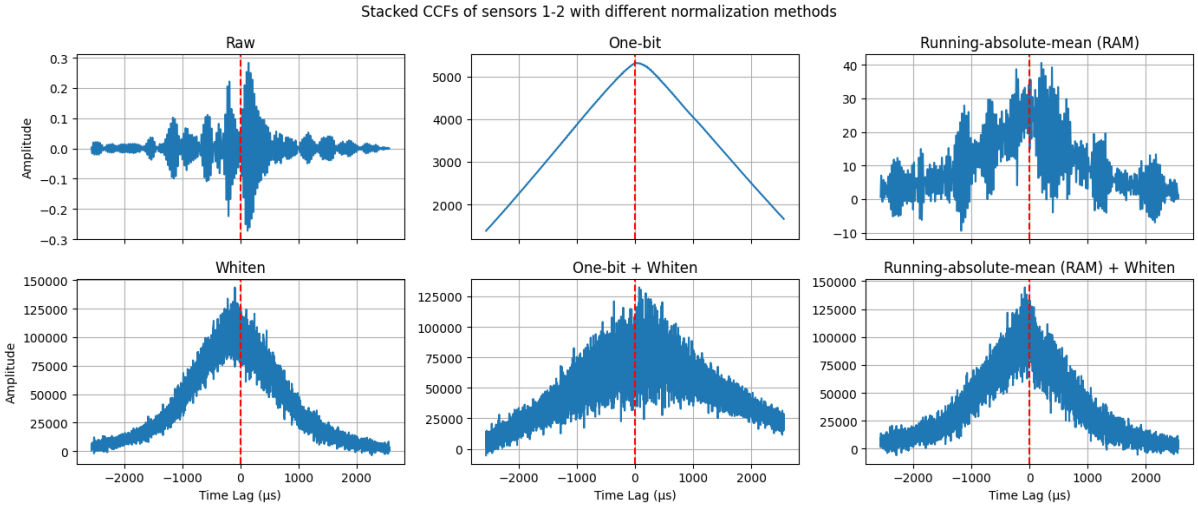

Figure 5.3: Normalization schemes for the stacked CCFs of sensor pair 1-2.

In Figure 5.3, the most coherent figure is the non-normalized raw stacked CCF. Among the six figures, it shows the clearest peaks and is least affected by the low-frequency component that introduces a wave-like pattern. Another observation is that the stacked CCF from one-bit normalization results in a single smooth peak because the normalization reduces amplitudes to either -1 or 1, removing any amplitude variations. Hypothetically, the non-normalized raw stacked CCF would be the best choice among the options. However, it is essential to conduct spectral whitening to equalize the frequency content of the CCFs, particularly for this real-world dataset with its significant noise uncertainties. Therefore, only the figures in the second row were considered and compared.

Among the three schemes in the second row, the better options appeared to be either the spectrally whitened scheme alone or the RAM + spectral whitening. However, it was difficult to determine which of these would be the best option, especially given the very incoherent results produced by the one-bit

normalization. Therefore, all three schemes were considered in the next step, which is the frequency analysis.

To determine a rough frequency range to investigate, the spectrograms of the hits per sensor and the frequency spectra of the sensor hits were examined. Below, only the spectrograms for sensors 1 and 3 are presented, as they were representative of their respective 'groups' observed in Figures 5.1 and 5.2.

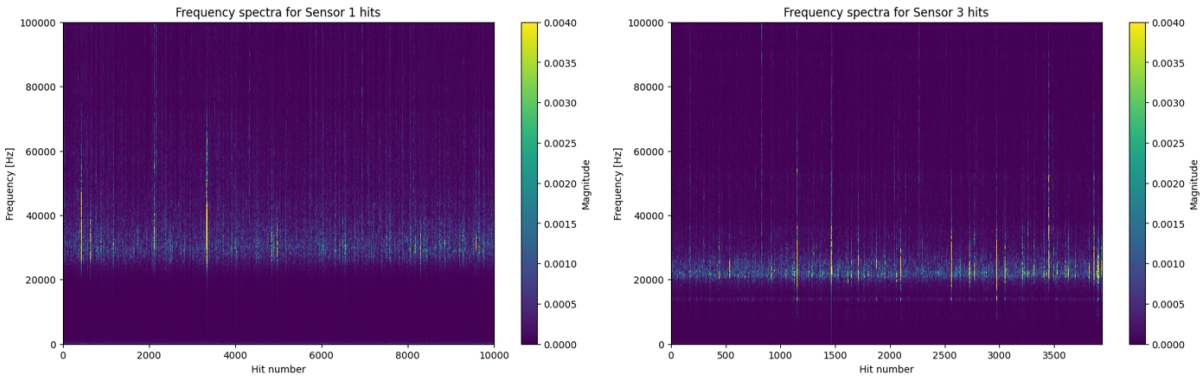

Figure 5.4: Frequency spectra of sensors 1 and 3.

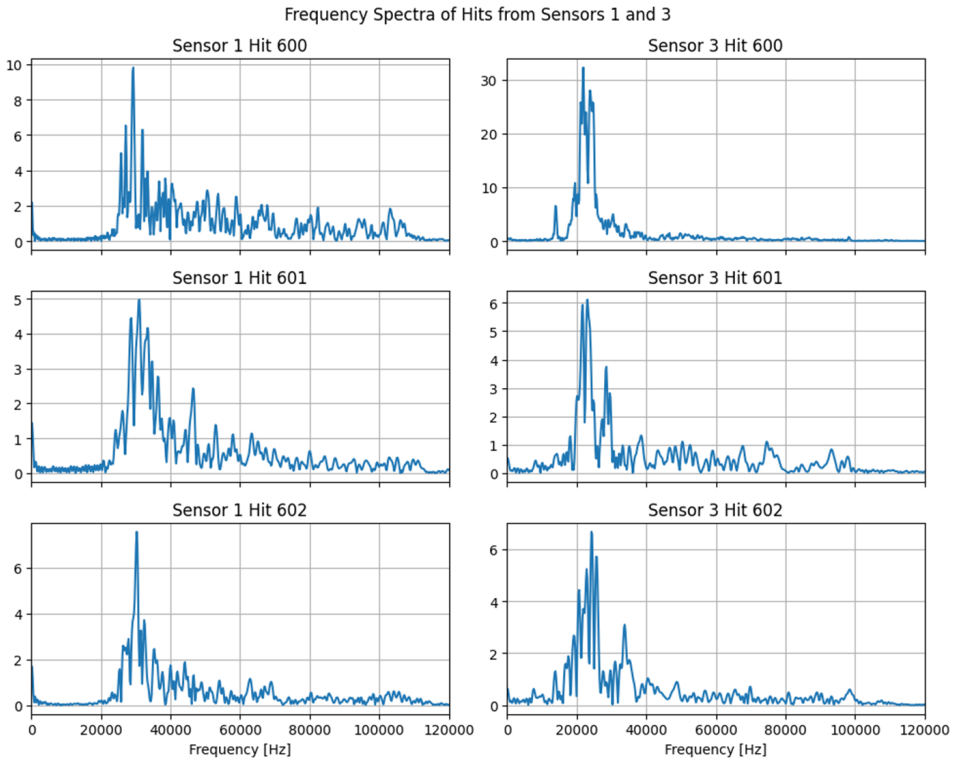

Figure 5.5: Frequency spectra of some hits of sensors 1 and 3.

In Figure 5.4, the spectrogram for the hits on sensor 1 shows a concentration in the 20 - 40 kHz frequency range, while for sensor 3, it is primarily within 20 - 30 kHz. Both spectrograms also display a few outliers exceeding 100 kHz. This observation is further supported in Figure 5.5, where the frequency distributions of a few randomly selected hits for both sensors are presented.

A common feature in the signals from both sensors is that, despite the high concentration within certain frequency ranges, a broad band of frequencies is still present. Smaller peaks around 100 kHz can also

be observed, particularly in hits 600 to 602 for sensor 1 and hits 601 and 602 for sensor 3.

Additionally, the amplitude of the frequency spectra is worth noting, especially for hit 600 of sensor 3. This hit stands out because the frequencies appear to flatten after 60 kHz. However, this is not due to the absence of higher frequencies but rather the strong concentration around 20 kHz, with an amplitude of about 30, which makes the other frequencies appear much lower by comparison. When compared to the other hits in Figure 5.5, hit 600 exhibits a different magnitude. The presence of hits like this emphasize the importance of applying spectral whitening.

Given these observations, the next step involved applying a series of frequency filters to the normalized GF estimation for sensor pair 1-2 using the three normalization schemes discussed previously. It was observed that including frequencies below 10 kHz introduced a low-frequency component that dominated the GF estimates. To address this, band-pass filtering with 10 kHz as the lower limit was used. The figure below shows the results of this band-pass filtering. Details on the low-pass and high-pass filtering applied can be found in Appendix A11 and A12.

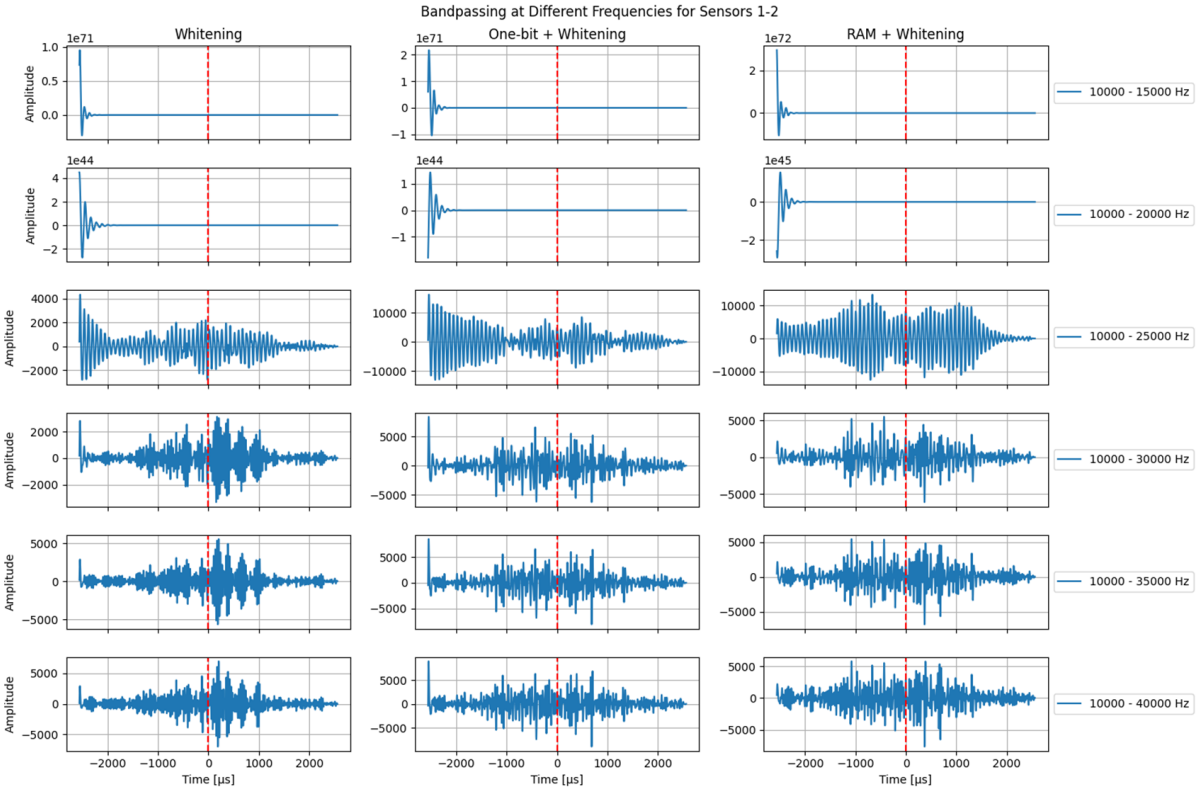

Figure 5.6: Band-passing for sensor pair 1-2.

