國立臺灣大學海洋研究所

博士論文

Institute of Oceanography

National Taiwan University

Doctoral Dissertation

探討周遭噪訊交互相關函數中雜訊的特性以及其應用

On the noise level of the ambient noise cross-correlation

function and its applications

陳映年

Ying-Nien Chen

指導教授：龔源成 博士

喬凌雲 博士

Advisor: YuanCheng Gung, Ph.D.

Ling-Yun Chiao, Ph.D.

中華民國 103 年 11 月

November 2014

國立臺灣大學博士學位論文
口試委員會審定書

探討利用周遭噪訊所得到的
交互相關函數中雜訊的特性以及其應用
On the noise level of the ambient noise
cross-correlation function and its applications

本論文係陳映年君（D98241004）在國立臺灣大學海洋研究所完成之博士學位論文，於民國 103 年 11 月 27 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

_____（簽名）（指導教授）

_____（指導教授）

_____

_____

_____

系主任、所長 _____（簽名）

i

摘要

地震紀錄中的噪訊主要來自於海浪與海床之間的交互作用所產生的微震訊號(microseisms)。近年來利用周遭噪訊法所得到的經驗格林函數已經被廣泛的應用在地震學的研究上，進而大幅增加地殼以及上部地函的解析能力。由於經驗格林函數的訊噪比可以簡單藉由增加地震連續紀錄的長度來提升，因此對於該函數中雜訊的特性至今尚未有量化的描述或系統性的探討。本研究首先提出測量經驗格林函數中「初始雜訊強度」的流程，藉此我們可以客觀的量化經驗格林函數的資料品質。理論上，「初始雜訊強度」與噪訊源的數量分佈息息相關，而經驗格林函數的強度則與噪訊數量以及其產生機制有關，因此，藉由比較經驗格林函數的強度的以及相對應的「雜訊」特性，我們可以進一步了解噪訊源產生的機制。分析了台灣以及韓國的寬頻地震站的資料之後，我們發現台灣短周期(3~5 秒)的噪訊強度主要與周遭海域的水深有關。此外，多數研究認為 7~9 秒的噪訊主要來自於遠方長浪引起的長周期的次級微震(long period secondary microseisms)，而本研究提出證據指出台灣以及韓國的噪訊主要是來自於近岸的海浪。過去研究必須同時仰賴地震資料以及海洋的觀測才能探討噪訊的產生機制，而利用本研究提出的「初始雜訊強度」，我們有能力可以單從地震噪訊的資料來探討噪訊源的特性。
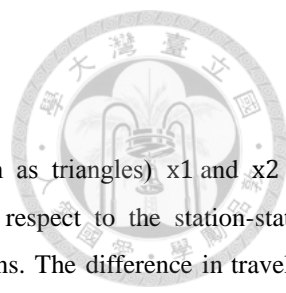
關鍵字: 周遭噪訊、雜訊強度、經驗格林函數

Abstract

Retrieving the Empirical Greens function (*EGF*) between two receivers by cross-correlating continuous records is now a well-recognized technique and the derived *EGFs* have been applied to various fields of seismology. However, little attention has been given to a more quantitative description on the noise behavior of the noise-derived cross-correlation functions (*CCF*), for its signal-to-noise ratio (*SNR*) can be improved easily by increasing the total correlation time. In this thesis, we propose a method to measure the noises within the *CCFs* and demonstrate the relationship between noises and the corresponding sources properties. We evaluate the original noise level (*ONL*) for *CCFs* in Taiwan and Korea. With the measured *ONL*, we can estimate data quality for any portion of the *CCF* in the time domain. Moreover, since the *ONL* is closely related to the noise source population and *EGF's* amplitude is sensitive to the excitation strength, combination of both measurements allows us to put better constraints to the noise sources. Using the approach, we conclude that (1) The dominant microseisms of period 5~10 sec observed in Taiwan and Korea are mostly contributed by Primary microseisms (PM), rather than long period secondary microseisms (LPSM) proposed by previous studies; (2) The high short period secondary microseisms (SPSM) level in Taiwan Strait is mainly caused by the bathymetry effect; (3) The low *ONL* in the SPSM band implies that sources for these dominant signals in *CCFs* are likely confined in the near-coast region; (4) The expected high source population of PM around Taiwan is well demonstrated by the strong *ONL* in the period ~6-9 seconds, although the PM signals are not present in the *CCF* records or the background seismic noises.

Keywords: ambient noise, cross-correlation function, microseisms, original noise level
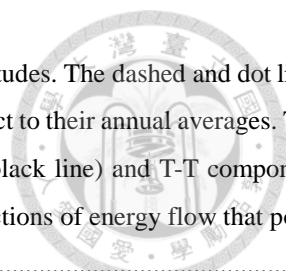
Contents:

List of figures

Chapter 1

Introduction

Microseisms had long been considered as worthless noises in seismic records until 1920s, when it was pointed out that these short-period (<20s) oscillations are closely related to weather phenomena (*Banerji*, 1924). Since then, characteristics of microseisms have been widely explored to study the nature of the corresponding triggering mechanisms [e.g., *Donn*, 1951; 1952; *Donn*, 1966; *Hasselmann*, 1963; *Haubrich et al.*, 1963]. Among these various topics, attention was first focused on the wave type of microseisms. Using particle motion analysis, Lee (1935) first demonstrated that Rayleigh waves are the dominant signals in microseisms, and this was concurred by later studies (e.g., *Rind and Down*, 1979). Results from F-K analysis [*Toksöz and Lacoss*, 1968] indicated that the short period (< 5 sec) microseisms consist of both body waves and higher mode surface waves, and their excitation sources are highly correlated to the low atmospheric pressure area in the weather map.

Many recent works have provided more evidences on the strong correlation between microseisms and ocean climates [e.g., *Bromirski and Duennebier*, 2002; *Bromirski et al.*, 1999; *Gerstoft and Tanimoto*, 2007; *Gerstoft et al.*, 2006], and further suggested that it might be possible to monitor the ocean climate through the microseism analysis [e.g., *Bromirski*, 2001; *Bromirski and Duennebier*, 2002; *Bromirski et al.*, 1999; *Cato and Tavener*, 1997; *Okeke and Asor*, 2000].
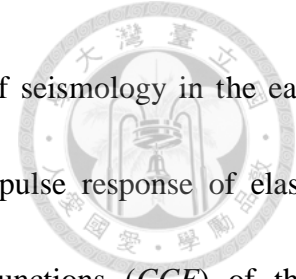
Utilization of microseisms was finally noticed in the community of seismology in the early 21th century. *Shapiro and Campillo* [2004] first showed that the impulse response of elastic waves between two wave sensors resembles the cross correlation functions (*CCF*) of their continuous records, which are mainly composed of microseisms. Empirical Greens functions (*EGFs*) thus retrieved are dominated by fundamental-mode surface waves (e.g., Shapiro et al., 2005) and have been quickly applied to seismic studies. The earthquake-free surface waves are ideal data for seismic tomography, and such approach has made revolutionary progress in resolving the crust and uppermost mantle structure worldwide [e.g., *Harmon et al.*, 2013; *Lin et al.*, 2009; *Poli et al.*, 2012; *Shapiro et al.*, 2005; *Yao et al.*, 2006; *Zhan et al.*, 2010]. The robust coda train trailing the major signal in *EGFs* has been applied to the detection of temporal perturbations in crustal elastic properties [e.g., *Brenguier et al.*, 2008; *Yu and Hung*, 2012], and the symmetry between causal and acausal *CCFs* could be used to correct the potential instrument errors, such as polarity reversal or time drift of internal clock [*Lukac et al.*, 2009b; *Stehly et al.*, 2007; *You et al.*, 2010] .

It is not surprising that the noise-derived *EGFs* are mostly used on the tomographic applications for the subsurface structure. On the other hand, progresses are relatively slow for the non-tomographic applications, and one of such issues is about the noise content in the noise-derived *EGFs*. It's well known the *CCFs* are not perfect Green functions, but details regarding to the unwanted noises in *CCFs* have never been studied. This is because that, for the tomographic applications, the noise level of *EGF*s can be easily reduced by increasing the cross-correlation time.

The main purpose of this thesis is to investigate the noises in the noise-derived *EGFs*.

To begin with, in Chapter 2 we present a theoretical derivation to demonstrate the explicit relationship between the noise-derived *EGF* and the Greens Function. In Chapter 2.2~2.4 we detail the relationship between CCF and the corresponding sensitivity zone of noise sources, and discuss the widely concerned issue: the precision of the phase velocity measurement using the noise-derived EGFs. The main formulas used this thesis are presented in Chapter 2.5, in which we detail the behavior of the noise of *CCF* and notice that the noise level is primarily sensitive to the source population.

Next, in Chapter 3 we investigate the spatial-temporal properties of SPSM around Taiwan[*Chen et al.*, 2011]. We conclude that the noise excitation in this frequency band (3-5 seconds) is highly correlated with the water depth of the surrounding ocean, and its signature is clearly shown in the resulting *CCFs*. We also observe that the temporal variation of *CCF* amplitudes is tied closely to the monsoon property.

In chapter 4, we propose a procedure to evaluate the noise level of *CCFs* and apply it to the *CCFs* derived in Taiwan and Korea. The obtained evolution of noise level follows the theoretical prediction nicely, and the resulting noise levels remain nearly constant for all station pairs analyzed. We discuss its implications and utilize the results to constraint the properties of SPSM, primary microseism (PM) and distant primary microseism (DPM).

Finally, we briefly summarize the key results of this thesis in Chapter 5.

Chapter 2
Theoretical Background

2.1. Introduction

Over the past 20 years, the theoretical bases to extract Greens Function (*GF*) from noise records have been developed by multiple lines of physics, including fluid dynamic physics [e.g., *Godin*, 1997; 2006; 2007], acoustic and thermal physics [e.g., *Derode et al.*, 2003; *Lobkis and Weaver*, 2001; *Weaver and Lobkis*, 2001], time-reversal theory[e.g., *Cassereau and Fink*, 1992; *Larmat et al.*, 2010], fluctuation theory [e.g., *Larose et al.*, 2008; *Weaver and Lobkis*, 2005], and normal mode analysis [e,g., *Tanimoto*, 2008; *Weaver and Lobkis*, 2005]. While these derivations vary from each other, the existence of a diffusive wave field is an important prerequisite for all of them.

In seismology, the continuous ambient noises can be generally considered as records from a quasi-diffusive wave field, since they are mainly generated by the everlasting and complex interactions between ocean waves and solid earth [*Shapiro and Campillo*, 2004]. As mentioned in Chapter one, microseisms induced from such interactions are dominated by fundamental mode surface waves, it's then not surprising that the noise-derived *CCFs* are also dominated by fundamental mode surface waves.

In a fully diffusive wave field, since the cross-correlation functions (*CCF*) of two records triggered by unrelated sources/modes cancel out with each other, the remaining *CCF* consists of only those from the same sources/modes. Naturally, the ambient seismic wave field is not perfectly diffusive, and the differences between the noise-derived *CCF* and the ideal *GF* are expected. In other words, there exist unwanted noises within the noise-derived *CCFs*, and we may further separate the noise-derived *CCF* into two components, the "Signals" and the "Noises", in which the "Signals" are related to the products from common sources and the "Noises" to the products from irreverent sources, respectively.

Because the noises in *CCFs* are generally decreasing with the correlation time, the accumulated *CCFs* in seismic study are thus commonly referred as empirical Greens Function (*EGF*), and it is assumed that *EGF* differs only by an amplitude factor from *GF*. However, results from ray-based derivation argue that there are phase discrepancies between *EGF* and *GF* under certain circumstances [e.g., *Boschi et al.*, 2013; *Tsai*, 2009]. In the following, we first clarify this issue by presenting the explicit relationship between *EGF* and *GF* using a ray-theory derivation (Section 2.2), and we conclude that, indeed, there are phase discrepancies between *EGF* and GF, but the difference can be essentially ignored when interstation distance is larger than three times of the wavelength in concern. The inference is consistent with the commonly used "three wavelengths" data selection criteria in the surface wave dispersion measurement.

The "three wavelengths" limitation for dispersion measurement can be relieved when the
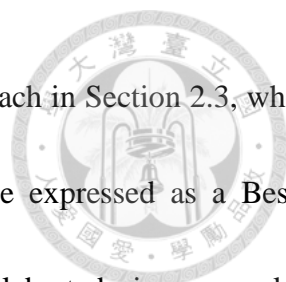
measurement is done in the frequency domain. We demonstrate this approach in Section 2.3, where we show that the real part of an azimuthal averaged coherency can be expressed as a Bessel Function, an important property first pointed out by Aki (1957) in his celebrated pioneer work in seismic interference using spatial auto-correlation. Such formulation provides an alternative method to measure the dispersion and attention of surface waves in the frequency domain without the "three wavelengths" limitations.

In about just one decade, the noise-based tomography has become a widely used technique and made striking progress in resolving the crust and uppermost mantle structure. In contrast, studies on the source mechanisms of the ambient seismic noises are much left behind, and our knowledge on this new field is still at the infant stage.

It is shown that the approximate spatial and temporal energy properties of the noise sources can be obtained by analyzing the noise-derived *CCFs*. Characteristics of the resulting source properties, in particular, about the excitation strength and energy spectrum, have been related to various mechanisms such as the bathymetry effects, types and heights of ocean waves, ocean currents and vigorous atmospheric perturbations…etc. [e.g., *Chen et al.*, 2011; *Gu et al.*, 2007; *Kraeva et al.*, 2009; *Stehly et al.*, 2006]. However, a more in-depth analysis between *CCFs* and the corresponding source characteristics seems to be lacking. One of the major topics of this thesis is to further explore this subject, aiming to better understand the nature of the ambient seismic noises using the noise-derived *CCFs*.
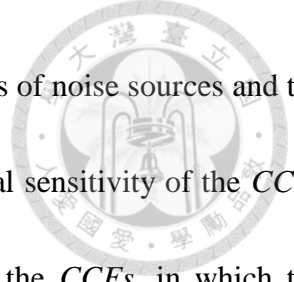
In this regard, we need to establish the link between various properties of noise sources and the observable characteristics of the resulting *CCFs*. We first assess the spatial sensitivity of the *CCFs* to given far-field sources by examining their effective contributions to the *CCFs*, in which the source energy is represented as the product of the excitation strength and source density (Section 2.4). Finally, we modify the expression of *CCFs* using a mode-based representation, and demonstrate the explicit relationship between the noise sources and the behavior of the *EGFs* and the noise level in the *CCFs* (Section 2.5).

Here, we first briefly introduce the basic formulations with a simple example, in which a far-field point source with a deterministic wave of frequency ω and two stations are considered, as shown in Figure 2-1. The displacement recorded by station x_i can be expressed as $U(x_i, t) = A(\omega) \cdot cos[\omega(t - t_i) + \varphi]$, where t_i is the travel time between source and station, $A(\omega)$ is the source amplitude, and φ is an arbitrary initial phase. In the case of a distant single source, *CCF* for records with length T $(T \gg 1/\omega)$ at stations x_1 and x_2 is written as follows

$$
\begin{aligned}
CCF_{x_1 x_2}(\tau, \omega) &= \frac{A(\omega)^2}{2T} \int_{-T}^{T} cos[\omega(t - t_1) + \varphi] cos[\omega(\tau + t - t_2) + \varphi] \, dt \\
&= \frac{A(\omega)^2}{2} \{cos[\omega(\tau - \Delta t)] - \frac{sin(2\omega T)}{2\omega T} cos[\omega(\tau - \Delta t) + 2\varphi]\} \\
&\approx \frac{A(\omega)^2}{2} cos[\omega(\tau - \Delta t)] \\
&= \frac{A(\omega)^2}{2} cos[\omega\tau - \phi],
\end{aligned}
\tag{2.1}
$$

where $\Delta t \equiv t_2 - t_1$ is the time delay between signals arriving at two stations, and ϕ is the corresponding phase shift. Notice that the information about the initial phase φ is washed away

through cross-correlating.