In Figure 5.6, it is observed that for all three normalization schemes, the GF estimates stabilize after the upper limit reaches 30 kHz. For the whitened and RAM + whitened schemes, noticeable differences continue up to 35 kHz, beyond which changes are minimal. Higher frequency signals beyond these upper limits are found to be incoherent, likely due to the increased attenuation of higher frequency components.

Therefore, the ideal frequency range was determined to be 10–30 kHz for the one-bit normalization scheme and 10–35 kHz for the other two schemes. Among these, the combination of one-bit normalization and spectral whitening produced the most coherent GF estimates, with the least high-frequency noise and the clearest arrivals.

For the subsequent analysis of GF estimations from each sensor pair, one-bit normalization with spectral whitening was used as the normalization scheme. The signals were then band-pass filtered with a lower limit of 10 kHz and an upper limit of 30 kHz.

5.2. Green's Function Estimation

In this section, it is explored whether the estimated GFs or impulse responses from the traffic noise, with the chosen pre-processing approach applied, yield valuable information regarding the medium between the sensors. To do this, GF estimations for each pair are presented in the figure below.

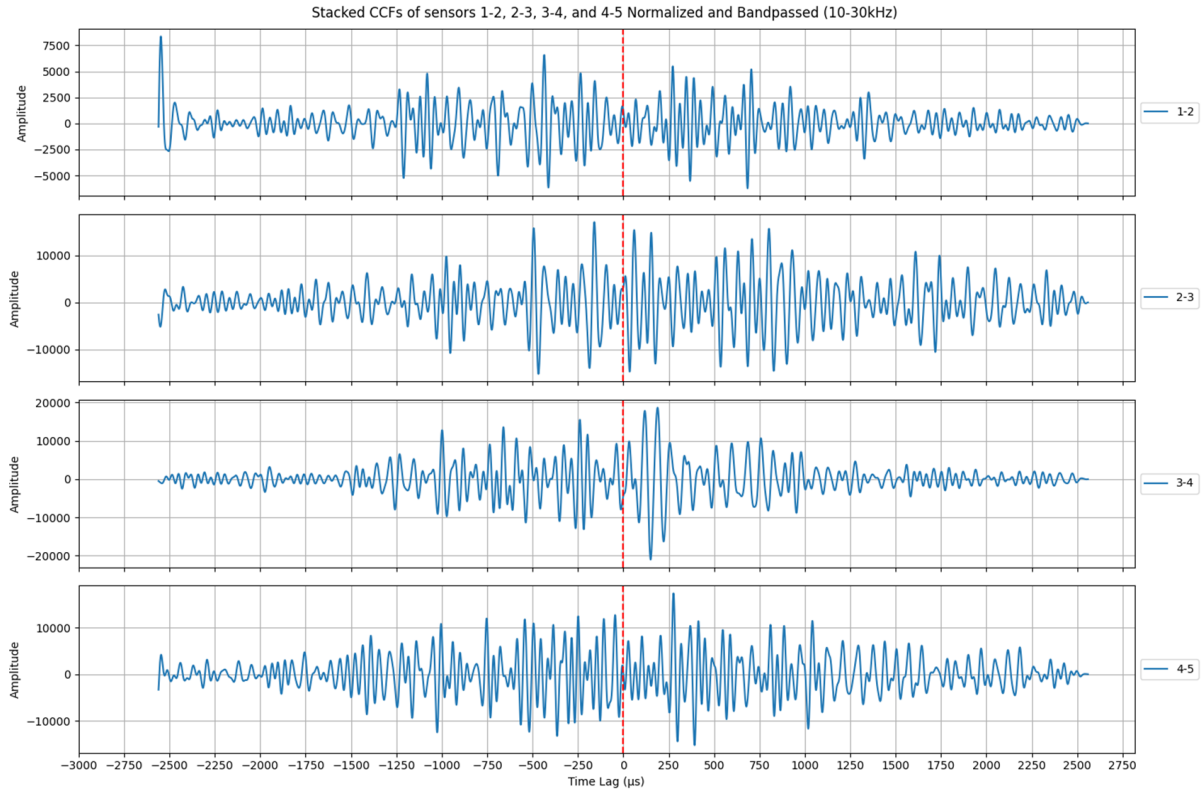

Figure 5.7: Normalized and Bandpassed (10-30kHz) GF estimations for all sensor pairs.

As depicted in Figure 5.7, despite implementing one-bit normalization, spectral whitening, and band-pass filtering within a reduced noise frequency range, the data still shows numerous peaks and noise that obscure meaningful information. Unlike the validation dataset, where recordings before and after crack occurrences allowed for establishing a baseline, the data in this study was based solely on recordings collected during installation. This situation resembles the challenges observed with the web sensor groups in the validation datasets, where some amplitude peaks were discernible but the arrivals remained unclear. In the validation dataset, knowledge of the structural condition enabled deductions linking the behavior of the GF estimations to the structural integrity of the member. For the Maastunnel dataset, no prior recordings were available, and knowledge of the condition of the road section was incomplete. Therefore, the results obtained in this study may serve as the 'baseline,' assuming no significant damages or defects were present at the time of measurement. However, without further tests, this assumption cannot be validated.

To further explore the GF estimations from this dataset, an additional investigation was conducted to attempt to improve signal coherence and resolution. In some practices where there is an unequal dis-

tribution of sources around given sensor pairs, averaging the causal and acausal parts of the stacked CCFs is done to enhance signal coherence and resolution, under the assumption that non-essential noises may be averaged out. This method was also explored in this study and is illustrated in Figure 5.8.

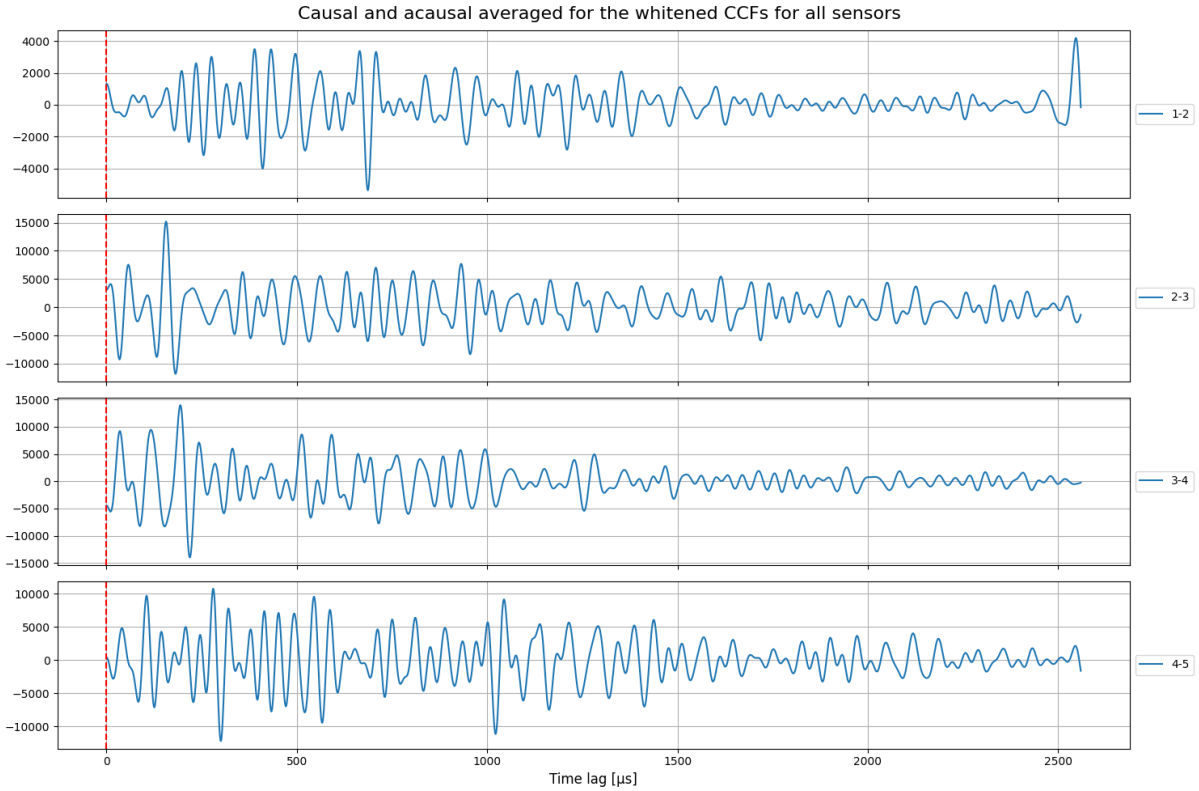

Figure 5.8: Averaged causal and acausal parts for all sensors.

In Figure 5.8, it can be observed that averaging the causal and acausal parts of the stacked CCFs does not lead to a significant improvement in coherence. This result should be noted for future studies, as further validation may be required to confirm its effectiveness.

In the following section, the GF estimations were decomposed using EEMD. The objective was to further decompose the estimates in order to extract distinct features, such as clear peaks in amplitudes, which could indicate significant events.

5.2.1. Decomposing Green's function estimations using EEMD

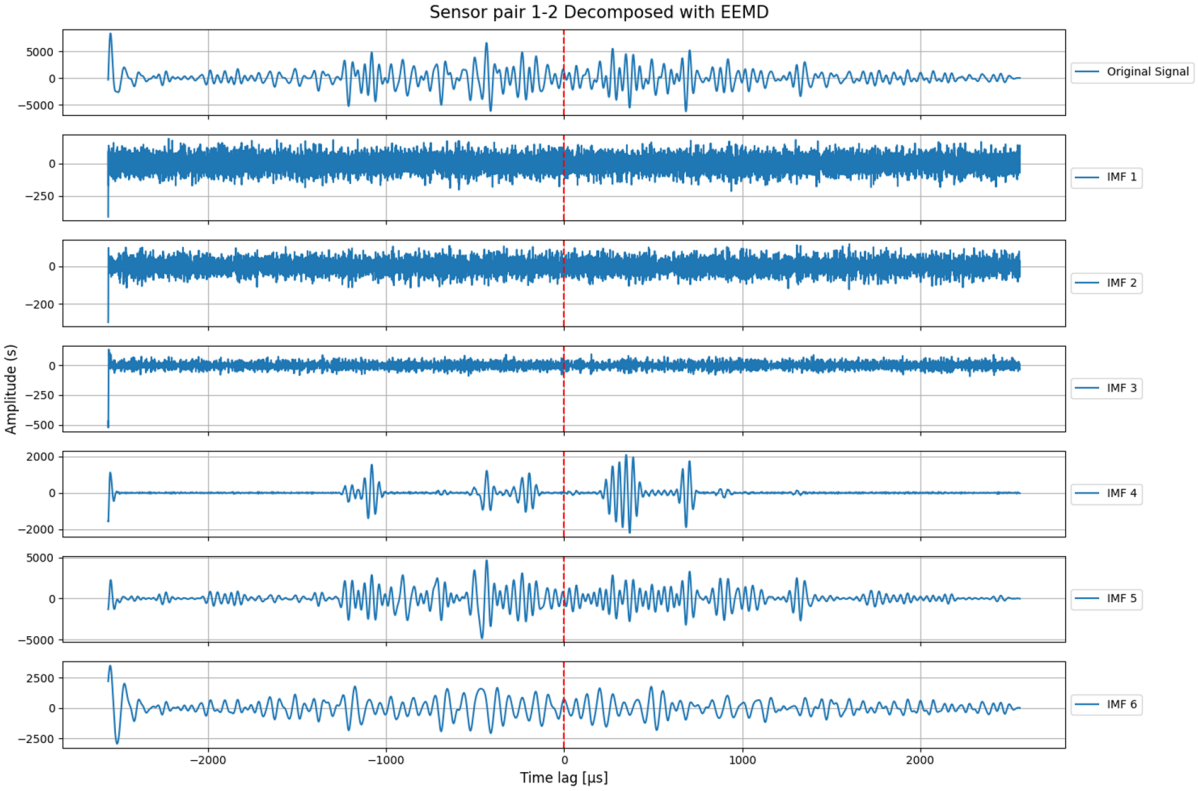

Figure 5.9: Decomposed Green's function estimation of pair 1-2 with EEMD.

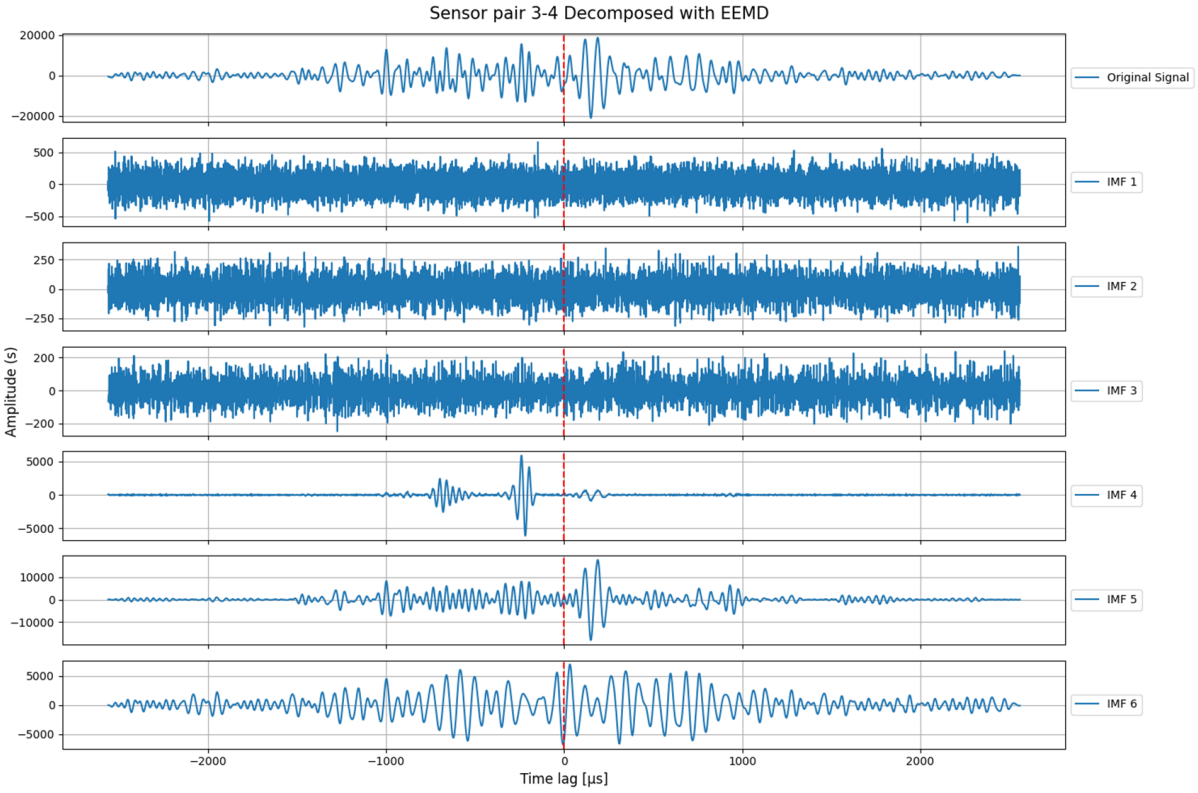

Figure 5.10: Decomposed Green's function estimation of pair 3-4 with EEMD.

As seen in Figures 5.9 and 5.10, the IMFs are extracted from the original signal after white noise is introduced. The first three IMFs represent high-frequency components due to the inherent high-frequency content of the signal in the time domain. By the time IMF4 is reached, these high-frequency components have been completely decomposed. Based on these observations, IMF4 was selected for further investigation across all sensor pairs. The comparison of four sensor pairs is illustrated in Figure 5.11. The EEMD-decomposed IMFs for sensor pairs 2-3 and 4-5 can be found in Appendix A.

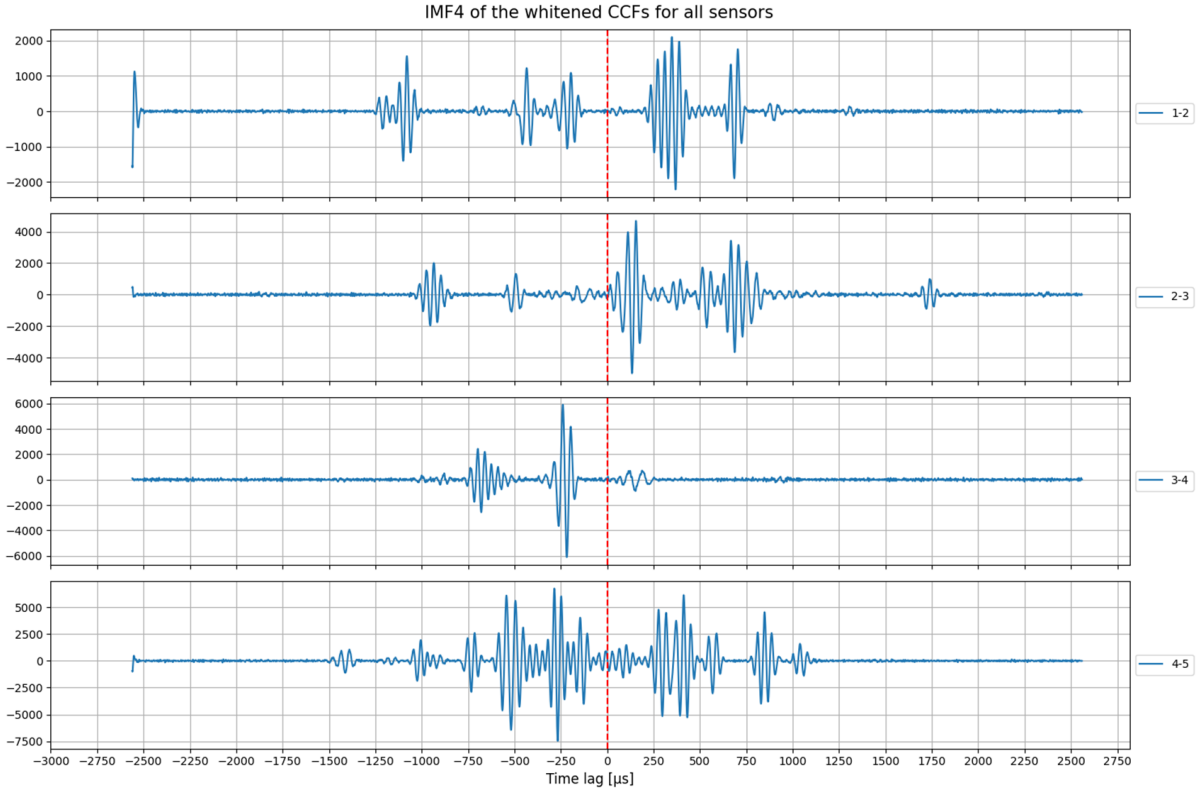

Figure 5.11: IMF4s of all sensor pairs.

In Figure 5.11, it can be observed that, compared to Figure 5.7, the GF estimations exhibit clearer arrivals due to successful decomposition into multiple IMFs using EEMD. The chosen mode function, IMF4, features high amplitude and relatively lower frequency components, making the peaks more discernible. For example, in sensor pair 1-2, multiple arrivals are evident at approximately -1000, -125, 250, and 600 microseconds. Although the arrival at -125 microseconds corresponds to a wave speed of 4000 m/s, which matches the P-wave speed estimated in Section 2.4.1, its accuracy is still uncertain and currently unverifiable. Furthermore, while clear peak arrivals were anticipated for pair 2-3, they were not expected for pairs 3-4 and 4-5 due to the cracks observed around those sensors.

Due to the inconsistencies observed and the current uncertainties, it was difficult to draw definitive interpretations at this stage.

In this section, it was demonstrated that EEMD had the capability to bring out and emphasize features such as peaks and arrival times of waves. However, it was challenging to assert certainty in the observations. Nonetheless, further investigation into EEMD in future studies may be worthwhile.

6

Comparative Analysis of Ambient Traffic Noise and Validation Datasets

In this chapter, comparisons of the findings between the two datasets are discussed in terms of inherent characteristics, appropriate pre-processing schemes, relevant frequency bands, and the coherence of resulting Green's Function estimations, as well as their applicability for Structural Health Monitoring (SHM).

6.1. Dataset Characteristics and Optimal Pre-processing Schemes

In terms of characteristics, the differences between the datasets were predictable, as seen in the plotted histograms for each dataset. The hits for the laboratory experiment were clear, with known sources, unlike the ambient traffic data, which involved randomly-sourced signals. Additionally, the sensor locations relative to the noise sources significantly impacted the datasets. Within the laboratory experiment itself, substantial differences were observed between the two sensor groups, and even more pronounced differences emerged when comparing the laboratory data to the ambient traffic data.

The web sensor group in the laboratory dataset more closely resembled the ambient traffic data, as both involved sound waves traveling through the medium. In contrast, the surface sensors in the laboratory considered only surface waves, which were more straightforward.

In terms of relevant frequencies, the validation dataset achieved sufficient coherence within the 20 to 60 kHz range after pre-processing. Conversely, the coherent range for the ambient traffic dataset was found to be 10 to 30 kHz. Despite both datasets featuring the same peak frequency ranges at around 20 to 40 kHz, the ambient traffic dataset featured a broader frequency spectrum with components that ranged from 0 to 110 kHz. For both datasets, including frequencies outside the determined bands in their GF estimations introduced dominant low-frequency trends and high frequency components that obscured wave arrivals.

In terms of normalization procedures, spectral whitening was applied to both datasets to normalize the energy and ensure consistent representation across the entire frequency bandwidth. However, for the ambient traffic dataset one-bit normalization was additionally used to enhance signal coherence. In contrast, none of the investigated temporal normalization methods were applied to the validation dataset,

as they were not beneficial.

6.2. Green's Function Estimations and Practical Applicability

As mentioned in the previous section, the relatively clear understanding of sound source distribution within the validation dataset made the analysis and interpretation of the GF estimations from different sensor pairs easier. This was particularly evident for the top layer sensors, where the dominant direction of energy could be directly linked to the location of the sensor relative to the chain path. Similarly, although less directly, this trend was also observed with the web sensor group. Despite multiple arrivals and peak amplitudes on both sides of the GF estimations, the more dominant side remained apparent.

In contrast, the ambient traffic noise dataset introduced greater complexity due to the random distribution of noise sources, which is more reflective of the ambient noise encountered in seismology. The continuous and unpredictable traffic flow presented a more realistic challenge for interpreting GF estimations. Despite these challenges, this complexity offers valuable insights into real-world applications, contrasting with the validation dataset where fewer uncertainties provided clearer explanations.

It is also worth noting that the GF estimations for the web sensor group in the validation dataset can still be compared to those from the ambient traffic noise dataset, particularly regarding multiple arrivals.