In a media with homogeneous velocity (c), Δt in eq. (2.1) can be expressed by $r \cdot cos\,(\theta)/c$, where θ and r represent the azimuth of source to the station-station line and the inter-station distance, respectively.

Next, we consider the case with continuous far-field sources coming from all azimuths, in which the strength of sources can be express as $A(\theta, \omega)$. With this, the displacement recorded by station x_i, can be rewritten as $U(x_i, t) = \sum_{\theta=0}^{2\pi} A(\theta, \omega) cos[\omega(t - t_i(\theta, \omega)) + \varphi(\theta, \omega)]$, where $t_i(\theta, \omega)$ is the delay time between source and stations, and $\varphi(\theta, \omega)$ is an arbitrary initial phase. Accordingly, *CCF* can be expressed as a simple integral form:

$$CCF_{x_1 x_2}(\tau, \omega) = \int_0^{2\pi} A(\theta, \omega)^2 \cdot cos[\omega(\tau - rcos(\theta)/c\,)]\,d\theta + Noise(\omega)$$

$$= \int_0^{2\pi} A(\theta, \omega)^2 \cdot cos[\omega\tau - \phi(\theta)]\,d\theta + Noise(\omega), \qquad (2.2)$$

where $\phi(\theta) \equiv r\omega cos(\theta)/c$, an azimuthal dependent phase shift in *CCF*.

As mentioned earlier, in a noise-derived *CCF*, the first term in eq. (2.2) represents the "Signal" and the second term represents the "Noise", respectively. The "Signal" is a summation of the "coherent source" product, and the "Noise" is a result of "unrelated source" product, as mentioned earlier. In the following, we will detail the phase and amplitude properties for the "Signal" (Section 2.2~2.4) and the behavior of the "Noise" (Section 2.5).
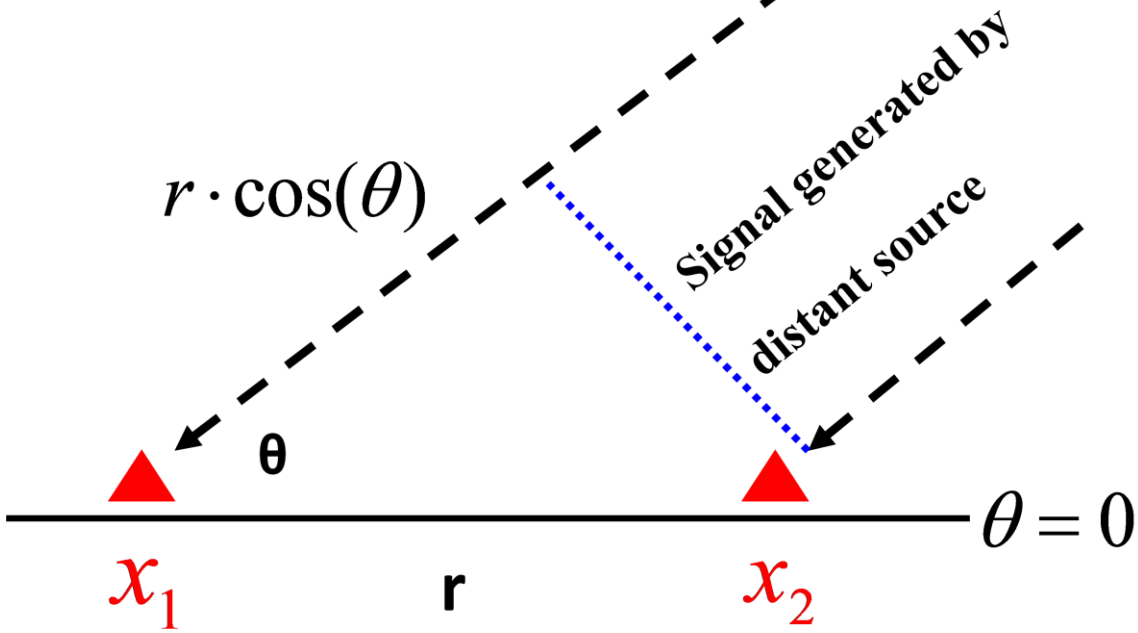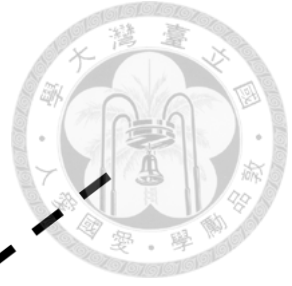
Figure 2-1: Schematic of a "far-field" source geometry considered. Stations (shown as triangles) x_1 and x_2 are separated by distance r. θ represents the azimuth of the far-field source with respect to the station-station direction, and dash lines indicate the traveling paths from the source to the stations. The difference in traveling distances to two stations, $r \cdot \cos(\theta)$ is also shown.

2.2. Explicit Relationship between *CCF* and the *GF*

In this section, we present a conceptual framework linking the *CCF* and the *GF*. To begin with, we first demonstrate that *CCFs* can be considered as an interference pattern resulting from all correlated distant sources surrounding the station pair. Ignoring the "Noise" in eq. (2-2), the obtained *CCF* is simply a summation of all coherent source products of azimuth-dependent phase delays. As shown in Figure 2.2, since the phase delays fluctuate quickly for sources away from the station-station line, contributions from these off-line sources cancel out with each other due to destructive interference, and the remaining *CCF* is dominated by sources along the station-station direction. With this simple illustration, we reach the same conclusion obtained by using the stationary phase analysis [*Snieder*, 2004]. Assuming the strength of distant sources within an effective zone is a function of frequency and azimuth, denoted by $\overline{A(\theta, \omega)}^2$, we may approximate the resulting *CCF* to *GF*, namely, $CCF \sim \overline{A(\theta, \omega)}^2 \cdot G(\omega)$.

In what follows, we first detail the dispersion measurement in the frequency domain, and then demonstrate the relationship between the *CCF* and the *GF* of the fundamental mode surface waves. Our treatment follows [*Tsai and Moschetti*, 2010] ,[*Tsai*, 2009; 2010; 2011] and [*Boschi et al.*, 2013].
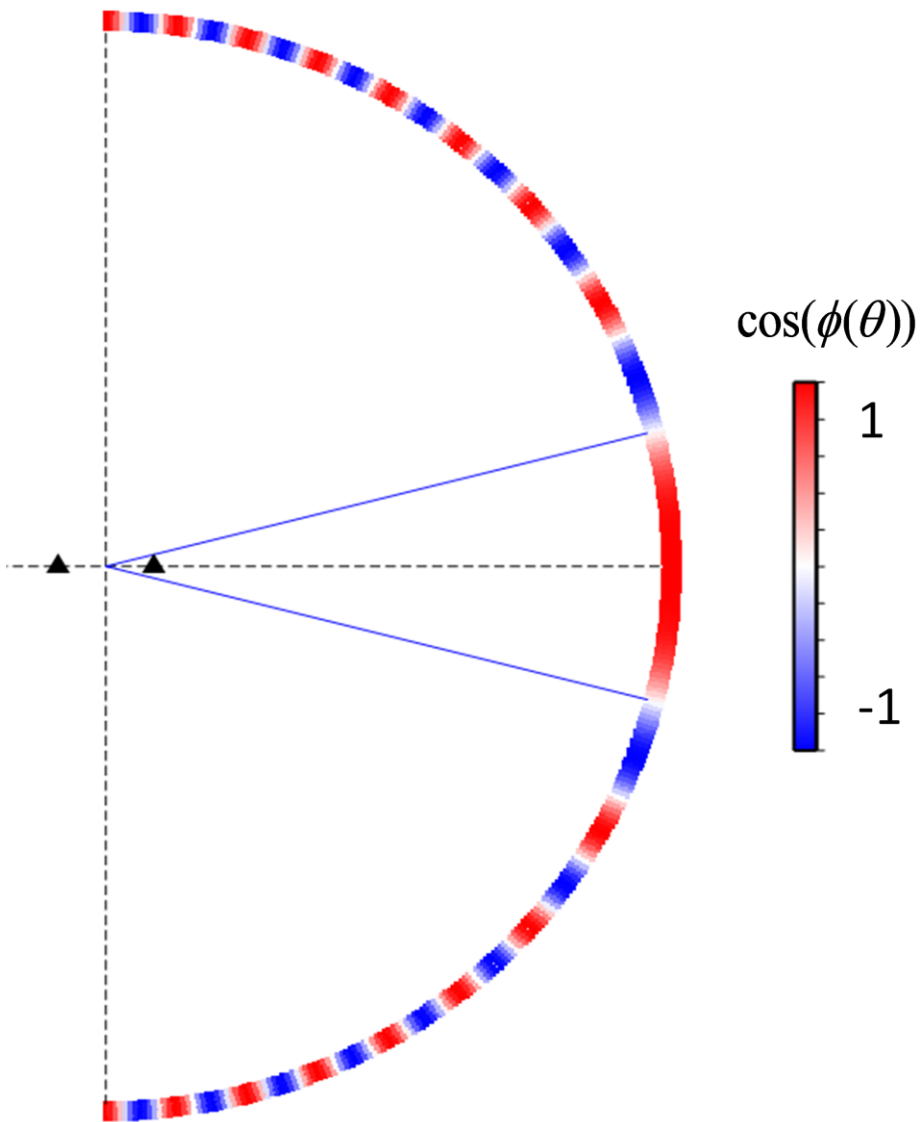
Figure 2-2: Demonstration of interference pattern in the resulting *EGF*. The stations are shown by black triangles, and the azimuthal-dependent phase shift $\cos(\phi(\theta))$ (equation (2.2)) is shown by scaled colors.

2.2.1. Retrieving Phase Velocity In Frequency Domain - Uniform Source Distribution

Aki (1957) first showed that the velocity between two stations could be estimated in the frequency domain through the spatial auto-correlation (SPAC), which is defined as $\phi(r) = \frac{1}{2T}\int_{-T}^{T} U(x_1, t)U(x_2, t)dt$, where U is displacement and r is the inter-station distance between x_1 and x_2. Apparently, the SPAC is simply the *CCF* at zero correlation lag time. Aki also noted that the azimuthal averaged SPAC in a non-uniformly distributed noise source condition, denoted by $\bar{\phi}(r) \equiv \frac{1}{2\pi}\int_0^{2\pi} \phi(|\xi| = r)dr$, is equivalent to any individual $\phi(r)$ obtained from an uniformly distributed one. Namely, the resulting SPAC is not influenced by the source inhomogeneity. Thus, the velocity (c) for a given frequency ω, is revealed through the relationship $\bar{\rho}(r, \omega) \equiv \frac{\bar{\phi}(r,\omega)}{\phi(0,\omega)} = J_0(\frac{r\omega}{c})$, where $\bar{\phi}(r, \omega)$ is an azimuthal averaged SPAC, and J_0 is the zeroth order Bessel function of the first kind. Since then, similar results were also derived by other studies, in which $\bar{\rho}(r, \omega)$ was commonly referred as averaged coherency [e.g., *Asten*, 2006; *Henstridge*, 1979; *Nakahara*, 2006].

A detailed derivation for the condition with an arbitrary noise source distribution will be given in Section 2.3. Here we first introduce Aki's basic idea and clarify the relation between the coherency and the *CCF* in the case with uniform source distribution. In such case, the source strength $A(\theta, \omega)$ can be simplified by A(ω). Ignoring the unwanted noise part, equation (2.2) is

now written as

$$CCF_{x_1 x_2}(\tau, \omega) = \int_0^{2\pi} A(\omega)^2 \cdot \cos[\omega(\tau - r\cos(\theta)/c)] \, d\theta$$

$$= 2\pi A(\omega)^2 \cdot J_0\left(\frac{r\omega}{c}\right) \cdot \cos(\omega\tau), \tag{2.3}$$

Note that $\phi(r, \omega)$ is equivalent to $CCF_{x_1 x_2}(0, \omega)$, and $\phi(0, \omega)$ can be related to *CCF* obtained

from a single station:

$$CCF_{x_1 x_1}(\tau, \omega) = CCF_{x_2 x_2}(\tau, \omega) = \int_0^{2\pi} A(\omega)^2 \cdot \cos(\omega\tau) \, d\theta \,]$$

$$= 2\pi A(\omega)^2 \cdot \cos(\omega\tau). \tag{2.4}$$

To reproduce the result proposed by Aki (1957), the *CCF* should be transformed into the

coherency by applying a spectrum-whitening normalization to the original data prior to cross

correlating. The whitened *CCF* can be written as

$$\widehat{CCF}(\omega, r) = \frac{CCF_{x_1 x_2}(\omega)}{\sqrt{CCF_{x_1 x_1}(\omega) \cdot CCF_{x_2 x_2}(\omega)}}. \tag{2.5}$$

Since equation (2.3) and (2.4) have the same phase, this division can be done in the time domain;

otherwise the division should be employed in the frequency domain. Substituting eq. (2.3) and eq.

(2.4) into eq. (2.5)

$$\widehat{CCF}(r, \omega, \tau) = \rho(\omega, r, \tau) = \frac{2\pi A(\omega)^2 \cdot J_0\left(\frac{r\omega}{c}\right)}{2\pi A(\omega)^2} \cos(\omega\tau) = J_0\left(\frac{r\omega}{c}\right) \cos(\omega\tau) \tag{2.6}$$

With an isotropic background energy flux, we have successfully established the relationship

between the $\widehat{CCF}$ and SPAC. Notice that in the frequency domain, the real part of $\widehat{CCF}$ is

equivalent to $\rho(r, \omega) \equiv \frac{\phi(r, \omega)}{\phi(0, \omega)} = J_0\left(\frac{r\omega}{c}\right)$. In addition, the coherency is a particular value of the

band-passed $\widehat{CCF}$ at zero lag time. Since the time domain cross-correlation is equal to the
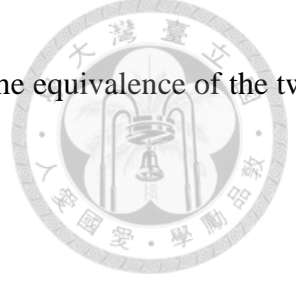
frequency domain multiplication, this is an alternative way to understand the equivalence of the two theories.

2.2.2. Phase Discrepancy Between The *CCF* And The *GF*

In the above derivation, while the velocities can be estimated in the frequency domain, the relationship between *CCF* and *GF* remains unclear. To this end, we first rewrite eq. (2.3) into the complex domain as

$$CCF_{x_1x_2}(\tau, \omega) = \Re[e^{i\omega\tau} \int_0^{2\pi} A(\omega)^2 \cdot e^{-i\omega(r\cos(\theta)/c)}d\theta \,]$$

$$= \Re[e^{i\omega\tau} \cdot 2\pi A(\omega)^2 \cdot (J_0\left(\frac{r\omega}{c}\right) - iH_0(\frac{r\omega}{c}))], \qquad (2.7)$$

where $\Re(\cdots)$ is the real part of its argument and H_0 is the zeroth order Struve function of the first kind. Now the exact travel time $(\tau(\omega))$ can be related to the pick of the band-passed *CCF* by

$$\omega\tau(\omega) + 2N\pi = \phi[J_0\left(\frac{r\omega}{c}\right) - iH_0(\frac{r\omega}{c})], \qquad (2.8)$$

where ϕ is the phase of the complex signal and $2N\pi$ represents the cycle skip phenomena in the phase velocity measurement. Next, using the high frequency approximation $(\frac{r\omega}{c} \gg 1)$, $J_0\left(\frac{r\omega}{c}\right)$ and $H_0\left(\frac{r\omega}{c}\right)$ in eq. (2.7) can be approximated as

$$J_0\left(\frac{r\omega}{c}\right) \approx \sqrt{\frac{2c}{r\omega\pi}}\cos\left(\frac{r\omega}{c} - \frac{\pi}{4}\right), \qquad (2.9)$$

$$H_0\left(\frac{r\omega}{c}\right) \approx Y_0\left(\frac{r\omega}{c}\right) = \sqrt{\frac{2c}{r\omega\pi}}\sin\left(\frac{r\omega}{c} - \frac{\pi}{4}\right), \qquad (2.10)$$

where Y_0 is the zero-order Bessel function of the second kind. Note that the geometrical spreading

14
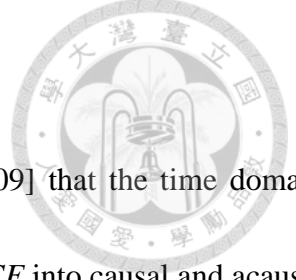
effect $(1/\sqrt{r})$ is included in such representation.