Additionally, as mentioned in Section 5.2, the ambient traffic noise dataset faced the challenge of lacking a verifiable baseline. This was the first attempt to reconstruct the dataset from that specific structure with those particular sensors, and the absence of active testing to support the findings made it difficult to make definitive claims. Nevertheless, the results from the ambient dataset were promising, as some features observed could be corroborated, as discussed in Section 5.2. Furthermore, with more data available, it is highly likely that the GF estimations using the algorithm in this research will achieve a higher signal-to-noise ratio, making the GF estimations more coherent.

For future work, it may be valuable to use the findings from this research as a basis for comparison with future studies.

7

Concluding Remarks

The previous chapters presented an exploratory study focused on addressing two primary questions: 1) whether the Green's function of a concrete medium can be estimated through the interferometry of ambient passive noise, and 2) how effective the estimated Green's function is in monitoring the structural health of concrete structures in both laboratory and real-world conditions. In this chapter, the findings from this research allow to answer these questions. Additionally, recommendations for further research, pertinent to this study but beyond the scope, will also be provided.

7.1. Conclusions

The main conclusions from this research are listed below:

- The Green's function of a concrete medium can be reliably estimated through interferometry using signals received by a pair of sensors, provided there is a sufficient volume of data to yield a higher signal-to-noise ratio (SNR). This Green's function estimation accurately reflects the arrivals of waves passing through the sensor pair.
- The analysis of the validation dataset revealed that the Green's function estimation is effective for monitoring changes in concrete, such as crack formation and strain variations. This effectiveness was observed for sensor pairs spaced up to 0.8 meters. However, the performance at greater distances requires further verification.
- The analysis revealed that the validation dataset achieved coherence between 20 and 60 kHz, while for the ambient traffic noise, it was coherent between 10 and 30 kHz, despite its broader frequency range. Spectral whitening was necessary for both datasets to equalize the distribution of frequencies. Additionally, one-bit normalization was necessary for the ambient traffic noise, whereas the validation dataset provided the best results without any temporal normalization.
- Upon comparing the Green's function estimations from the two datasets, it is evident that the estimations vary significantly depending on the placement of sensors relative to the noise source. This difference was particularly evident when comparing the estimations from different sensor groups within the validation dataset.
- Based on the current findings of this research, it cannot be conclusively inferred that ambient traffic noise is suitable for structural health monitoring (SHM) of real structures. While coherent patterns

in the Green's function estimation from the sensor pairs in the concrete tunnel were observed after signal processing, SNR remained poor. This likely resulted from the limited amount of available data and insufficient recording time.

7.2. Recommendations and Outlooks for Future Works

In addition to the conclusions, the following are the recommendations for those who wish to pursue this topic further:

- Collect data over a longer period per recording and across more days to obtain a more representative Green's function.
- Explore using a lower sampling frequency to get more manageable data file sizes to accommodate longer recordings and faster data processing.
    - In this study, a sampling frequency of 120 kHz would have been theoretically sufficient for the chosen frequency ranges for both datasets, allowing for a reduction in the number of samples by factors of approximately 20 and 40 for the validation and real-world datasets, respectively. However, the analysis showed that having more samples was more advantageous, as it improved the accuracy and spatial resolution of the results (see Figure A.15 in Appendix A). Therefore, a proper sampling frequency remains to be explored.
- Explore alternative sampling methods, such as continuous recording over multiple fixed time windows. Synchronized and continuous data recording eliminates the need to align 'hits' and more closely resembles seismic noise interferometry.
- Investigate alternative decomposition methods to denoise data and emphasize specific features. For instance, consider further exploring Ensemble Empirical Mode Decomposition (EEMD), which was used in this study.
- Reevaluate frequency ranges suitable for analyzing real-world traffic noise data and explore additional normalization techniques.
- Conduct an active experiment where one sensor - assuming it is a transducer - can act as a source so that the true Green's function can be compared to the one estimated through interferometry.

References

[1] Mohamed Abdo. *Structural Health Monitoring, History, Applications and Future. A Review Book*. Jan. 2014. ISBN: 978-1-941926-07-9.

[2] Hirotogu Akaike. “Information Theory and an Extension of the Maximum Likelihood Principle”. en. In: *Selected Papers of Hirotugu Akaike*. Ed. by Emanuel Parzen, Kunio Tanabe, and Genshiro Kitagawa. New York, NY: Springer, 1998, pp. 199–213. ISBN: 978-1-4612-1694-0. DOI: 10.1007/978-1-4612-1694-0_15. URL: https://doi.org/10.1007/978-1-4612-1694-0_15 (visited on 07/03/2024).

[3] Anthony Bedford and Douglas S. Drumheller. *Introduction to Elastic Wave Propagation*. en. Cham: Springer Nature Switzerland, 2023. ISBN: 978-3-031-32874-9 978-3-031-32875-6. DOI: 10.1007/978-3-031-32875-6. URL: https://link.springer.com/10.1007/978-3-031-32875-6 (visited on 04/10/2024).

[4] GD Bensen et al. “Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements”. In: *Geophysical Journal International* 169 (June 2007), pp. 1239–1260. DOI: 10.1111/j.1365-246X.2007.03374.x.

[5] Peter Bormann, Eric Engdahl, and Rainer Kind. “Seismic Wave Propagation and Earth models”. In: Jan. 2012. DOI: 10.2312/GFZ.NMSOP-2_ch2.

[6] Michel Campillo and Anne Paul. “Long-Range Correlations in the Diffuse Seismic Coda”. In: *Science (New York, N.Y.)* 299 (Feb. 2003), pp. 547–9. DOI: 10.1126/science.1078551.

[7] A. Carpinteri et al. “Reliable onset time determination and source location of acoustic emissions in concrete structures”. en. In: *Cement and Concrete Composites* 34.4 (Apr. 2012), pp. 529–537. ISSN: 09589465. DOI: 10.1016/j.cemconcomp.2011.11.013. URL: https://linkinghub.elsevier.com/retrieve/pii/S0958946511002149 (visited on 05/18/2024).

[8] Hao Cheng et al. “Comparing the stretching technique and the wavelet cross-spectrum technique for measuring stress-induced wave-velocity changes in concrete”. In: *Automation in Construction* 158 (Feb. 2024), p. 105221. DOI: 10.1016/j.autcon.2023.105221.

[9] *Experiments*. URL: https://concrete-prediction-contest.tudelft.nl/index.php/information/experiments (visited on 07/04/2024).

[10] Andreas Fichtner, Daniel Bowden, and Laura Ermert. “Optimal processing for seismic noise correlations”. en. In: *Geophysical Journal International* 223.3 (Sept. 2020), pp. 1548–1564. ISSN: 0956-540X, 1365-246X. DOI: 10.1093/gji/ggaa390. URL: https://academic.oup.com/gji/article/223/3/1548/5893300 (visited on 07/17/2024).

[11] Danny Fu. “Application of Coda Wave Interferometry with embedded Smart Aggregate in a pre-stressed cast-in-situ concrete continuous slab bridge”. en. In: (2022). URL: https://repository.tudelft.nl/islandora/object/uuid%3A7de3512e-d984-490d-ba2d-38df4c036197 (visited on 03/04/2024).

[12] Ramesh Gomasa et al. “A review on health monitoring of concrete structures using embedded piezoelectric sensor”. In: *Construction and Building Materials* 405 (Nov. 2023), p. 133179. ISSN: 0950-0618. DOI: 10.1016/j.conbuildmat.2023.133179. URL: https://www.sciencedirect.com/science/article/pii/S0950061823028969 (visited on 03/04/2024).

[13] J.c. Groos, S. Bussat, and J.r.r. Ritter. “Performance of different processing schemes in seismic noise cross-correlations”. en. In: *Geophysical Journal International* 188.2 (2012). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-246X.2011.05288.x, pp. 498–512. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.2011.05288.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.2011.05288.x (visited on 03/04/2024).

[14] Justin Harris. *Non-Destructive Evaluation of Concrete Using Electrical Resistivity and Ultrasonic Wave Propagation*. en. Google-Books-ID: KtJ7zwEACAAJ. University of Maine, 2022.

[15] Norden Huang et al. “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454 (Mar. 1998), pp. 903–995. DOI: 10.1098/rspa.1998.0193.

[16] D. S. Hughes and J. L. Kelly. “Second-Order Elastic Deformation of Solids”. In: *Physical Review* 92.5 (Dec. 1953). Publisher: American Physical Society, pp. 1145–1149. DOI: 10.1103/PhysRev.92.1145. URL: https://link.aps.org/doi/10.1103/PhysRev.92.1145 (visited on 03/04/2024).

[17] Yapeng Li et al. “Acoustoelastic effect simulation by time–space finite element formulation based on quadratic interpolation of the acceleration”. en. In: *Wave Motion* 93 (Mar. 2020), p. 102465. ISSN: 01652125. DOI: 10.1016/j.wavemoti.2019.102465. URL: https://linkinghub.elsevier.com/retrieve/pii/S0165212518305031 (visited on 08/19/2024).

[18] Ivan Lillamand et al. “Acoustoelastic effect in concrete material under uni-axial compressive loading”. In: *NDT & E International* 43.8 (Nov. 2010), pp. 655–660. ISSN: 0963-8695. DOI: 10.1016/j.ndteint.2010.07.001. URL: https://www.sciencedirect.com/science/article/pii/S0963869510000861 (visited on 03/04/2024).

[19] Xiling Liu et al. “Research on the Attenuation Characteristics of High-Frequency Elastic Waves in Rock-Like Material”. In: *Materials* 15.19 (Sept. 2022), p. 6604. ISSN: 1996-1944. DOI: 10.3390/ma15196604. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9570857/ (visited on 08/03/2024).

[20] Zinoviy Nazarchuk, Valentyn Skalskyi, and Oleh Serhiyenko. “Propagation of Elastic Waves in Solids”. en. In: *Acoustic Emission: Methodology and Application*. Ed. by Zinoviy Nazarchuk, Valentyn Skalskyi, and Oleh Serhiyenko. Cham: Springer International Publishing, 2017, pp. 29–73. ISBN: 978-3-319-49350-3. DOI: 10.1007/978-3-319-49350-3_2. URL: https://doi.org/10.1007/978-3-319-49350-3_2 (visited on 04/10/2024).

[21] Andy St-Onge. “Akaike Information Criterion Applied to Detecting First Arrival Times on Microseismic Data”. In: 30 (Jan. 2011). DOI: 10.1190/1.3627522.

[22] Wonsiri Punurai. “Cement-based materials’ characterization using ultrasonic attenuation”. Publication Title: Ph.D. Thesis ADS Bibcode: 2006PhDT.....61P. PhD thesis. Jan. 2006. URL: https://ui.adsabs.harvard.edu/abs/2006PhDT.....61P (visited on 04/10/2024).

[23] He Ren, Xi Chen, and Yong Chen. “Structural Health Monitoring and Influence on Current Maintenance”. In: Dec. 2017, pp. 173–184. ISBN: 978-0-12-812668-4. DOI: 10.1016/B978-0-12-812668-4.00009-5.

[24] Muhammad Ryan. *Decomposing Signal Using Empirical Mode Decomposition — Algorithm Explanation for Dummy*. en. Nov. 2019. URL: https://towardsdatascience.com/decomposing-signal-using-empirical-mode-decomposition-algorithm-explanation-for-dummy-93a93304c541 (visited on 07/02/2024).

[25] Johannes Salvermoser, Céline Hadziioannou, and Simon C. Stähler. "Structural monitoring of a highway bridge using passive noise recordings from street traffic". eng. In: *The Journal of the Acoustical Society of America* 138.6 (Dec. 2015), pp. 3864–3872. ISSN: 1520-8524. DOI: 10.1121/1.4937765.

[26] Christoph Sens-Schönfelder and U Wegler. "Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia". In: *Geophysical Research Letters* 33 (Nov. 2006), p. L21302. DOI: 10.1029/2006GL027797.