As indicated by several authors [e.g., *Boschi et al.*, 2013; *Tsai*, 2009] that the time domain

relationship between the *CCF* and the *GF* can be verified by separating *CCF* into causal and acausal

parts. Next, we present a similar derivation by taking the causal part of *CCF* as an example.

Using a positive windowing function, defined as $\mathcal{W}(\tau) = \begin{cases} 1, \tau \geq 0 \\ 0, \tau < 0 \end{cases}$, the causal *CCF* is given

by

$$CCF^p(\tau, \omega) = \mathcal{W}(\tau) \cdot CCF(\tau, \omega).$$

$$= \Re[e^{i\omega\tau} \int_0^{\pi/2} A(\omega)^2 \cdot e^{-i\omega(r\cos(\theta)/c)} d\theta\,]$$

$$= A(\omega)^2 \cdot \Re\left[e^{i\omega\tau} \cdot \frac{\pi}{2} \cdot \left(J_0\left(\frac{r\omega}{c}\right) - iH_0\left(\frac{r\omega}{c}\right)\right)\right] \tag{2.11}$$

Note that the source integration over azimuth is taken from 0 to $\pi/2$, as only the causal part of

CCF is considered. Substituting eq. (2.9) and eq. (2.10) into eq. (2.11) gives

$$CCF^p(\tau, \omega) \approx A(\omega)^2 \cdot \Re\left[e^{i\omega\tau} \cdot \frac{\pi}{2} \cdot \left(J_0\left(\frac{r\omega}{c}\right) - iY_0\left(\frac{r\omega}{c}\right)\right)\right]$$

$$= A(\omega)^2 \cdot \sqrt{\frac{c\pi}{2r\omega}} \Re(e^{i(\omega\tau - \frac{r\omega}{c} + \frac{\pi}{4})}) \tag{2.12}$$

Similarly, calculating the $\widehat{CCF}$ for the causal part yields

$$\widehat{CCF}^p(\omega, r, \tau) = \sqrt{\frac{c}{8r\omega\pi}} \Re\left(e^{i\left(\tau\omega - \frac{r\omega}{c} + \frac{\pi}{4}\right)}\right) = \sqrt{\frac{c}{8r\omega\pi}} \cdot \cos\left(\tau\omega - \frac{r\omega}{c} + \frac{\pi}{4}\right). \tag{2.13}$$

It is worth mentioning that there is a phase-shift $\pi/2$ between the *CCF* and the *GF* given by

Dahlen and Tromp [1998]. Such discrepancy implies that a time derivative of the *CCF* agrees with

the *GF (CCF'(t)~ − G(t) + G(−t))*, and further confirms the conclusions those derived from the

other approach [*Lobkis and Weaver*, 2001; *Snieder*, 2004; *Tsai*, 2010; *Weaver and Lobkis*, 2002].

Given the above representation (Eq. 2.10), the phase velocities for any given period T can be estimated in the time domain from CCF ($c(T) = \frac{r}{\tau + 8/T}$).

Next, we discuss the potential errors in the dispersion measurement arising from the aforementioned high frequency approximation. In Figure 2-3a, we compare the travel time estimates using the exact Bessel function (equation 2.8) and its approximation (equation 2.13). While errors are unavoidable in both approaches, the exact phase velocity can be evaluated by comparing the measurements obtained from Eq. (2.8) and Eq. (2.13), as shown in Fig. 2-3b. Our results also demonstrate that the discrepancies in two measurements are essentially negligible (<1%) as the ratio between traveling distance and wavelength ($\frac{r}{\lambda}$) is larger than 3, and this explains why the "three-wavelength" data selection criterion is commonly used in the dispersion measurement of surface waves (e.g., *Yao et al.* [2006].)

In the far-field source assumption, we may treat the formulation from the point of view of delay time. Since $\Delta t = r \cdot \frac{cos(\theta)}{c}$, $\theta(\Delta t) = cos^{-1}(c\Delta t/r)$ and $\frac{d\theta(\Delta t)}{d(\Delta t)} = -\frac{c/r}{\sqrt{1-\left(c \cdot \frac{\Delta t}{r}\right)^2}} = -\frac{c/r}{sin(\theta)}$.

Using the approach of Tsai (2009), an alternative expression for the equation (2.3) can be written as

$$CCF_{x_1 x_2}(\tau, \omega) = \int A(\omega)^2 \left|\frac{d\theta(\Delta t)}{d(\Delta t)}\right| \cdot cos[\omega(\tau - \Delta t)] \, d(\Delta t)$$

$$= \int A(\omega)^2 \cdot \frac{c/r}{sin(\theta(\Delta t))} \cdot cos[\omega(\tau - \Delta t)] \, d(\Delta t). \tag{2.14}$$

It is clear that, in terms of delay time Δt, the contribution to CCF is strongly dependent on the source azimuth. It indicates that the CCF is primarily contributed by sources along the

station-station line, consistent with the conclusion shown in Figure 2-2.
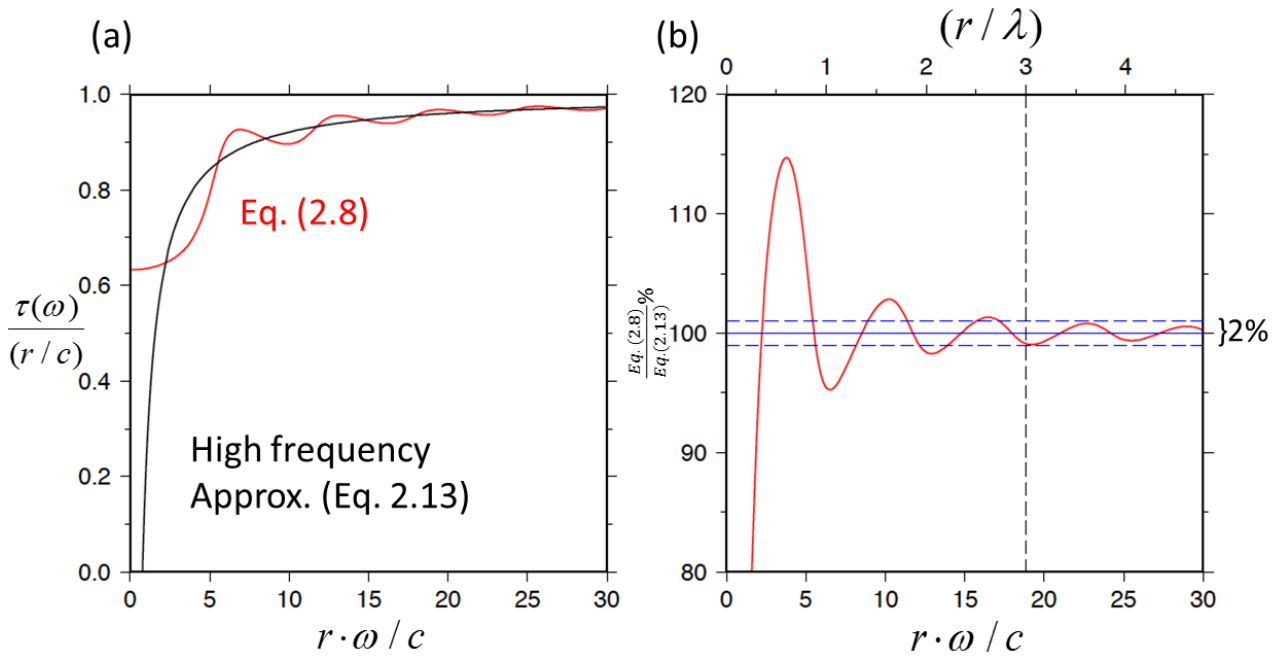
Figure 2-3: Comparisons of the travel time measurement obtained from formula with and without a high frequency approximation. (a) Modified from *Tsai* [2009]. The estimated travel time $\tau(\omega)$ normalized by the true travel time (r/c) between stations as a function of frequency ω.   (b) The ratio of the travel time measurements obtained from Eq. (2.8) and Eq. (2.13) as a function of frequency and the corresponding wavelength r/λ (top x-axis).

2.3 On The Amplitude Of The *EGF*

While the phase of *CCFs* can be easily addressed by assuming a far-field homogenous source distribution, the expression for *CCF* amplitudes is much more complicated. Apparently, *CCF* amplitudes are affected by various factors, such as the inter-station distance (Eq. 2.13), source distribution and the intrinsic attenuation of the structure.

Tsai [2011] presented the proper coherency expressions for cases with different noise source distributions, and the results indicate that, even under a homogeneous energy flux condition, the expressions for cases with a "Far-Field" and a "Far-Field + Near-Field" source distribution are quite different.

On the other hand, recent works on *CCF* amplitudes [*Prieto et al.*, 2009; *Prieto et al.*, 2011; *Tsai*, 2011] showed that the attenuation coefficient $\alpha(\omega)$ can be obtained by modifying equation (2.6) to

$$\Re[\rho(\omega, r)] = e^{-r\alpha(\omega)} J_0\left(\frac{r\omega}{c}\right). \tag{2.15}$$

The attenuation coefficient $\alpha(\omega)$ is related to a path averaged quality factor Q ($\alpha(\omega) \equiv \omega/(2UQ)$, where U is group velocity). Note that such expression is only valid for the cases with both Near-Field and Far-Field noise sources. For instance, they could be applied to arrays sitting by the islands or near the coastal area.

For an array located far away from the coast, a pure Far-Field source condition can be assumed,

and we may express the equation (2.6) as

$$\Re[\rho(\omega, r)] = \frac{J_0(\frac{r\omega}{c})}{I_0(\alpha(\omega)r)}, \tag{2.16}$$

where I_0 is a modified zero-older Bessel function of the first kind [*Tsai*, 2011].

Note that the equation (2.15) and (2.16) are only appropriate for a homogeneous energy distribution. As mentioned in Section 2.2.1, Aki [1957] pointed out that the azimuthal averaged coherency $\bar{\rho}(r, \omega) \equiv \frac{\overline{\phi(r,\omega)}}{\phi(0,\omega)}$ is equivalent to that retrieved in an isotropic source condition. In the following, we detail the derivations.

According to the results given in previous section that *CCFs* are dominated by source along the station-station direction, any *CCF* obtained from an environment with arbitrary source distribution can be approximately expressed by adding an azimuthal-dependent source density term $\rho_s(\vartheta)$, where ϑ is the azimuth for the station-station direction. Accordingly, the *CCF* in Eq. (2.4) for station pairs along azimuth ϑ can be approximated as

$$CCF_{x1x2}(\tau, \omega) \approx \rho_s(\vartheta)A(\omega)^2 J_0(\frac{r\omega}{c}) \cos(\omega\tau). \tag{2.17}$$

Using an averaged source density defined as $\bar{\rho}_s \equiv \frac{\int_0^{2\pi} \rho_s(\vartheta)d\vartheta}{2\pi}$, we can rewrite equation (2.4) as

$$CCF_{x1x1}(\tau, \omega) = CCF_{x2x2}(\tau, \omega) \approx \bar{\rho}_s A(\omega)^2 \cdot \cos(\omega\tau). \tag{2.18}$$

Finally, the real part of an azimuthal averaged coherency is given by

$$\Re[\bar{\rho}(\omega, r)] \equiv \frac{\overline{\phi(r, \omega)}}{\phi(0, \omega)} = \frac{\left(\int_0^{2\pi} \rho_s(\vartheta)\, d\vartheta/2\pi\right)\left(A(\omega)^2 J_0\left(\frac{r\omega}{c}\right)\right)}{\bar{\rho}_s A(\omega)^2}$$

$$= J_0\left(\frac{r\omega}{c}\right) \tag{2.19}$$
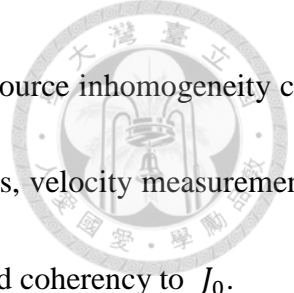
It agrees with the result shown in equation (2.6). Note that the effects of source inhomogeneity can be removed by taking an azimuthal (spatial) average of the coherency, thus, velocity measurements can be made in the frequency domain by fitting the real part of the observed coherency to J_0.

While this approach have been applied to measurements for velocity and attenuation (*Prieto, Lawrence and Beroza*, 2009; *Prieto et al.*, 2011; *Weemstra et al.*, 2013), the application is limited to area with dense seismic network to guarantee that the azimuthal averaged coherence can be estimated properly. Otherwise, one can only measure velocities for some particular frequencies at $J_0\left(\frac{r\omega}{c}\right)=0$, for those are insensitive to the source effect.

2.4 The Sensitivity Zone Of *CCF*

As demonstrated in Section 2.2 and Fig 2.2, *CCF* is dominated by sources within a narrow

sensitivity zone along the station-station line. Similarly, here we evaluate the effective sensitivity

zone of coherency.

In Section 2.2, we have presented two major properties of *CCF* resulting from using a far-field

source condition: (1) Due to violent interference, *CCF* is dominated by sources within a narrow

zone along the station-station direction. (2) The accuracy of velocity measurement is highly

dependent on the high frequency approximation ($r\omega/c$). In the following, we simulate the effective

sensitivity zone of coherency, and we conclude that the above two properties of *CCF* are actually

closely related.

For simplicity, we consider a Far-Field source distribution with strength as a function of

azimuth. Since the coherency is mostly sensitive to the source along the station-station direction

(Sec. 2.2), we evaluate the perturbations of coherency amplitude by increasing the source strength

of varying zone sizes centering at azimuth $\theta = 0$. Specifically, the noise source strength for the

zone size of interest θ_o, is given by $A_s(\theta)|_{\theta_o} \equiv \begin{cases} 2, & \theta_o \geq \theta \geq 0 \\ 1, & \theta > \theta_o \end{cases}$ (see Fig. 2-4), thus, the

corresponding coherency can be expressed as

$$\rho(\omega, \theta_o, r) \sim \frac{1}{A_s(\theta)|_{\theta_o}} \Re\left[e^{i\omega\tau} \int_0^\pi A_s(\theta)|_{\theta_o} \cdot e^{-i\omega(r\cos(\theta)/c)} d\theta\right]. \tag{2.20}$$

Here, we take a station pair with inter-station distance 400 km in a homogeneous medium
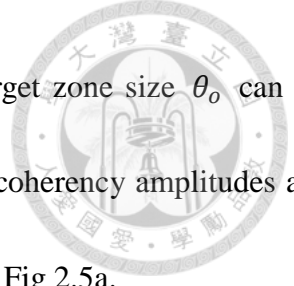
(velocity = 4 km/s) as an example (see Fig. 2-5). Coherency for any target zone size θ_o can be estimated (see Fig. 2-5a) using eq. (2.20). We first note that, indeed, the coherency amplitudes are only sensitive to a narrow zone around the station-station line, as shown in Fig 2.5a.

By taking the gradient of coherency amplitude along the azimuth axis, we may define the effective sensitivity zone (θ_{0E}) for each periods at their first turning point of gradient ($\frac{\delta\rho}{\delta\theta_o}|_{\theta_{0,}=\theta_{0E}} = 0$), as shown in Fig 2.5c.

Instead of using the azimuthal width, we may also describe the effective sensitivity zone in terms of a properly defined delay times, $\Delta t(\theta_0) \equiv \frac{r}{c} - \frac{r\cdot cos(\theta_0)}{c}$, a differential travel time related to the zone size considered, θ_o. Similarly, we show the variations of coherency amplitude w.r.t. $\Delta t(\theta_0)$ in Fig 2.5b, the converted amplitude gradient in Fig 2.5d, and a room-in image in Fig 2.5e. The delay times (Δt_E) related to the effective sensitivity zone widths are also defined in the same manner, which is marked as solid line in Fig 2.5e. Note that there is a simple linear relationship between Δt_E and periods.

In the case we consider, the exact travel time between stations is $T_a = \frac{r}{c} = 100$ sec., and the delay time related to the first beam zone is $\Delta t_E(\theta_{0E}) = T_a(1 - cos(\theta_{0E})) \sim \frac{T}{4}$, corresponding to a $\frac{\pi}{2}$ phase shift (see Fig. 2-5e). The physics behind the effective sensitivity zone is now clear, the phase shift is larger than $\frac{\pi}{2}$ for waves coming from sources beyond azimuth θ_{0E}, as a result, the coherence amplitudes are reduced by the destructive interferences.

Finally, the sensitivity zone size θ_0 can be directly related to the parameter $\frac{r\omega}{c}$, which has

22
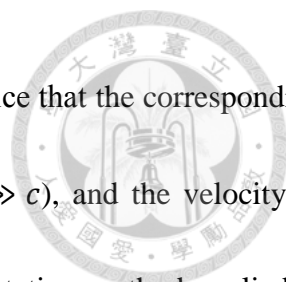
been discussed in section 2.2 for the "high frequency approximation". Notice that the corresponding

sensitivity zone is getting narrower for higher frequency situation ($r\omega \gg c$), and the velocity of

resulting *CCF* is getting close to that obtained by the conventional two-station method applied in

the dispersion measurement of surface waves. .
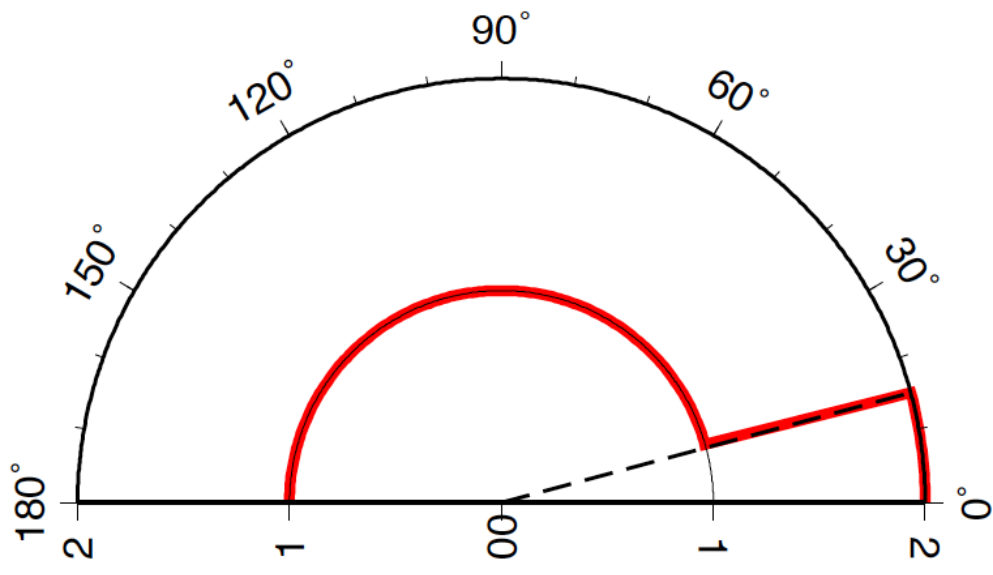
Figure 2-4: An example of a noise source strength of $\theta_o = 15°$, with $A_s(\theta)|_{\theta_o}$ plotted in polar

coordinates (with θ in degrees).
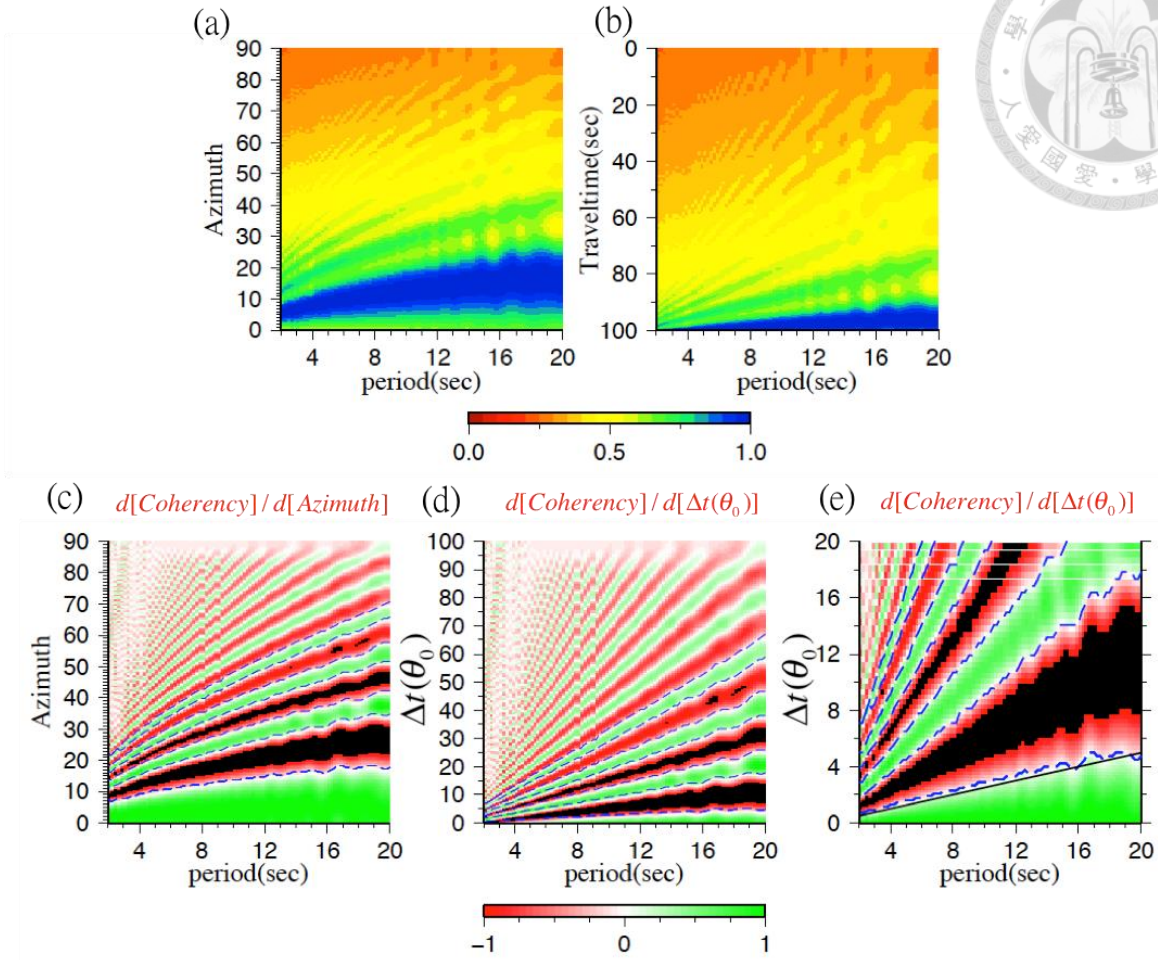
Figure 2-5: Example of the coherency as a function of a target bean zone of interest. (a) variation of coherency w.r.t the bean zone size (θ_o), in which the coherency is normalized at each individual frequency. (b) An alternative expression for panel (a) by replacing y-axis, the azimuth (θ_o) in panel (a) with travel time delay ($\Delta t(\theta_0) = \frac{r \cdot cos(\theta_0)}{c}$). (c) Results of gradient analysis for (a), in which contours of gradient zero are marked by dashed lines. (d) similar to panel (c) but represents the gradient pattern for the proper defined delay time. (e) A room-in image of the panel (d) for conveniently observing the first beam zone. The black solid line shows the corresponding phase shift $\pi/2$.

2.5 On The Remanent Fluctuations Of *CCFs*

It is known that the causal and acausal signals of noise-derived *CCFs* are mainly related to sources on the opposite directions along the station-station line, characteristics of *CCF* amplitude asymmetry can be used to explore the source heterogeneities of the ambient noise [e.g. *Brzak et al.*, 2009; *Chen et al.*, 2011; *Gu et al.*, 2007]. However, the connection between the *CCF* amplitude asymmetry and the nature of noise excitation is still ambiguous. Note that the *CCF* amplitude mentioned here is related to the "Signal" part (Eq. 2.2) of *CCF*.

Mathematically, since the *CCFs* are product of two wave records from sources integrated over time and space, the same *CCF* amplitudes can be generated from numerous spatial/temporal source combinations. For instance, the amplitude resulting from one source with amplitude 10 is identical to that from 100 sources with amplitude 1. In terms of ocean waves, the most dominant sources of ambient seismic noises, contribution to the *CCF* amplitudes from a huge wave may have similar effects to that from endless ripples.

Because of the highly non-unique relationship between *CCF* amplitudes and source characteristics, one may only extract very limited source information from the "Signal" amplitude of *CCF*. Here we show that additional source information can be retrieved from the "Noise" part of *CCF*. As discussed in Sec 2.2, the "Noise" part is resulted from the products of unrelated sources.

Various researches based on the Fluctuation Theory, normal mode analysis, acoustic
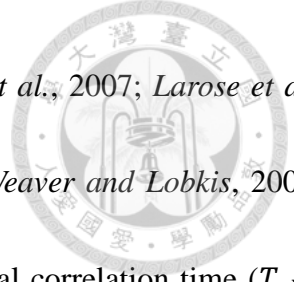
experiment and numerical simulation [e.g., *Derode et al.*, 2003; *Larose et al.*, 2007; *Larose et al.*, 2008; *Lobkis and Weaver*, 2001; *Sabra et al.*, 2005; *Tsai*, 2010; 2011; *Weaver and Lobkis*, 2005] pointed out that the noise level of a *CCF* generally decrease with the total correlation time (T_{all}). Similar feature has been confirmed in the seismological applications, in which the empirically defined signal-to-noise ratio (*SNR*) for the noise-derived *EGF*s also increase with T_{all} [*Bensen et al.*, 2007; *Gerstoft et al.*, 2006; *Sabra et al.*, 2006; *Seats et al.*, 2012]. However, the above *SNR* is only a roughly defined measure used to describe the *EGF* quality as seismic data. Here we aim to study the "Noise" in *CCF* in a more quantitative manner.

As mentioned earlier, by examining the *CCF* signal amplitude, one can't identify whether it is related to the source numbers or the source strength. Here we propose a new method to distinguish the effects from these two source parameters.

Considering a case with *N* independent noise sources, the displacement recorded by station at x in a specific cross-correlation time window j can be expressed as

$$U_j(x, t, \omega) = \sum_{s=1}^{N} A_s(\omega) \cos\left[\omega\left(t - \frac{r_{sx}}{c(\omega)}\right) + \phi_{sj}\right], \tag{2.21}$$

where A_s represents the amplitude triggered by the source s, r_{sx} represents the distance between source s and station, and ϕ_{sj} is the corresponding phase of the noise source s in the jth time window. As a result, the *CCF* obtained from the time window j is

$$CCF_j(\omega, \tau) = \sum_{s=1}^{N} A_s^2(\omega) \cos[\omega(\tau - \Delta t_s)] +$$

$$\sum_{k=1}^{N-1} \sum_{l=k+1}^{N} A_k(\omega) A_l(\omega) \cos(\omega\tau + \phi_{klj}) \qquad (2.22)$$

, where $\Delta t_s = \frac{r_{sx} - r_{sy}}{c(\omega)}$ represents the travel-time delay from the source s to two stations, and the

phase shift $\phi_{klj} = \frac{\omega(r_{kx} - r_{ly})}{c(\omega)} + (\phi_{kj} - \phi_{lj})$ reveals a phase delay resulting from delay time of two

unrelated sources (k and l) traveling to two stations separately and a difference of the initial phase.

It is notice that the first term is identical to the signal in the *CCF* of eq. (2.2), which is the product

of *N* correlated sources, and the second term consists of much more elements, as it is contributed

from combinations of all unrelated sources ($N(N-1) \sim N^2$)

Assuming that the distribution and triggering time of noise sources are both random (ϕ_{kj} and

ϕ_{lj} are randomly chosen), and considering the fact that a random walk of N^2 steps is equivalent to

travel a total distance of N units, the summation of the N(N-1) elements in equation (2.22) can be

simplified as a product of the amount of noise source N and an average value $\overline{A_x^2} = \frac{1}{N(N-1)} \sum A_k A_l$

[*Tsai*, 2011, eq. (19)]. As a result, the second term in the equation (2.22) can be expressed as

$$\sum_{k=1}^{N-1} \sum_{l=k+1}^{N} A_k(\omega) A_l(\omega) \cos(\omega\tau + \phi_{klj}) = N\overline{A_x^2} \cos(\omega\tau + \phi_j^{avg}), \qquad (2.23)$$

where ϕ_j^{avg} represent a final phase shift and are randomly distributed within $0 \sim 2\pi$.

It is noteworthy that the ϕ_j^{avg} are not a constant but related to the time widow j used in cross

correlation. Consequently, it behaves as a meaningless random oscillation and we name this term

"Noise" or "remanent fluctuations" of a *CCF*. We define an ensemble averaged $\overline{CCF^M}$ through

stacking M independent CCF_j as

$$\overline{CCF}^M = \frac{1}{M}\sum_{j=1}^{M} CCF_j. \qquad (2.24)$$

Since the ϕ_j^{avg} are randomly chosen ($1 \leq j \leq M$), the sum of the noise term will decrease as

$\sqrt{M}$ [*Derode et al.*, 2003; *Larose et al.*, 2008; *Lobkis and Weaver*, 2001; *Tsai*, 2010; 2011]. Next,

considering the fact that normal modes triggered by noise sources are degenerate and

non-equipartitioned, equation (2.22) can be approximately expressed as a coherency form:

$$\overline{CCF}^M(\omega,\tau) \sim \sum_n \frac{\overline{A_n^2(\omega,\theta)}}{A^2(\omega)} G_n(\omega,\tau) + \left(\frac{N(\omega)}{\sqrt{M}}\right)\left(\frac{\overline{A_x^2(\omega)}}{A^2(\omega)}\right)\cos\left(\omega\tau + \phi_M^{avg}\right), \qquad (2.25)$$

where suffix n represents the mode index.

In terms of the above mode formulation (Eq. 2.25), there is a direct link between the mode

type of resulting *CCF* and the mode type of noise sources. The pressure fluctuation on the seabed

resulting from ocean gravity waves tends to excite fundamental mode surface waves more

efficiently; naturally, *CCFs* obtained from continuous seismic data are generally dominated by

fundamental mode surface waves. On the other hand, one may retrieve *EGF* containing body waves

through cross correlating a particular time window with energetic body wave noise sources, for

instance, the earthquake coda waves or recorded excited by some special weather conditions [*Boué*

et al., 2014; *Landès et al.*, 2010; *Lin et al.*, 2013; *Zhang*, 2010].