[27] Roel Snieder et al. "Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem". en. In: *Physical Review E* 78.3 (Sept. 2008), p. 036606. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.78.036606. URL: https://link.aps.org/doi/10.1103/PhysRevE.78.036606 (visited on 07/18/2024).

[28] A. Sofi et al. "Structural health monitoring using wireless smart sensor network – An overview". In: *Mechanical Systems and Signal Processing* 163 (Jan. 2022), p. 108113. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2021.108113. URL: https://www.sciencedirect.com/science/article/pii/S0888327021004957 (visited on 03/04/2024).

[29] L. Stehly, M. Campillo, and N. M. Shapiro. "A study of the seismic noise from its long□range correlation properties". en. In: *Journal of Geophysical Research: Solid Earth* 111.B10 (Oct. 2006), 2005JB004237. ISSN: 0148-0227. DOI: 10.1029/2005JB004237. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2005JB004237 (visited on 05/29/2024).

[30] *TUNNELBOUW - Open Beelden*. URL: https://www.openbeelden.nl/media/1197110 (visited on 07/04/2024).

[31] Xin Wang, Ernst Niederleithinger, and Iris Hindersmann. "The installation of embedded ultrasonic transducers inside a bridge to monitor temperature and load influence using coda wave interferometry technique". en. In: *Structural Health Monitoring* 21.3 (May 2022). Publisher: SAGE Publications, pp. 913–927. ISSN: 1475-9217. DOI: 10.1177/14759217211014430. URL: https://doi.org/10.1177/14759217211014430 (visited on 03/04/2024).

[32] C. P. A. Wapenaar et al. "Tutorial on seismic interferometry: Part 1 — Basic principles and applications". en. In: *Geophysics, 75 (5), 2010* (2010). Publisher: Society of Exploration Geophysicists. ISSN: 0016-8033. DOI: 10.1190/1.3457445. URL: https://repository.tudelft.nl/islandora/object/uuid%3Aa7510463-5446-4a49-a8fc-81f44db1d984 (visited on 03/04/2024).

[33] Jian-Xin Wei, Bang-Rang Di, and Pinbo Ding. "Effect of crack aperture on P-wave velocity and dispersion". In: *Applied Geophysics* 10 (June 2013). DOI: 10.1007/s11770-013-0379-z.

[34] Zhaohua Wu and Norden Huang. "Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method". In: *Advances in Adaptive Data Analysis* 1 (Jan. 2009), pp. 1–41. DOI: 10.1142/S1793536909000047.

[35] Xiaotao Yang et al. "Optimal stacking of noise cross-correlation functions". In: *Geophysical Journal International* 232.3 (Mar. 2023), pp. 1600–1618. ISSN: 0956-540X. DOI: 10.1093/gji/ggac410. URL: https://doi.org/10.1093/gji/ggac410 (visited on 03/04/2024).

[36] Yingjie Yang et al. “Ambient noise Rayleigh wave tomography across Europe”. In: *Geophysical Journal International* 168.1 (Jan. 2007), pp. 259–274. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2006.03203.x. URL: https://doi.org/10.1111/j.1365-246X.2006.03203.x (visited on 03/04/2024).

[37] I. N. Yermolov. *Akustichestke Metody Kontrolya (Acoustic methods of testing)*. Vol. 2. Moskva: Vysshaya shkola, 1991.

[38] A. Zeiler et al. “Empirical Mode Decomposition - an introduction”. en. In: *The 2010 International Joint Conference on Neural Networks (IJCNN)*. Barcelona, Spain: IEEE, July 2010, pp. 1–8. ISBN: 978-1-4244-6916-1. DOI: 10.1109/IJCNN.2010.5596829. URL: http://ieeexplore.ieee.org/document/5596829/ (visited on 05/18/2024).

A

Appendix A

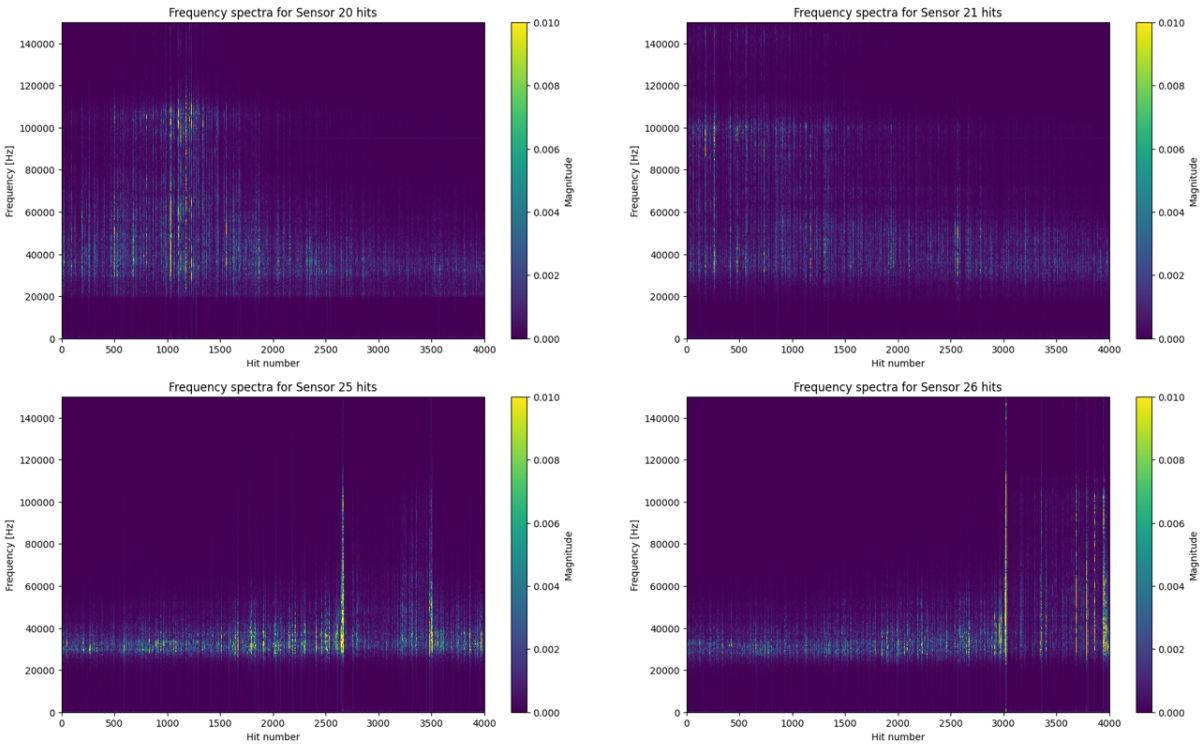

Figure A.1: Top sensor group spectograms (AE)

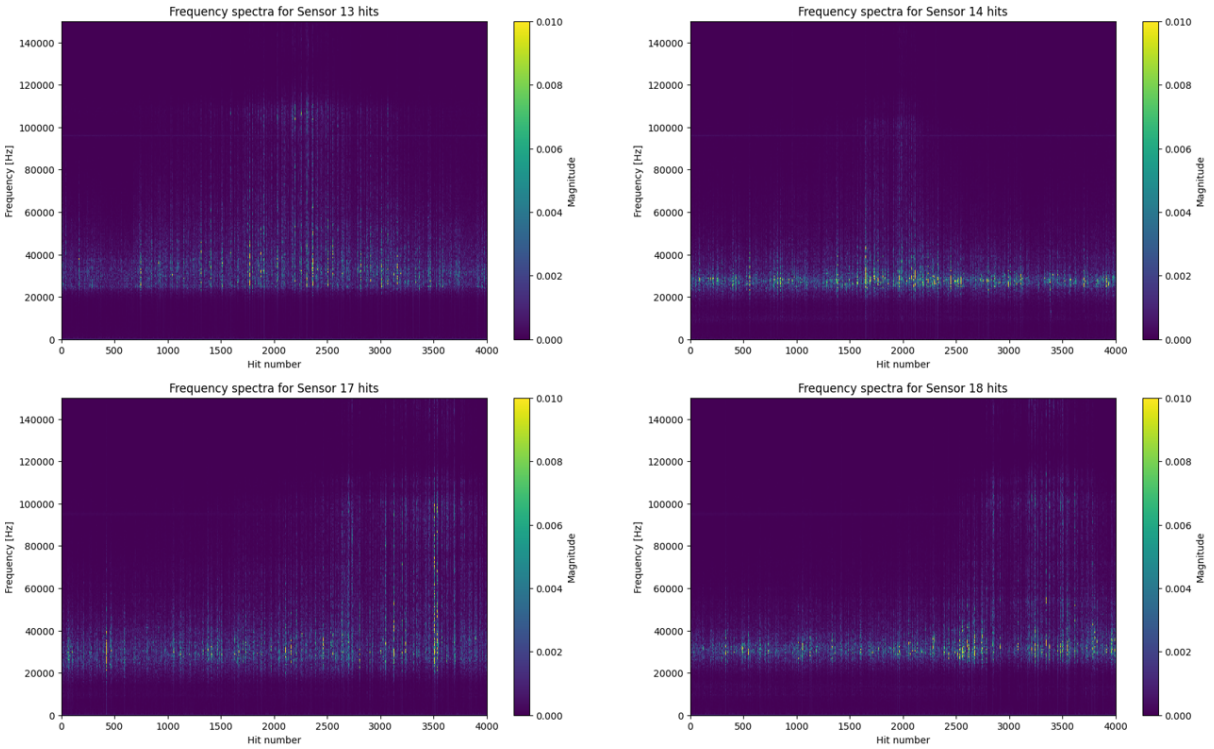

Figure A.2: Web sensor group spectograms (AE)

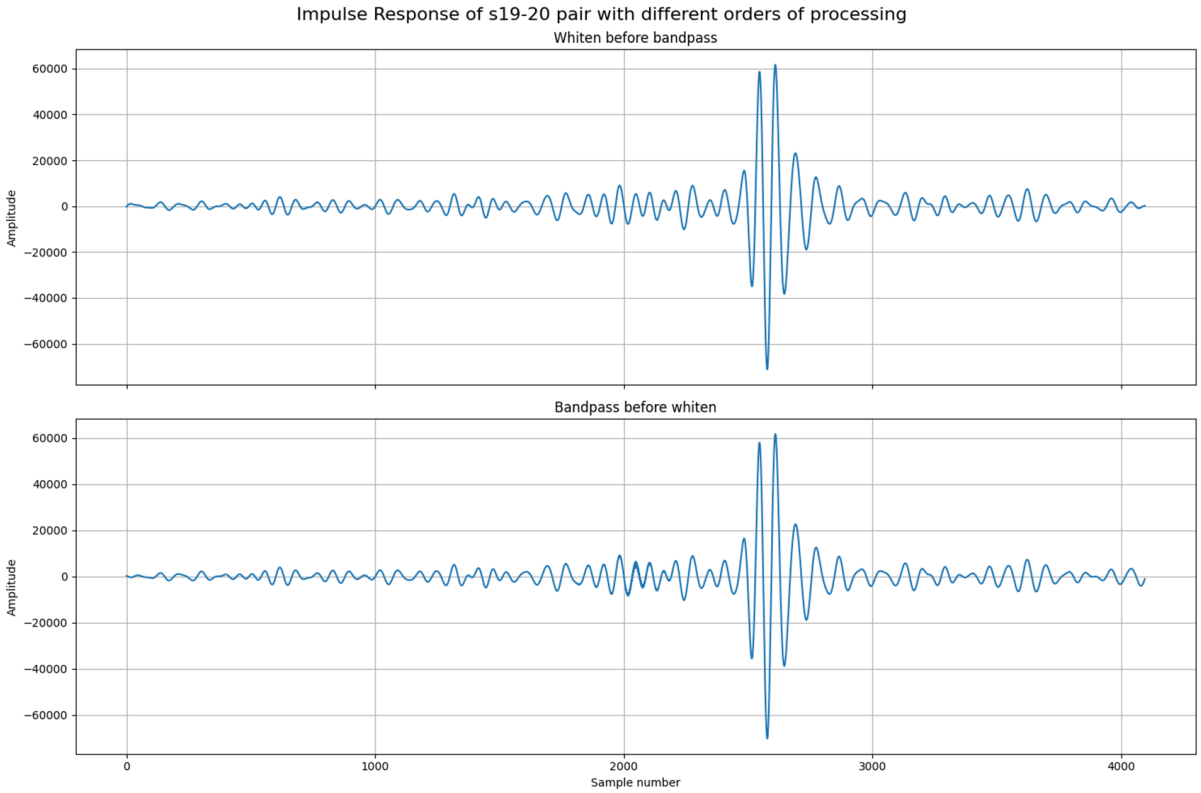

Figure A.3: Bandpassing at different points of the processing (AE).