It is interesting to note that, if the noise source property (amplitude and distribution) does not

vary with time, the signal part of *CCF* does not vary with correlation time either. Precisely speaking,

the main propose of stacking, a common process in the noise study, is to suppress the noise part,
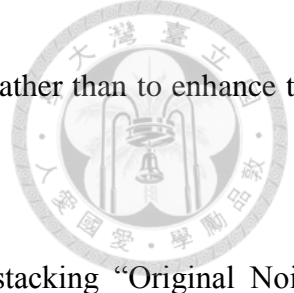
which decays with correlation time as show by the 2nd term in Eq. 2.25, rather than to enhance the signal itself.

For convenience, we name the noise level of a coherency prior stacking "Original Noise Level, ONL", which is defined as $ONL(\omega) \equiv N(\omega)(\frac{\overline{A_x^2(\omega)}}{\overline{A^2(\omega)}})$. Note that, in order to reduce the sensitivity to the ambient noise level, the defined ONL has been normalized by an average source strength $\overline{A^2(\omega)}$.

In Chapter 4 we will estimate the ONL for $CCFs$ derived in Taiwan and Korea, and detail its applications.

Chapter 3

Characteristics of Short Period Secondary Microseisms (SPSM) in Taiwan – the influences of bathymetry and monsoons on SPSM excitations

This chapter has been published in Geophysical Research Letters [*Chen et al.*, 2011] under the title: "Characteristics of Short Period Secondary Microseisms (SPSM) in Taiwan – the influences of bathymetry and monsoons on SPSM excitations."

Abstract

Taking advantage of a unique opportunity provided by a dense array of coastal short-period seismic stations and the diverse bathymetry around Taiwan, we examine how the long-range coherent ambient noises are influenced by surrounding ocean settings using the cross-correlation functions (*CCFs*) between pairs of stations. The effective energy of the *CCFs* derived from three components of short-period seismometer data falls within the frequency range of the short period secondary microseism (SPSM). The spatial variations mapped from the amplitude asymmetry of *CCFs* and source migration images evidently demonstrate that the SPSM strengths are closely linked to the drastic changes in offshore ocean characteristics and result in much stronger SPSM in the shallow and narrow Taiwan Strait than in deep open seas of eastern Taiwan. The temporal variations of the

CCF strengths exhibit very good correlations with the wind speeds and wave heights, explicitly indicating the observed SPSM is dominated by local sources generated from wind-driven ocean waves around offshore Taiwan.

3.1. Introduction

It has been shown that the cross-correlation function (*CCF*) of continuous records from two wave sensors resembles the elastic impulse response between them [e.g., *Lobkis and Weaver*, 2001;. *Shapiro and Campillo*, 2004; *Snieder*, 2004; *Weaver and Lobkis*, 2002]. Theoretically, in a fully diffusive wave field, the causal (time positive portion) and acausal (time negative portion) signals of the derived *CCFs* are identical in both their phases and amplitudes. In the application to seismic data, prior to cross-correlation, various normalizations applied to raw data are indispensable in order to suppress non-diffusive signals such as earthquakes and instrumental irregularities. The phase symmetry with respect to zero time lag is usually fairly attained in the resulting *CCFs*, and the violation of phase symmetry is mostly related to errors of the internal clock of seismometers [e.g., *Lukac et al.*, 2009a; *You et al.*, 2010]. On the other hand, compatible amplitudes for causal and acausal signals are less common, and it is due to the fact that the source strengths of background noises are not spatially homogeneous. Because the causal and acausal signals of *CCFs* are mainly excited by the sources from opposite directions along the line of the station pair, the
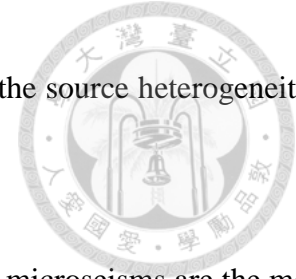
characteristics of *CCF* amplitudes has become a useful means to explore the source heterogeneities of ambient noises [e.g., *Brzak et al.*, 2009; *Gu et al.*, 2007].

Seismic ambient noise exists across a wide frequency band, in which microseisms are the most common signals in seismograms at periods between ~3 to 20 seconds. Microseisms are primarily generated by the coupling between ocean waves and seafloor [e.g., *Cessaro*, 1994; *Longuet-Higgins*, 1950]. While some studies on the amplitude asymmetry of *CCFs* suggest that short-period microseisms (<10 seconds) are mostly excited in the coastal area [e.g., *Stehly et al.*, 2006; *Yang and Ritzwoller*, 2008], other investigation found that they could be effectively generated by resonance of the compression waves at favorable depth in deep oceans as well [*Kedar et al.*, 2008].

The island of Taiwan is located on the arc-continent collision boundary between the Eurasia plate and Philippine Sea plate. Owing to this complex tectonic setting, the north-south trending mountain ranges run through the central Taiwan and divide the island into the western and eastern parts surrounded by very diverse bathymetry as shown in Figure 3-1a. The seafloor morphology off southern and eastern Taiwan is strongly associated with the juxtaposition of oblique collision and subduction between two plates and the water depth steeply increases off the coast. In contrast, to the western and northern Taiwan the narrow Taiwan Strait basin, part of the passive South China continental margin, has relatively flat and much shallower bathymetry with an average water depth of about 70 meters only [*Liu et al.*, 1998].
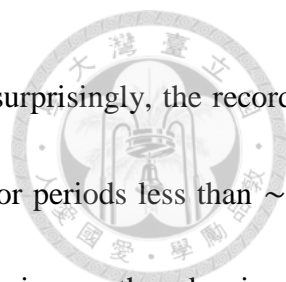
In Taiwan, most seismic stations are deployed near the coasts. Not surprisingly, the recorded background seismic noise level is much higher than the global average for periods less than ~8 s [*Lin et al.*, 2010]. Although the energetic background noise may obscure micro earthquake signals, it turns out to be an advantage in the implementation of ambient noise tomography [*You et al.*, 2010]. Moreover, the combination of dense coastal stations and diverse offshore morphology provides a unique opportunity to explore how the long-range coherent ambient noises are influenced by the fluctuations in offshore settings.

In this study, we analyze *CCFs* derived from three components of data recorded by short-period seismic stations of the Central Weather Bureau Seismic Network (CWBSN). The major energy of the resulting *CCFs* falls within the period range of about 2 to 6 seconds. Microseisms within this frequency band are known as short period secondary microseism (SPSM) [*Bromirski et al.*, 2005; *Stephen et al.*, 2003].

We characterize the spatiotemporal property of the SPSM intensity by the amplitude asymmetry of *CCFs* and their source distributions in offshore regions determined by a migration imaging method. Results from both approaches clearly demonstrate that SPSM are much stronger in shallow water settings. Moreover, the close connection between the long-range coherent seismic ambient noises and atmosphere perturbations is manifested by the excellent temporal coherence between the monthly-averaged *CCF* amplitudes and the wind speeds and wave heights.
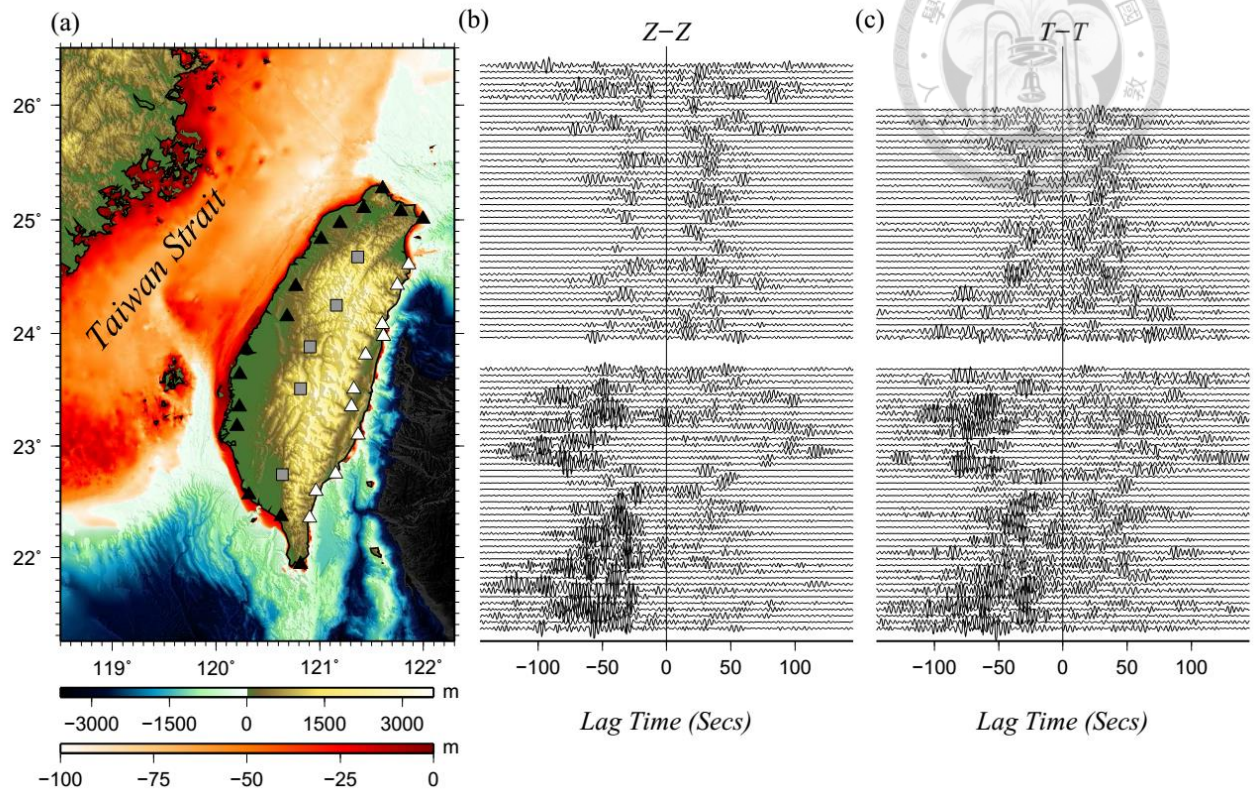
Figure 3-1: (a) Map of the study region and stations used. The west and east coast stations are denoted by black and white triangles, respectively. The inland stations are denoted by gray squares. The topography and bathymetry are also shown as defined by the color bars. Different color scale is used for the depth range between 0 and 100m to better illustrate the shallow bathymetry in the Taiwan Strait. (b) *CCFs* derived from the vertical component. The causal and acausal signals of *CCFs* are related to energy coming from the inland and coastal direction, respectively. The *CCFs* (top) from stations along the east coast and (bottom) for the stations from the west coast. The amplitude of each trace is normalized by the smaller peak amplitude between the corresponding causal and acausal signals. (c) Same as Figure 1b, except for the transverse component.

3.2. DATA

Among a total of 78 short-period CWBSN stations distributed over the island, we select 31 representative stations for this study; 26 of them are located near the coastal area with an approximately uniform spacing along the circum-island coast line, and the rest situated in the inland area along the Central Mountain Ranges (Figure 3-1a).

It is well known that noise-derived *CCFs* are dominated by fundamental mode Rayleigh waves for the vertical (Z) and radial components, and by Love waves for the transverse (T) component [e.g., *Campillo and Paul*, 2003]. To take both wave types into account, we derive *CCFs* of the Z-Z and T-T components using three components of continuous records in the year 2006. The main purpose of this study is to examine how the noise excitations are influenced by different offshore settings, hence, only the station pairs composed of one coastal and one inland station are used in the *CCF* analysis.

Figure 3-2 shows the spectral contents of the resulting *CCFs* on the T-T and Z-Z components which both concentrate within the narrow SPSM band at the peaks of around 4 seconds. (See auxiliary material for instrument response and details on data processing.)
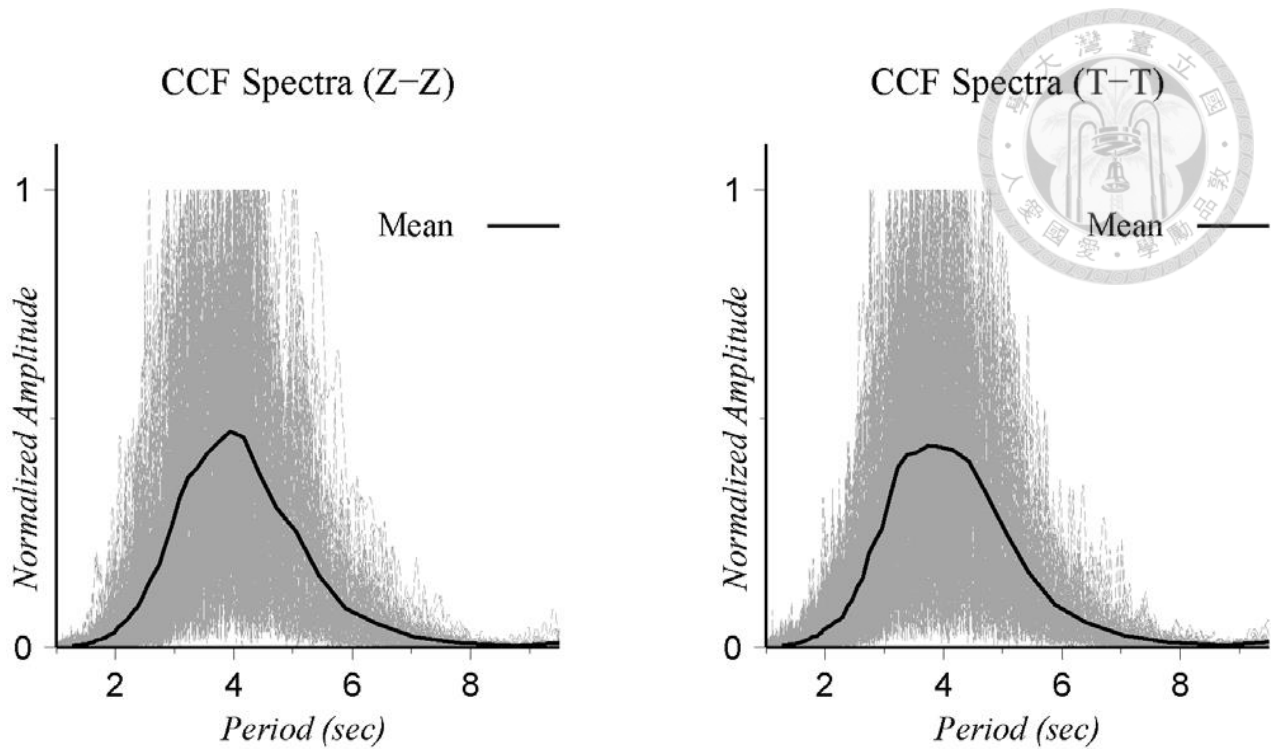
Figure 3-2: Amplitude spectra of the noise-derived *CCFs* for the Z-Z and T-T components. All the spectra are normalized to emphasize their frequency contents. The gray lines show the spectra of annual *CCF* stacks for all the available pairs, and black curves represent the corresponding mean spectra.

3.3. Spatiotemporal Characteristics of SPSM

3.3.1. Spatial Variations of CCF Amplitude Asymmetry

We first explore the spatial variations of the noise excitations using the amplitude asymmetry of annual *CCFs*. Only the *CCFs* with effective emergence of the empirical Greens function (*EGF*) are used in this analysis. 86 pairs of the *CCFs* from the Z-Z components and 79 pairs from the T-T components meet the selection criterion (see auxiliary material), and are shown in Figure 3-1b and
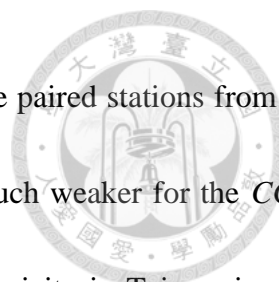
3-1c, respectively. It is fairly clear that the *CCFs* obtained with one of the paired stations from the west coast show strong amplitude asymmetry, while the asymmetry is much weaker for the *CCFs* derived from the eastern coastal stations. It's worth mentioning that seismicity in Taiwan is very active along the east coast and very quiet on the west coast [e.g., *Wu and Rau*, 1998]. The above results suggest that transient earthquake signals have little contribution to the emergence of *EGF*s after raw data normalization.