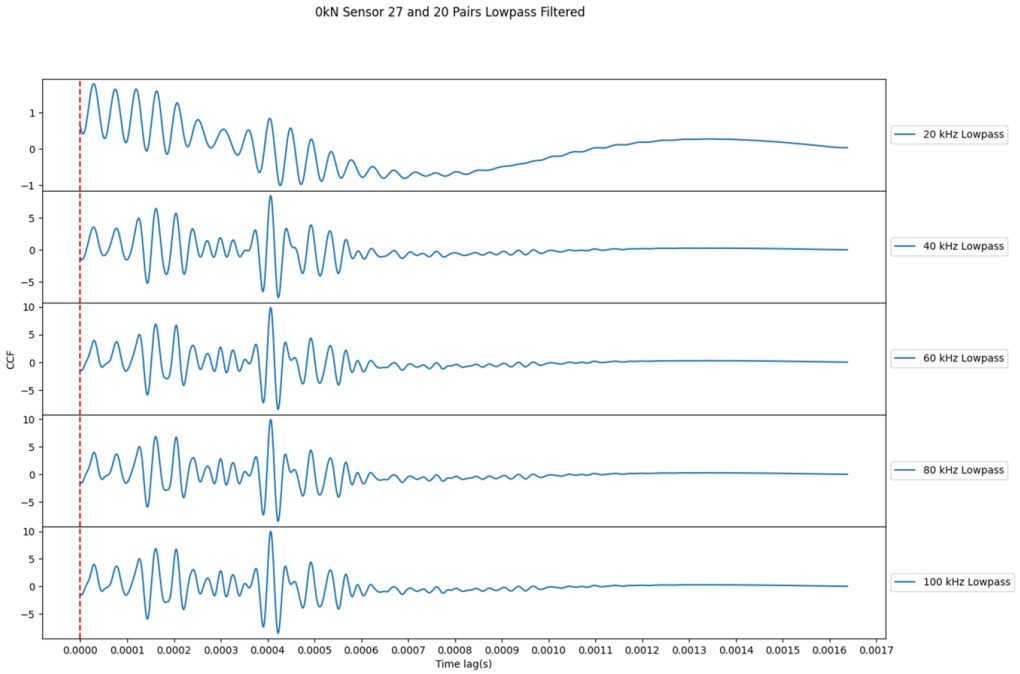

Figure A.4: Low-pass filtering of GF estimation of sensor pair 27-20 (AE).

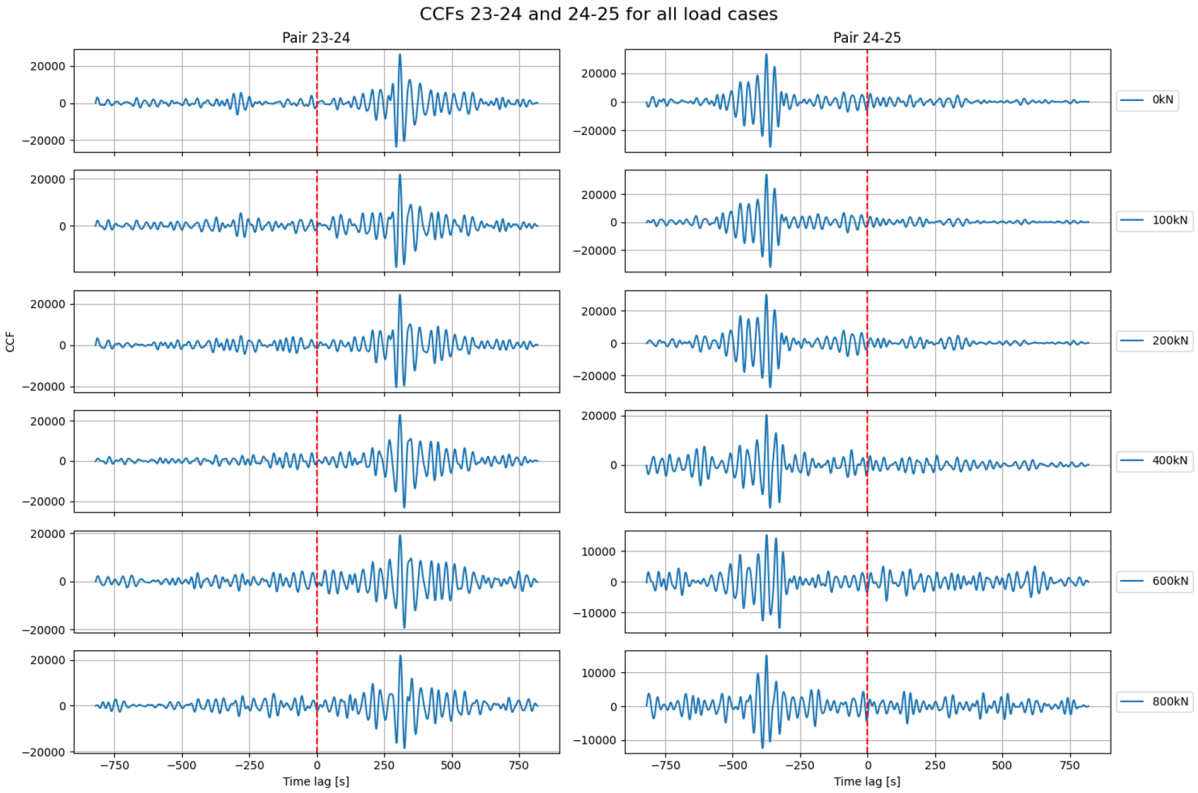

Figure A.5: Estimated GFs for s23-24 and s24-25 pairs.

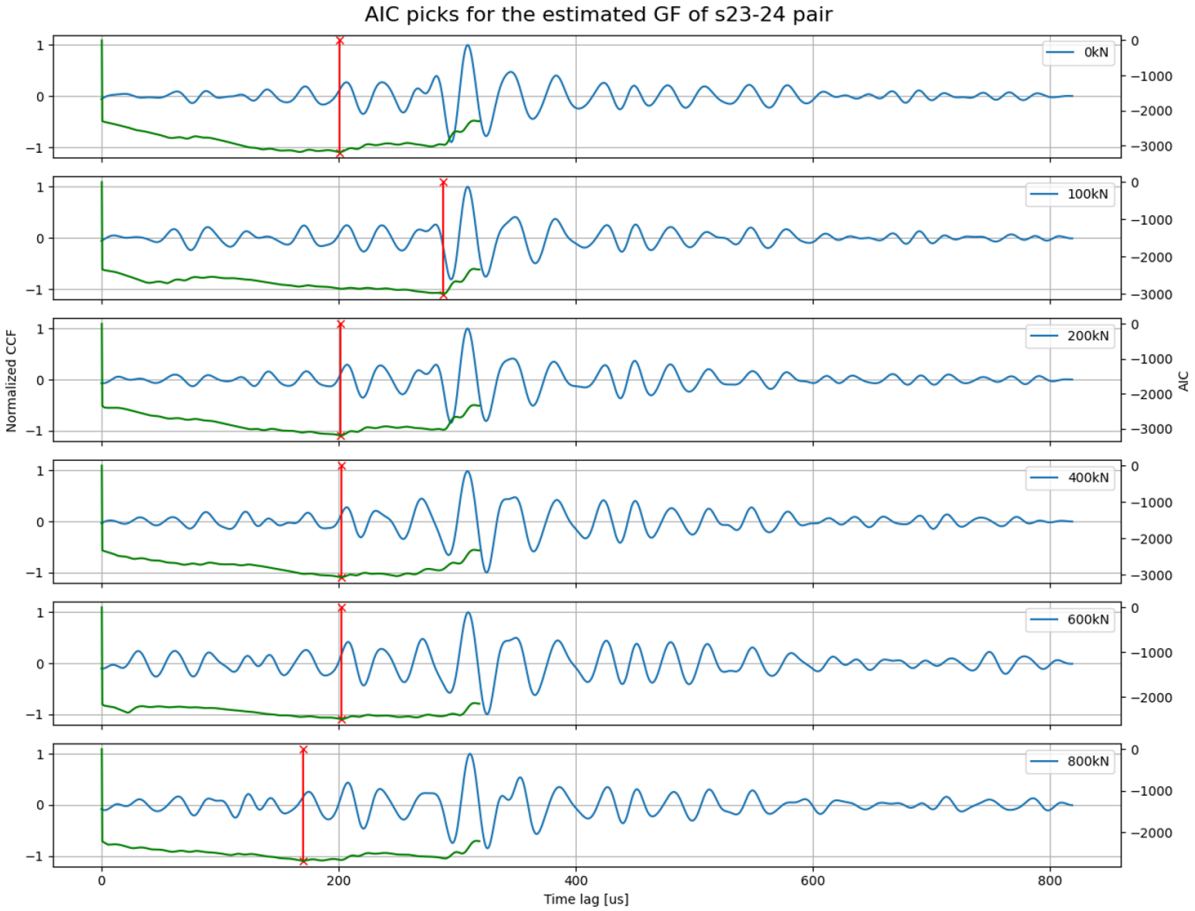

Figure A.6: AIC Picks for the estimated GF of s23-24 pair.

Table A.1: Arrival times and wave speeds for pair 23-24 at different load cases.

	Arrival time (μs)	**Filtered Speed (m/s)**
0kN	182.40	2741.2
100kN	100kN	165.60
200kN	200kN	182.00
400kN	400kN	12.80
600kN	600kN	99.60
800kN	800kN	186.40

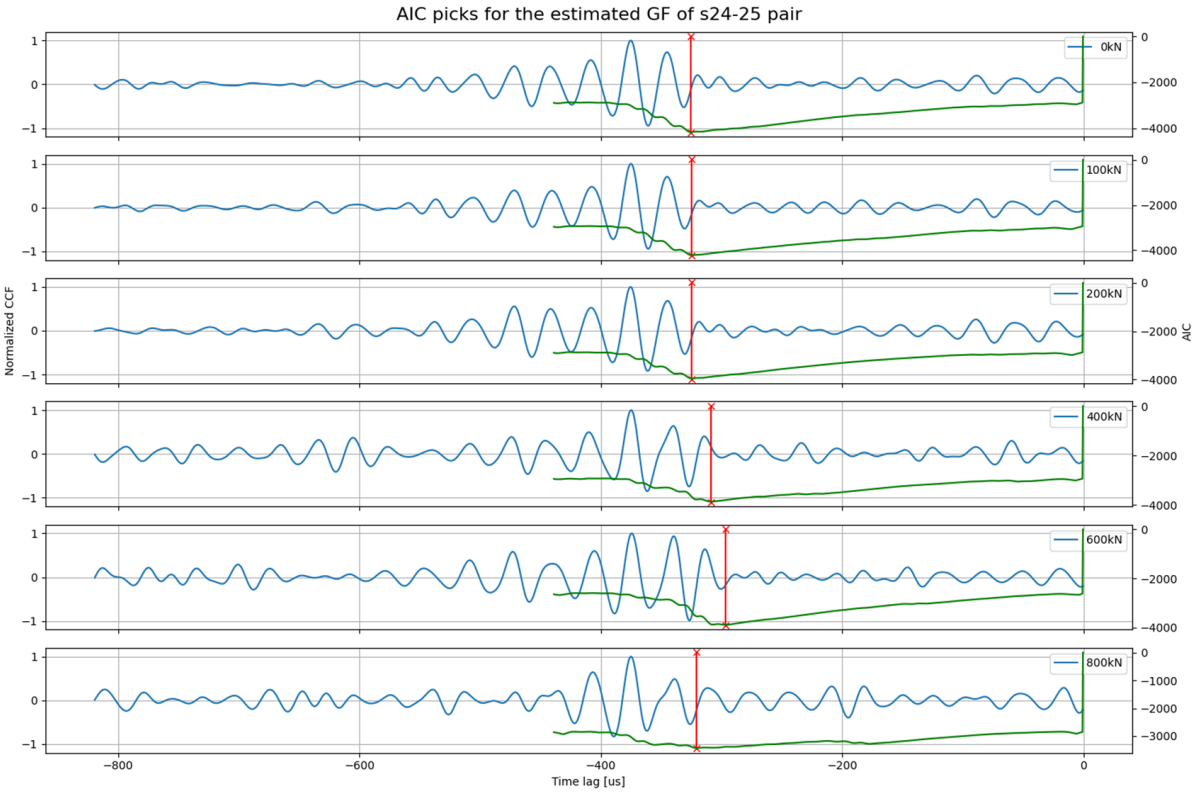

Figure A.7: AIC Picks for the estimated GF of s24-25 pair.

Table A.2: Arrival times and wave speeds for pair 24-25 at different load cases.

	Arrival time (μs)	**Filtered Speed (m/s)**
0kN	312.40	1600.5
100kN	308.40	1621.3
200kN	308.40	1621.3
400kN	226.00	2212.4
600kN	341.60	1463.7
800kN	9.60	52083.3

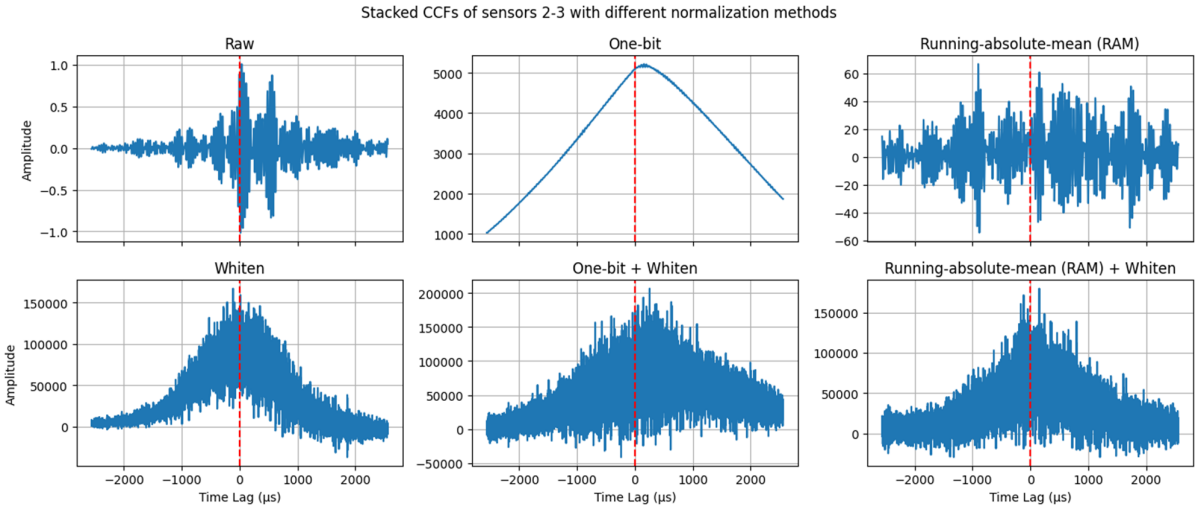

Figure A.8: Normalization schemes for the stacked CCFs of sensor pair 2-3.

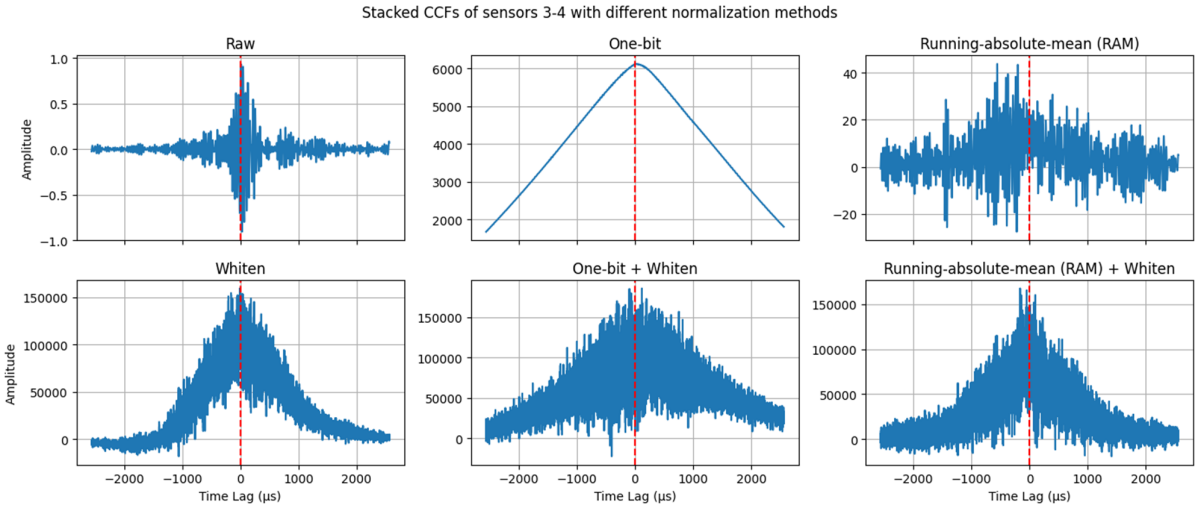

Figure A.9: Normalization schemes for the stacked CCFs of sensor pair 3-4.

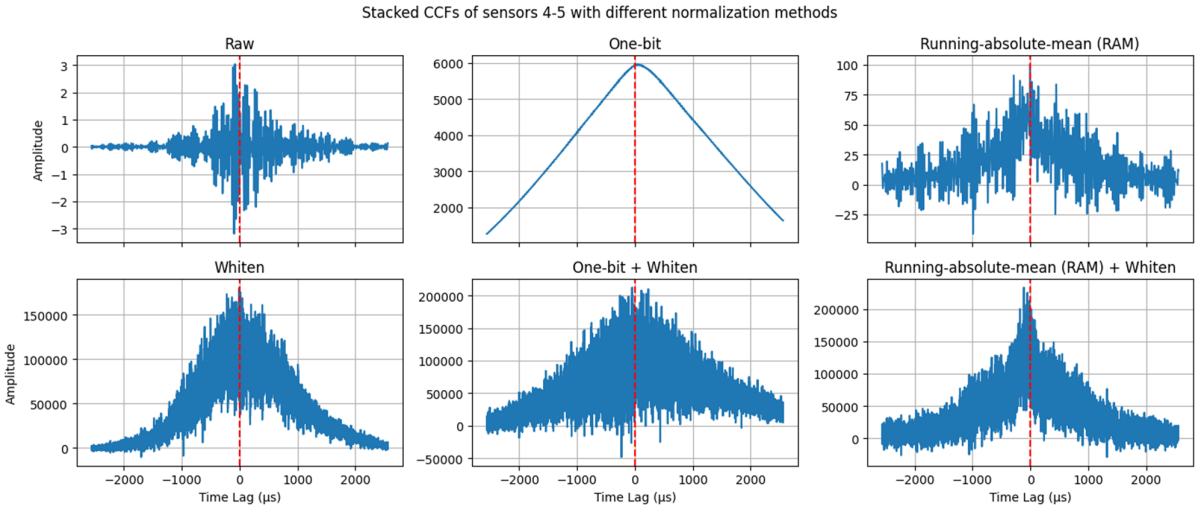

Figure A.10: Normalization schemes for the stacked CCFs of sensor pair 4-5.

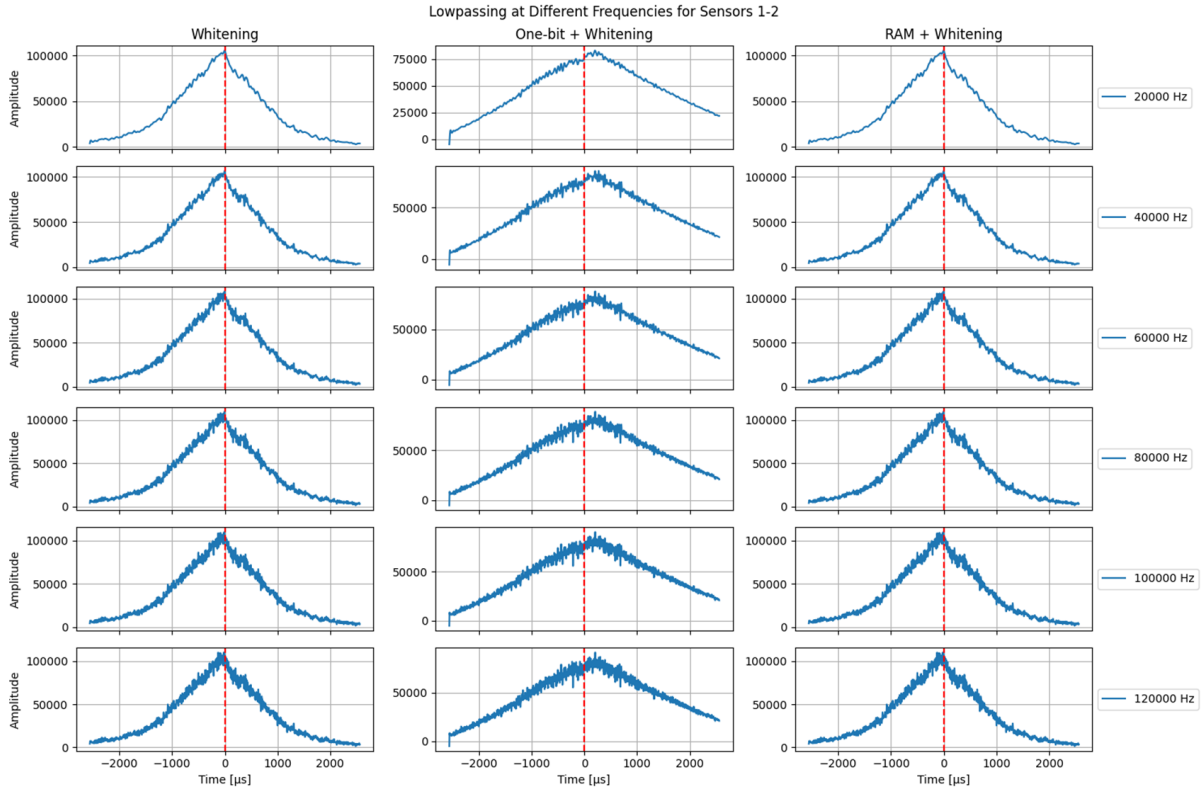

Figure A.11: Lowpassing for sensor pair 1-2.