The amplitude asymmetry is further quantified by taking the ratio of the peak amplitudes of a *CCF* in one of the causal or acausal signals excited from the coastal direction to that in the other signal excited from the inland direction. The results are displayed as vectors on the locations of coastal stations, with their lengths proportional to the amplitude ratios and direction to the energy flow along the lines of station pairs. Results from Rayleigh waves (Figure 3-3a) and Love waves (Figure 3-3b) demonstrate very similar patterns wherein the SPSM excitations from the Taiwan Strait and northern coast are much stronger than those from the eastern and southern coasts. The color images of the offshore area in Figure 3-3 are the lateral variations of the SPSM intensity resulting from the noise source migration (see Section 3.3.2).

We have also conducted a similar experiment using 10 coastal broad-band stations to ensure that the above observations on the spectral content and amplitude asymmetry are not biased by the band-limited response of short-period stations (see auxiliary material).
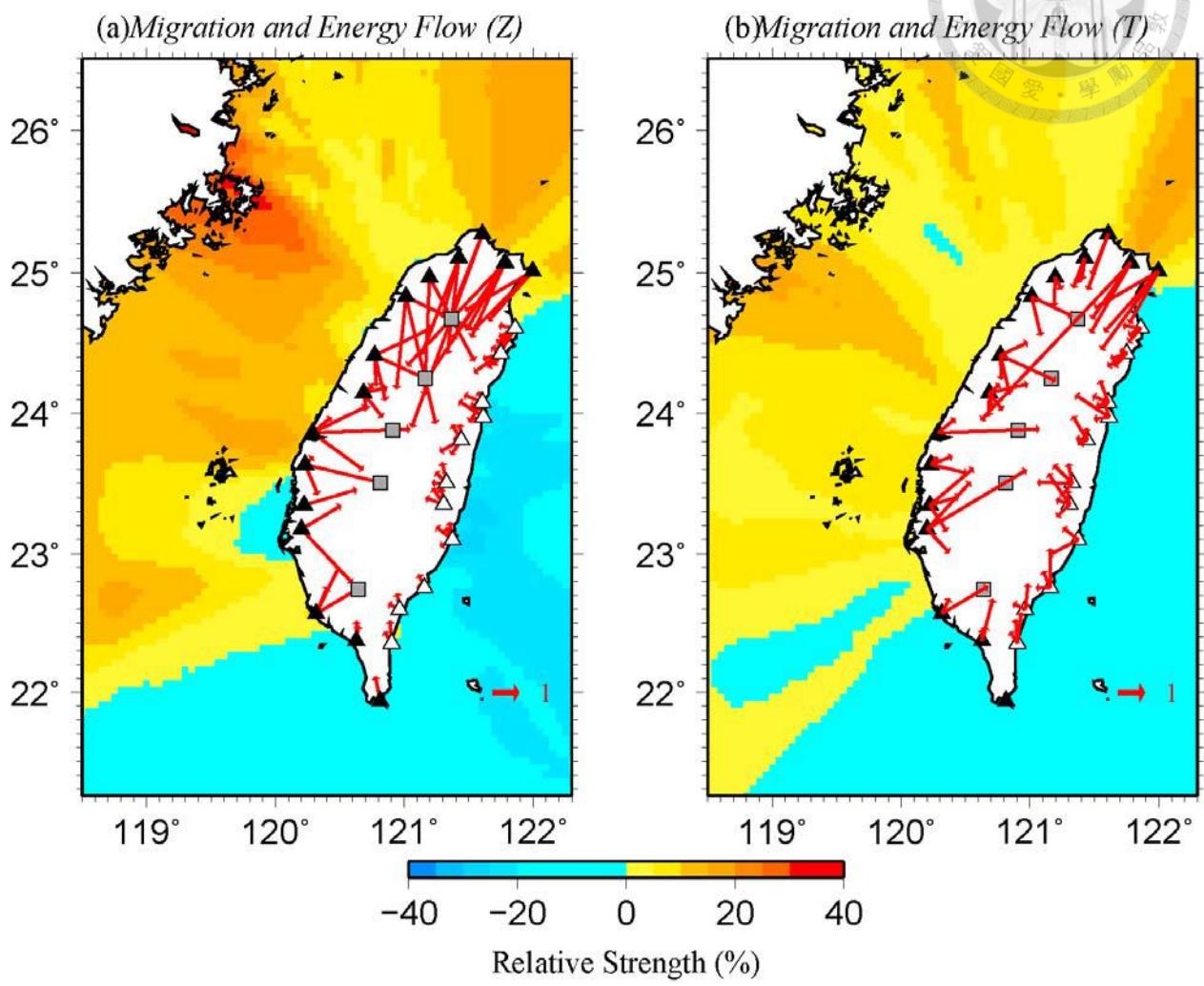
Figure 3-3: Results of migration imaging of SPSM strengths and relative peak amplitudes of the annual *CCF* stacks for the (a) Z-Z and (b) T-T components. The relative strengths of *CCFs* are shown as vectors on the locations of coastal stations, with their lengths proportional to the amplitude ratios (scale shown in the lower-right corner of each panel) and directions to the energy flow along the line of station pairs (see text for details). The results of migration images are expressed as perturbations with respect to the regional average.

3.3.2. Spatial Variations of SPSM Sources From Migration Imaging Method

To further look into the spatial distribution of noise sources, we implement the migration imaging method [*Brzak et al.*, 2009]. Briefly speaking, with the noise-derived 2-D surface wave models [*Chen et al.*, 2009] in the Taiwan island and an average wave velocity for the oceanic area, the excitations of potential SPSM sources are evaluated based upon the strength of cross-correlated signals among all available station pairs, and the whole offshore area in Figure 3-1 are estimated at every 0.05 by 0.05 degree grid cells.

All the annual *CCFs* on the Z-Z and T-T components are used for the source migration analysis, and the lateral variations of the resulting SPSM intensities are presented as relative perturbations with respect to the regional average (Figure 3-3). The distribution of source intensity is consistent with the pattern of the *CCF* amplitude asymmetry, indicating that the SPSM are stronger in the west and north offshore Taiwan. (More details on the migration imaging method and the stability test are given in the auxiliary material)

3.3.3. Temporal Variations of CCF Amplitudes and Their Correlation to Wind Speeds and Wave Heights

The temporal variations of overall *CCF* amplitudes are estimated by the relative differences of the average peak amplitudes between individual monthly stacks and annual stacks. We then compare the monthly variations of *CCF* amplitudes with wind speeds and offshore ocean wave heights for the same time period to further demonstrate their close relationships. The monthly means of 20 island-wide coastal meteorological observatories and 6 offshore buoys in 2006 are used to derive the temporal variations of wind speeds and wave heights with respect to their annual means (see auxiliary material for data sources of wind speeds and wave heights). The comparison is shown Figure 3-4, which reveals strong positive correlations among the variations in *CCF* amplitudes, wind speeds and ocean wave heights.
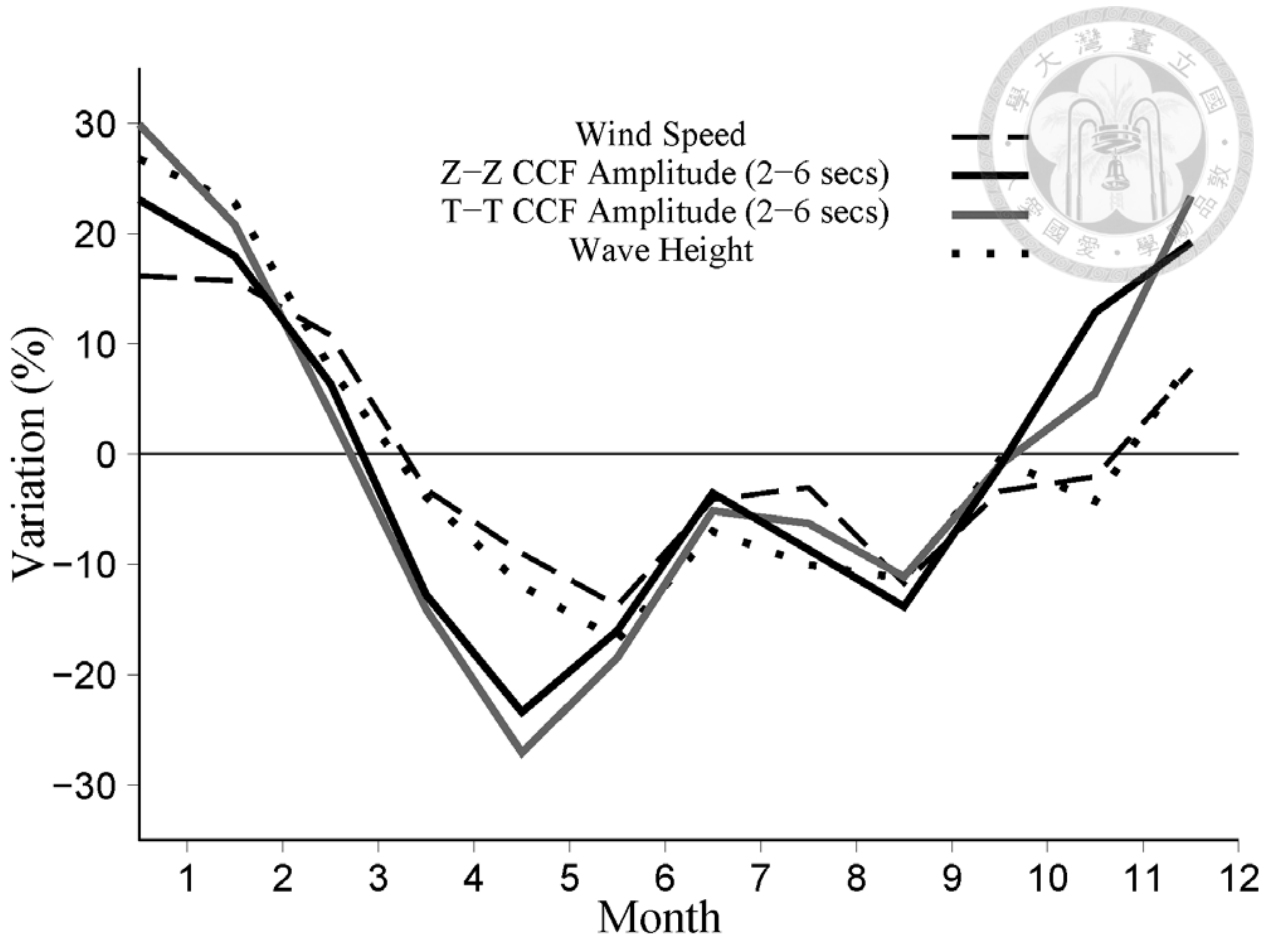
Figure 3-4: Comparison of the temporal variations between wind speeds and *CCF* amplitudes. The dashed and dot lines represent the variations of overall monthly wind speeds and wave heights with respect to their annual averages. The solid lines are the variations of overall *CCF* amplitudes for the Z-Z component (black line) and T-T component (gray line). Only either the causal or acausal *CCF* signals corresponding to the directions of energy flow that point from the coastal stations to inland stations are taken into account.

3.4. Discussions and Conclusions

We have examined the spatiotemporal properties of SPSM around the offshore Taiwan using noise-derived *CCFs*. Both the spatial and temporal variations of the SPSM excitations exhibit characteristic features which are likely linked to the diverse offshore settings and monsoon migration in Taiwan.

The strong east–west asymmetry in SPSM intensity is evidently demonstrated by the *CCF* amplitude ratios and migration images, and such pattern matches very well with the corresponding drastic changes of seafloor morphology and ocean characteristics surrounding Taiwan shown in Figure 3-1a. Although the shallower water depth and the narrow ocean strait are plausible causes for the stronger SPSM in the Taiwan Strait and northern offshore, there are other possibilities to explain our observations, such as the wave heights or frequency content of ocean waves.

We first examine the wave heights in offshore Taiwan from data recorded by six offshore buoys in the year 2006 (see Figure S3-4 in the auxiliary material). Wave heights recorded in the northern end and southern end of Taiwan are relatively higher (106 cm) and lower (63 cm), respectively, consistent with the observed SPSM intensity. However, there is no significant difference in annual means of wave heights (∼90 cm) between the west and east coasts.

Second, there might be a doubt that the observed east–west asymmetry in SPSM excitations is merely a consequence of the difference in frequency contents between waves in the Taiwan Strait
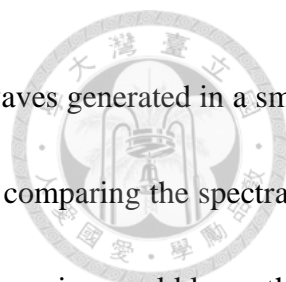
and the eastern open sea, as it is well known that the frequency of ocean waves generated in a small ocean basin is higher than in the deep open sea. This could be verified by comparing the spectra of *CCF* signals from the east and west coasts, since waves with different frequencies would leave their signatures on the resulting secondary microseisms. The comparison made in Figure S3-5 of the auxiliary material shows that the frequency contents of the signals from both coasts are very similar. Another line of evidence is provided by the strong temporal correlation between *CCF* amplitudes and local wind speeds (Figure 3-4). It implies that the observed SPSM are dominated by waves generated locally in offshore area, and the distant lower frequency ocean swells in deep sea should have little influence.

Given the above arguments, we thus propose that much of the observed spatial variations in SPSM intensity are related to the contrasting offshore settings between the east and west coasts. Stronger excitations take place in the narrow Taiwan Strait where water depth is very shallow, while the excitations are relatively weak in the eastern offshore area, an open sea with water depth increases rapidly off the coast.

The correlation between bathymetry and excitation strength is in agreement with the generation mechanism of microseisms. Because ocean wave energy is transferred to microseisms through the interaction of ocean swells with a shoaling ocean bottom, and the transferred energy decays quickly with increasing water depth, reaching to nearly zero at the depth equal to half the
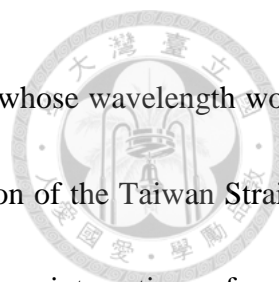
wavelength. The observed SPSM correspond to ∼8 second ocean waves, whose wavelength would be ∼80 meters. As shown in Figure 3-1a, the water depth of a large portion of the Taiwan Strait is less than 40 meters, and such shallow bathymetry may enhance the nonlinear interactions of waves through the stronger bottoming friction, and introduce more effective SPSM excitation.

The narrow Taiwan Strait might be a privileged place for the generation of secondary microseisms as well. Because secondary microseisms are induced by nonlinear interactions between oppositely traveling ocean waves of similar wavelengths [*Longuet-Higgins*, 1950], their excitations should be more vigorous through the coastal reflections of ocean waves from both sides of the strait. Furthermore, the decay of ocean waves due to geometrical spreading in a confined basin like the Taiwan Strait is much less than that in an open ocean, and this could be another possible cause for the observed stronger SPSM in the west offshore of Taiwan.

On the other hand, the strong temporal correlation among *CCF* amplitudes, wind speeds and wave heights confirms that the SPSM strength is closely related to the local wind fields and ocean waves [e.g., *Bromirski and Duennebier*, 2002; *Gerstoft and Tanimoto*, 2007; *McNamara and Buland*, 2004; *Traer et al.*, 2008]. It is known that the wind fields over Taiwan are strictly tied to the East Asia monsoon system, in which the northeast monsoon commences in September, prevails from October to January, and weakens continuously thereafter; the southwest monsoon only prevails in June to August, and is much weaker than the northeast monsoon. The observed temporal

variations of SPSM strength can thus be well explained by the seasonal migration of the two

monsoons.