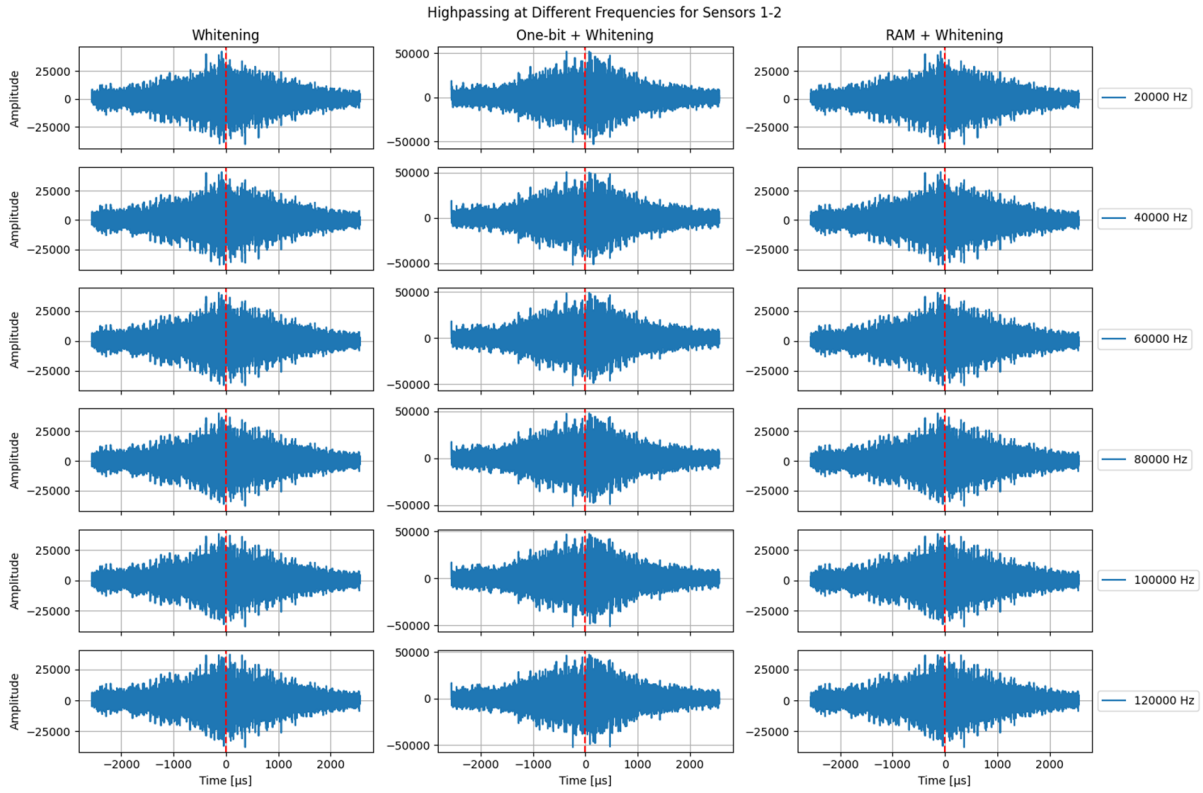

Figure A.12: Highpassing for sensor pair 1-2.

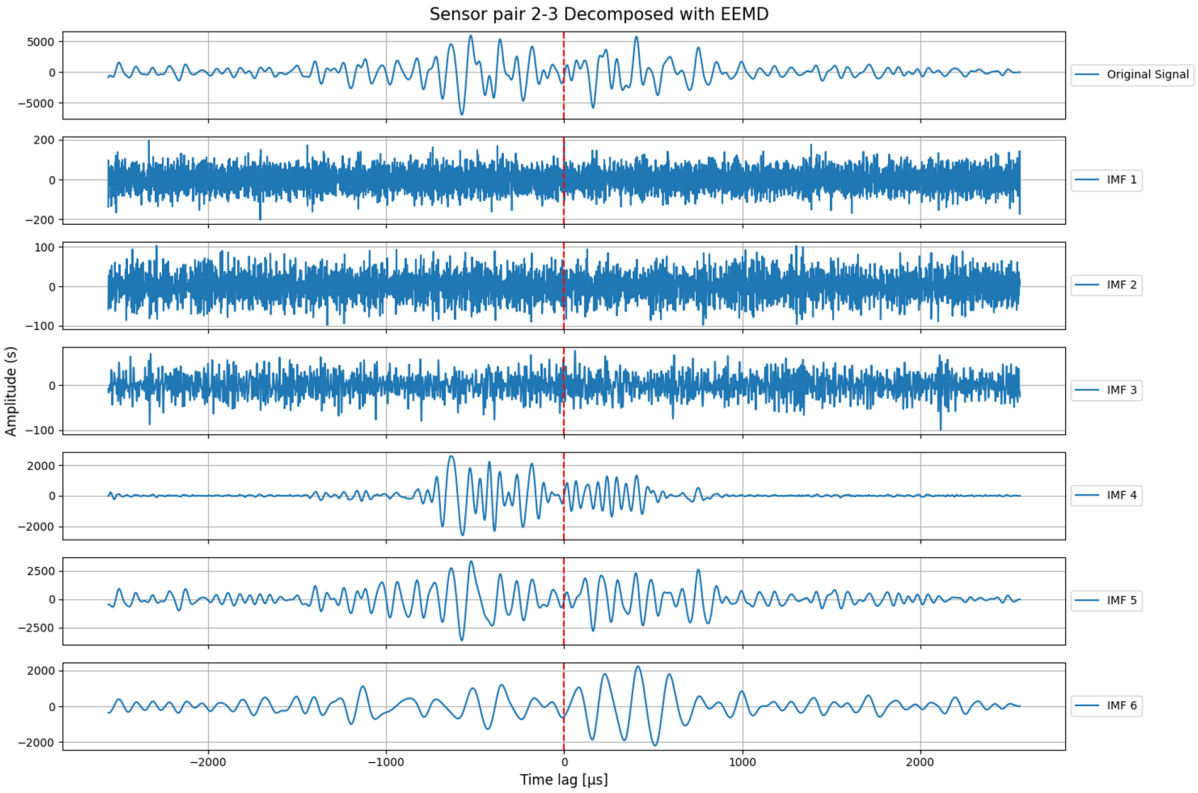

Figure A.13: IMF4 of pair 2-3 (MT)

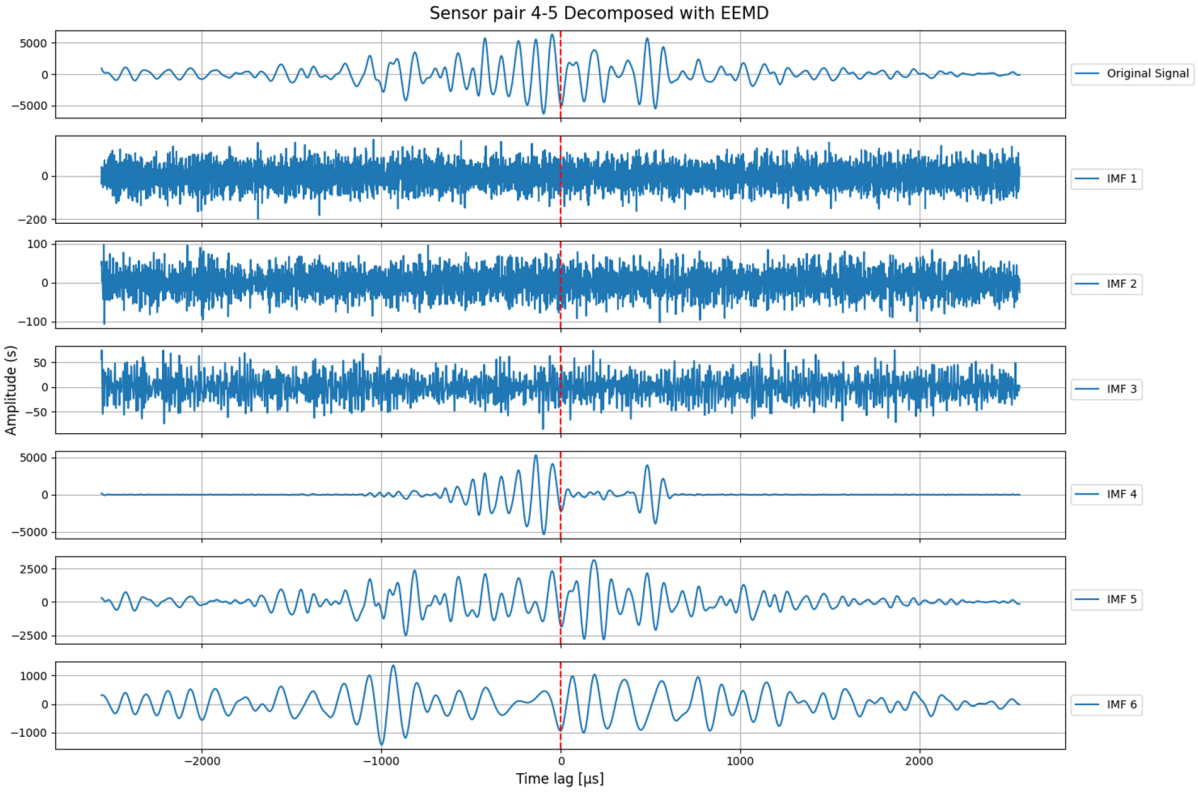

Figure A.14: IMF4 of pair 4-5 (MT)

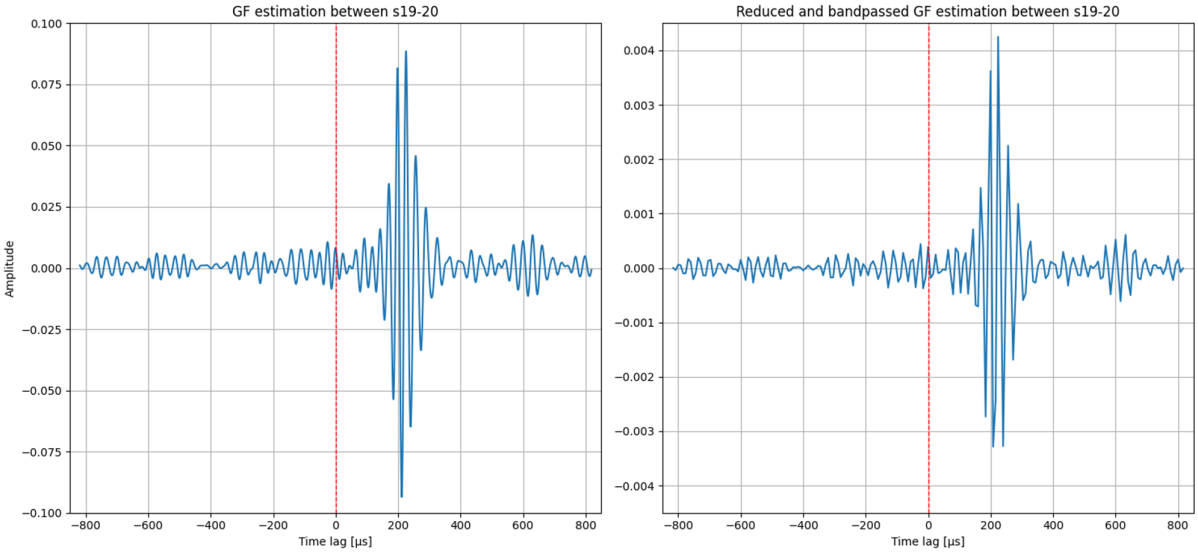

Figure A.15: Green's function estimations for sensor pair 19-20 in the validation dataset for data with sampling frequencies of 2.5 MHz (left) and 125 kHz (right).