Chapter 4

On the Noise of the ambient noise cross-correlation function and its applications

In Chapter 2 we present the basic ideas for the "Signal" and the "Noise" of *CCF* (coherency) separately; in Chapter 3, based on the asymmetry of *CCF's* amplitudes observed from the stations in the island of Taiwan, we demonstrate that the SPSM strengths are closely linked to the offshore ocean characteristics. In this chapter, we are going to provide a quantitative description on the behavior of the "Noises" within the *CCF*, and show that, similar to the "Signals", the "Noises" are also closely related to some unique properties of source excitations. In particular, we show that, after the temporal variations of sources are removed by random stacking, the resulting temporal evolution of *CCF* noise level is well predicted by the theoretical formulations.

4.1. Introduction

Since 2004, the early studies using ambient seismic noises are mostly focused on their tomographic applications, in which the noise-derived *EGF*s are used similarly as data [e.g., *Bensen et al.*, 2007; *Gerstoft et al.*, 2006; *Seats et al.*, 2012], and the data quality are usually determined by its signal to noise ratio (*SNR*). In such applications, only the major signals are useful, and the rest of *CCFs* are considered as useless noises and should be suppressed by increasing the total cross-correlation time.

As a matter of fact, the later part tailing the major signals of *CCFs* is not just random noises, they contain the coda signals of the Greens function, and one may reconstruct *EGF* by cross-correlating the *CCF* coda [*Froment et al.*, 2011; *Stehly et al.*, 2008].

In practice, since the total cross-correlation time is not infinite, the existence of "Noises" within *CCFs* is inevitable. However, while the applications of *CCFs* have been extended to their coda part, only very few studies have focused on the true "Noises" within the *CCFs* [e.g., *Derode et al.*, 2003; *Larose et al.*, 2008]. In the following, we present a procedure to estimate the noise level of *CCFs* and illustrate how the estimated noise levels are related to the source characteristics.

For simplicity, the expression for the coherency in Eq. (2.25) can be written as

$$CCF(\omega, T) = \frac{\overline{A(\theta, \omega)^2}}{A(\omega)^2} G(\omega) + \frac{ONL(\omega)}{\sqrt{T}} . \tag{4.1}$$

The first term on the RHS of above equation is the "Signal" of *CCFs*, where $G(\omega)$ is the Greens
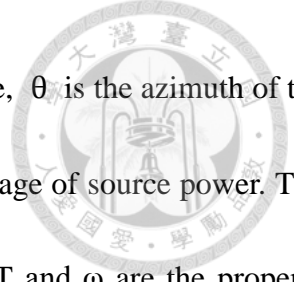
function, $\overline{A(\theta,\omega)^2}$ is the average source power within the sensitivity zone, θ is the azimuth of the source relative to the station-station line, and $\overline{A(\omega)^2}$ is an azimuthal average of source power. The second term in Eq. (4.1) is the product of the unrelated sources, where T and ω are the properly defined total cross-correlation time and angular frequency, respectively. *ONL* is the "original noise level", defined as $ONL(\omega) \equiv N(\omega)\left(\frac{\overline{A_x^2(\omega)}}{\overline{A^2(\omega)}}\right)$, in which $N(\omega)$ is the number of sources. (Please see Section 2.5 for more details).

It is clear that the *CCF* noises are represented by the second term on the RHS of Eq. (4.1). A well-known feature about the *CCF* noises is that *CCF SNR* generally grows with the cross-correlation time, as is shown in Eq. (4.1). Also note that, given a constant source condition, ONL is not a time-dependent term, and is proportional to the number of sources. Thus, ONL may provide us not only the exact noise level of *CCF* but also the information about the source population of ambient noises, which may help us to further characterize the excitation mechanisms. We should point out that such information is inaccessible from the "Signal" part of *CCFs*.

In nature, ONL varies with time, because the source population/strength of ambient noises are highly related to the atmospheric conditions, accordingly, the "stable source distribution" is obviously not a reasonable assumption. It turns out that we cannot precisely evaluate ONL at any given time without knowing the exact source conditions. Nevertheless, we may measure the Expected Original Noise Level ($\widehat{ONL}(\omega)$), i.e., the ONL averaged over a given time period. Details of this procedure will be given in the Section 4.2, and the applications of the estimated $\widehat{ONL}(\omega)$

will be presented in Section 4.3 and Section 4.4.

In this study, we evaluate $\widehat{ONL}(\omega)$ for *CCFs* derived from the vertical component continuous seismic data recorded at 37 and 17 broadband stations in Taiwan and Korea, respectively (Figure 4-1). Using the method presented by *Welch* [1967], the daily coherency is computed with a 4000s time window and a 50% overlapping moving window.
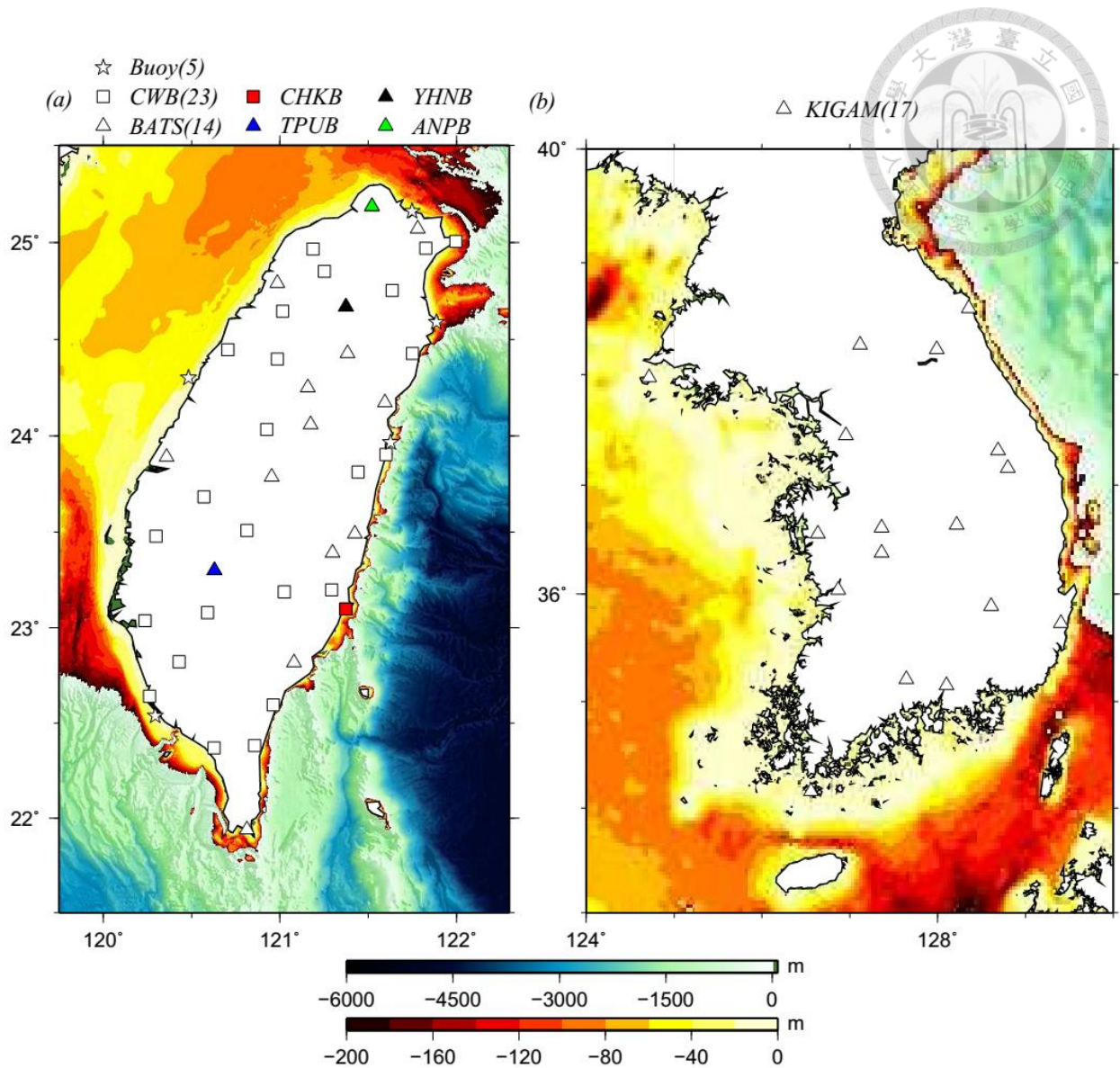
Figure 4-1: Map and bathymetry of the study region (a) Taiwan and (b) Korea, and the distribution of stations. Station pairs of CHKB-TPUB and ANPB-YHNB are used to demonstrate convergence of *CCF* with time in Section 4.2 and Section 4-3. Five offshore buoys used in the estimations of the significant wave period around Taiwan are marked by star.
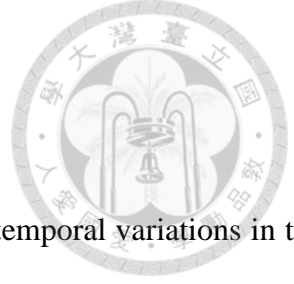
4.2. The Expected Original Noise Level $\widehat{ONL}(\omega)$

To begin with, we first consider a simple case in which there are no temporal variations in the source condition. In such case, the constant ONL can be easily derived by examining the time-dependent residuals of *CCFs* as follows.

Taking the final $CCF(T_{ref})$ as the reference *CCF*, the noise part of $CCF(T_{ref})$ is $\frac{ONL(\omega)}{\sqrt{T_{ref}}}$; likewise, the noise part of *CCF* stacked over any given correlation time T is $\frac{ONL(\omega)}{\sqrt{T}}$. Note that, just like *CCF*, the function of noise level is a time series. For convenience, here we evaluate ONL by the root-mean-square of the waveform residual. With a properly defined time-dependent term $(\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{T_{ref}}})$, $ONL(\omega)$ can be easily expressed as:

$$ONL(\omega) = \left\| CCF(\omega, T_{ref}) - CCF(\omega, T) \right\| / (\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{T_{ref}}}), \tag{4.2}$$

where $\|\cdot\|$ is a root-mean-square operator. Examples of waveform residuals developing with correlation time are shown in Figure 4-2.

As mentioned in Section 4.1, the constant source condition is not a practical assumption in nature. For instance, in Taiwan, the excitation strengths of SPSM are strongly influenced by the near coast wind speeds and wave heights [*Y-N Chen et al.*, 2011]. This implies that the resulting "Signal" part of any straight stacked *CCF* might not be consistent with that of the reference *CCF*, and the retrieved $\widehat{ONL}(\omega)$ through Eq. (4.2) will be biased. This is well demonstrated in Figure 4-3, where the residuals (red line) between the reference *CCF* and the straight-stacked *CCFs* are, indeed,
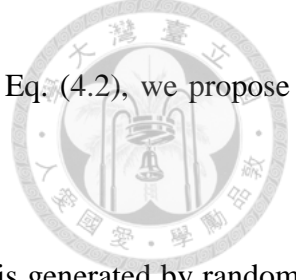
not decreasing with time monotonically. To evaluate the $\widehat{ONL}(\omega)$ using Eq. (4.2), we propose to suppress the source temporal variations as follows.

Instead of straight stacking, the *CCF* of any target correlation time is generated by randomly stacking over the same correlation time of individual *CCFs*. By doing so, we aim to homogenize the source conditions over time, yielding similar "Signals" for the reference *CCF* and *CCF* at any target correlation time, i.e., $\frac{\overline{A(\theta,T)}}{A(T)} \approx \frac{\overline{A(\theta,T_{ref})}}{A(T_{ref})}$.

In this study, we randomly stacked 50 *CCFs* for each target correlation time, and the $\widehat{ONL}(\omega)$ is given by the averaged residual of the 50 estimates. In Figure 4-3a, we compare the evolution of residuals over time for the random-stacked *CCFs* and the straight-stacked *CCFs*. It clearly shows that the relative noise level obtained from random-stacked *CCFs* decays exactly with $\sqrt{T}$ (shown by black cross). On the other hand, the pattern of the straight-stacked *CCFs* is obviously perturbed by the monsoon properties around Taiwan.

Of course, the source temporal variations cannot be perfectly removed by the above approach. Taking this into account, Eq. (4.2) is modified as

$$\widehat{ONL}(\omega) \cdot \left(\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{T_{ref}}}\right) = \left\|CCF(\omega,T_{ref}) - CCF(\omega,T)\right\| + err(\omega), \tag{4.3}$$

In Figure 4-3b, we replace the X-axis in Figure 4-3a, the total correlation time, with $\left(\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{T_{ref}}}\right)$. In such presentation, the $\widehat{ONL}(\omega)$ is simply identical to the regression slope. The stable slope in Figure 4-3b, shown by the black dashed line, implies that the error caused by source uncertainties (Eq. 4.3) must be either small or having little temporal variations and have little

influences on our measurements.

The case shown in Figure 4-3 is not a particular example, similar results are obtained in all the *CCFs* used in this study. More examples from *CCFs* in Taiwan and Korea are presented in Figure 4-4 and Figure 4-5 respectively.

We should point out, the effective sensitivity zones of *ONL* and the corresponding *CCF* signals are different. For *CCF* signals, as discussed in the Ch. 2, since the fundamental mode surface waves are confined within a narrow time widow, their energies are primarily associated with sources along the station-station lines. On the other hand, since much longer time windows are used to evaluate *ONL* in this study and the residuals are usually evenly distributed (Figure 4-2b), the associated sensitivity zone would be much wider.
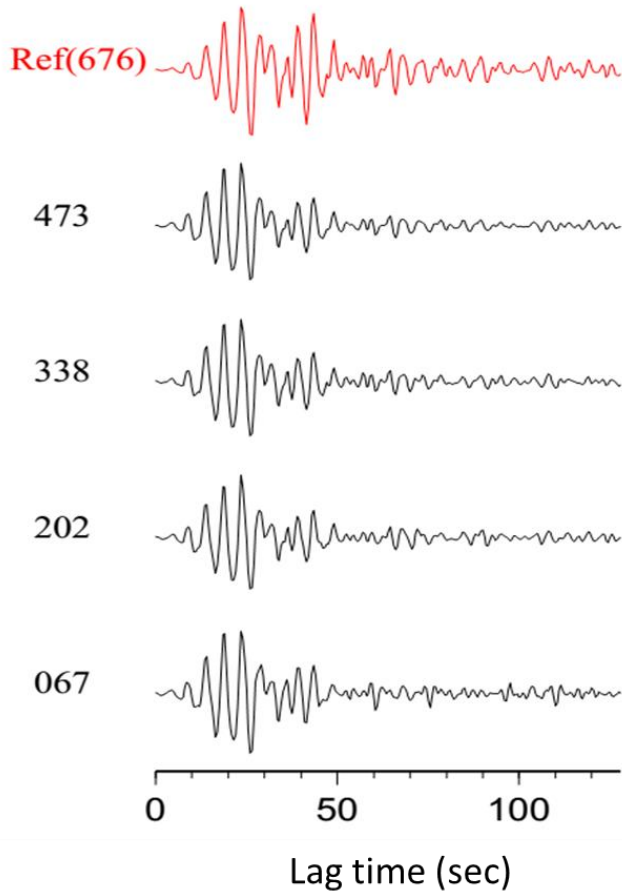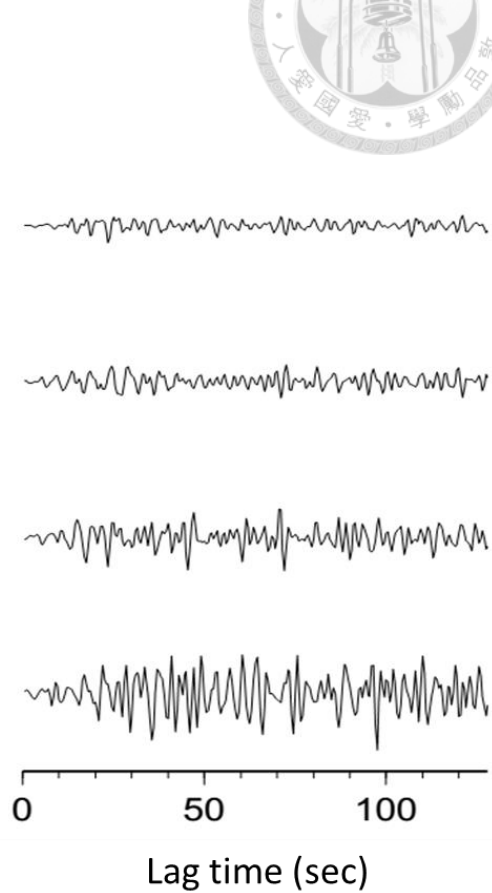
Figure 4-2: Evolution of *CCF* of pair CHKB-TPUB (inter-station distance 79 km). The total correlation time are shown in the left side of each trace. (a) Comparisons of waveforms between the reference *CCF* (the top trace) and the averaged one. (b) A demonstration of the waveform residual between the target *CCF* and the reference one. The amplitude is normalized by the same value. Apparently, amplitude of the waveform residual is decreasing with total correlation time.
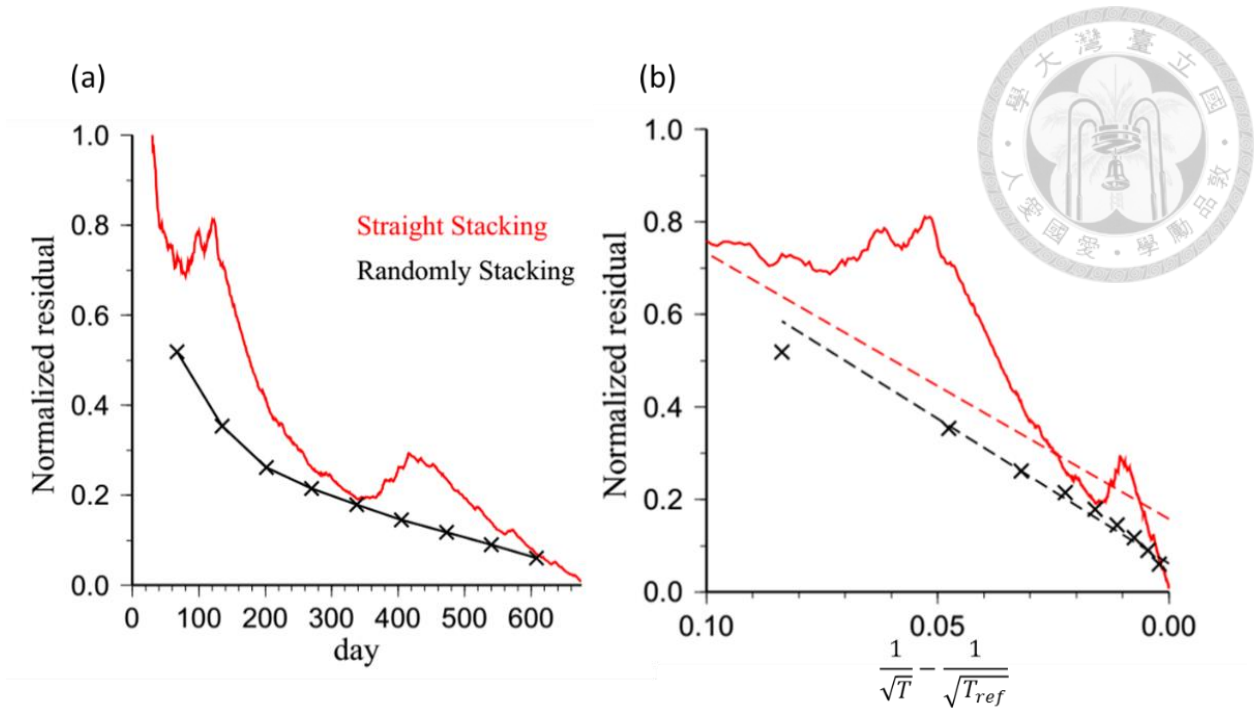
Figure 4-3: A comparison of remnant noise level derived from waveform residuals for straight (red) and randomly (black) stacks for the same station pair CHKB-TUPB in figure 4-2. (a) Waveform residual (rms) as a function of integration time. The cross is the average residual measurement of the randomly stacked *CCF*. (b) Waveform residual as a function of the properly defined time-dependent term, and the dash lines are the regression results.
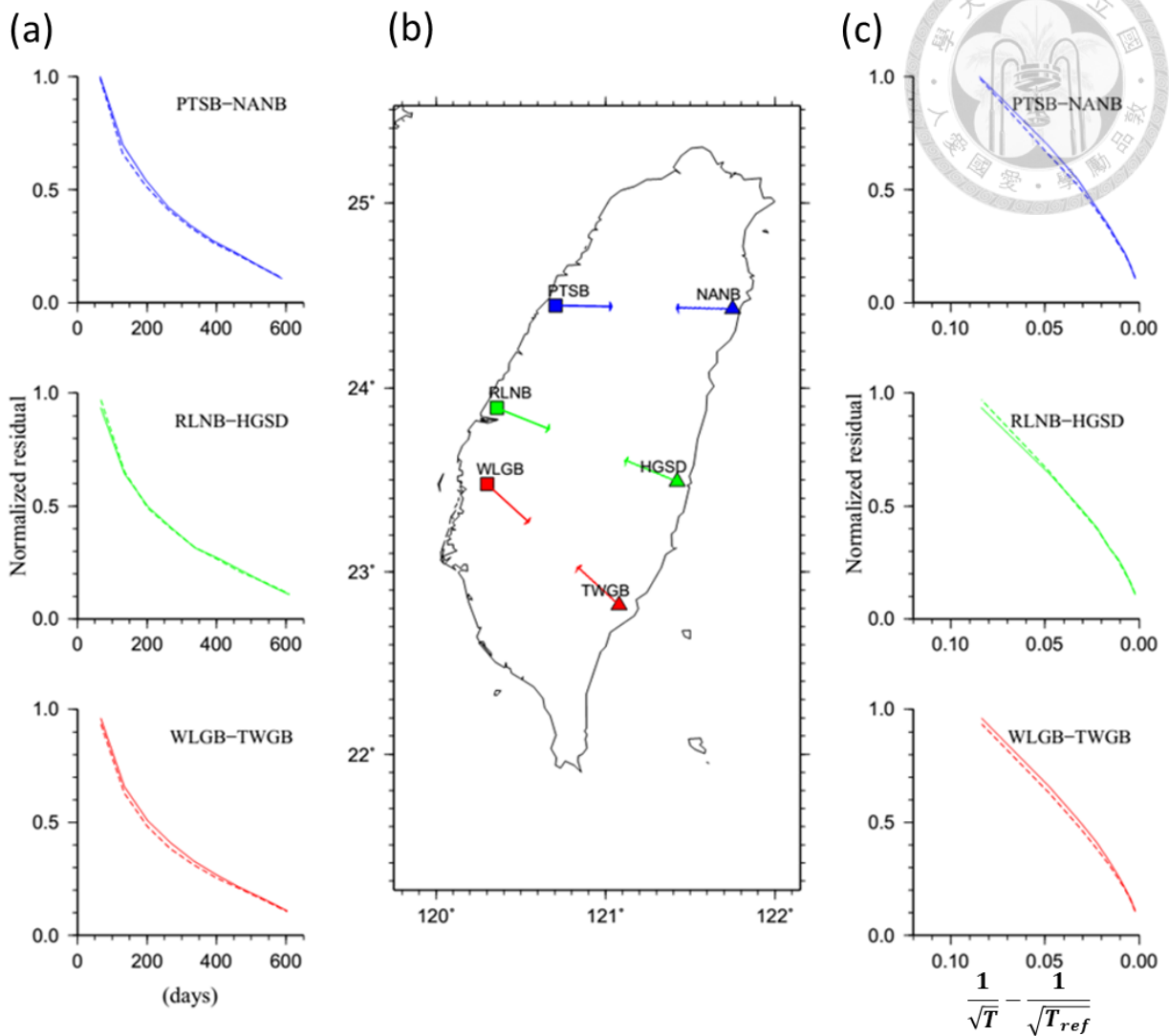
Figure 4-4: The evolution of remnant noise level of CCF for 6 representative coastal stations. The CCFs are band-pass filtered to extract the SPSM signals (2~5 seconds) (a) The remnant noise level as a function of correlation time. The solid and dash lines are related to energy coming from western and eastern direction, respectively. (b) Station distribution. (c) Remanent noise level as a function of the properly defined time-dependent term.
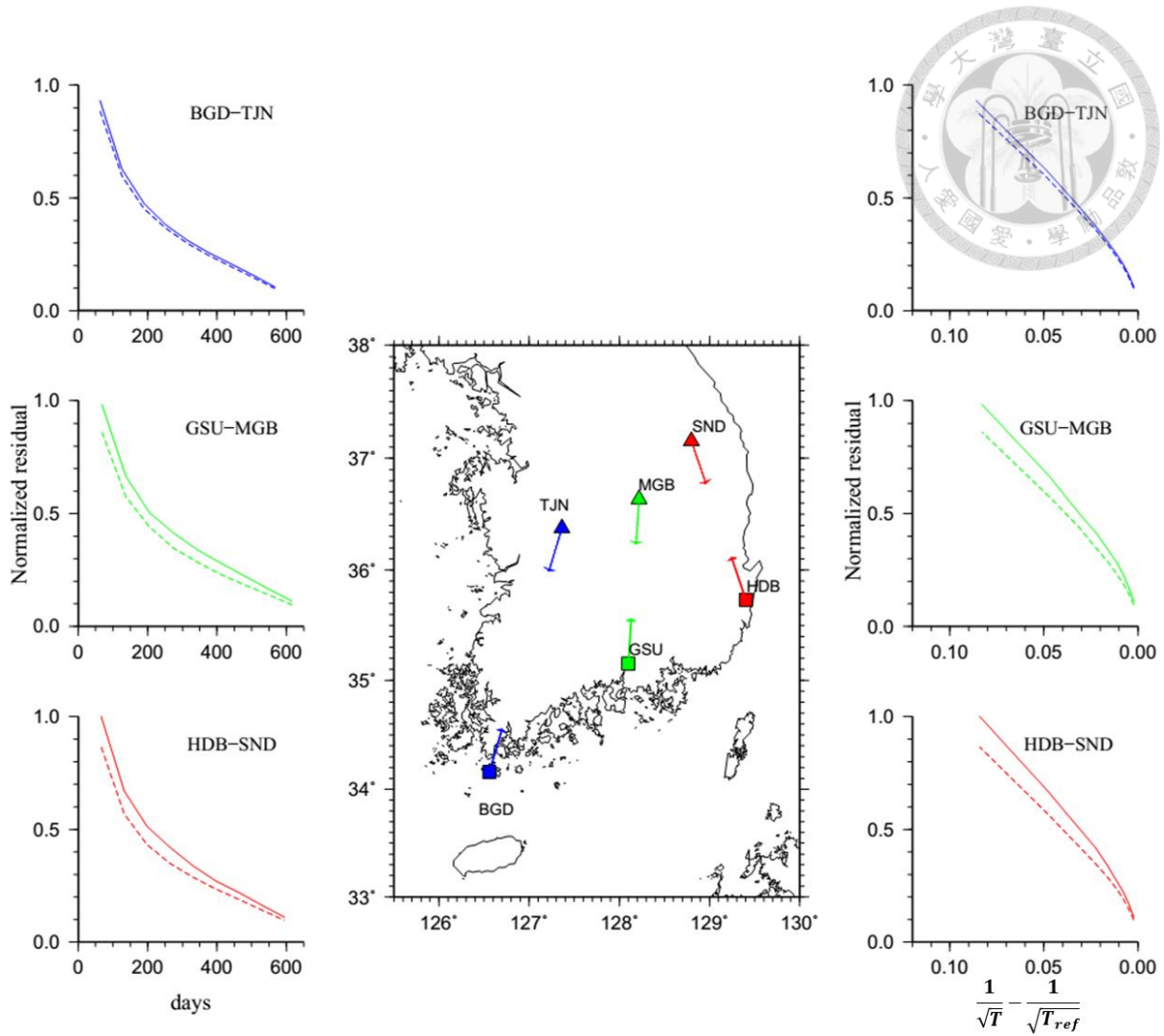
Figure 4-5: Same as Figure 4-4, but for cases in Korea. (a) Remanent noise level of *CCFs* (2~5 sec) decays with time. Solid and dash lines relate to energy coming from south (near coastlines) and north (inland) directions. (b) Stations used in this demonstration. (c) Remanent noise level varies with a properly defined time-dependent term.
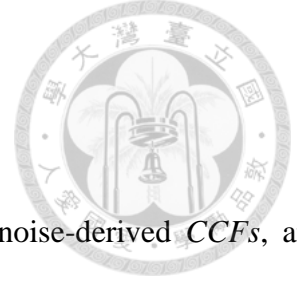
4.3. The noise level of *CCF*

We have defined the estimated original noise level, $\widehat{ONL}$, of the noise-derived *CCFs*, and confirmed that $\widehat{ONL}$ can be robustly evaluated using the randomly stacked *CCFs*. In the following, we compare the resulting $\widehat{ONL}$ derived in Taiwan and Korea, and discuss the applications and implications of the results.

4.3.1. The Exact Signal-to-Noise Ratio of *CCFs*

The noise-derived *EGFs* have been widely applied to seismic tomography, in which the major signals of *EGFs* are taken as fundamental mode surface waves. To ensure the reliability of the obtained dispersion, an empirically defined *SNR* is usually assigned to each *EGF* as a measure of data quality. The *SNR* of *EGFs* is generally defined as

$$\text{SNR} \approx \frac{W_s(t)\overline{A(\theta,\omega)}G(t)}{W_n(t)\overline{A(\theta,\omega)}G(t)} = \frac{W_s(t)G(t)}{W_n(t)G(t)}, \tag{4.4}$$

where $W_s(t)$ is the box function windowing the target major signals, $W_n(t)$ is the box function widowing the noises or the entire *CCF* trace, and *t* is the lag time. Notice that the source strength $\overline{A(\theta,\omega)}$ disappears in such expression. Moreover, considering the fact that amplitudes of major signals of *CCFs* generally decay with inter-station distance due to intrinsic attenuation and geometrical spreading, the above criteria favors data with shorter inter-station distance, as shown in Figure 4-6a. In other words, the implicit distance-dependent selection criteria may result in a biased
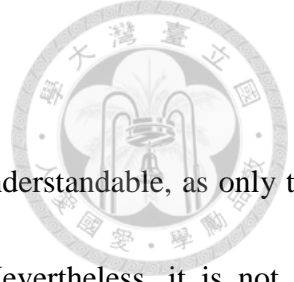
data collection.

Using the above simply defined *SNR* to measure *CCF* quantify is understandable, as only the fundamental mode surface waves are used as data in those studies. Nevertheless, it is not an appropriate approach to measure the relative strength between the signals and noises within the *CCFs* because signals and noises always coexist in *CCFs*, namely, there are noises in the signal window $W_s(t)$ and signals also exist in the noise window $W_n(t)$.

Recently, it has been recognized that the tailing waves of *CCF* are valuable. Studies have showed that *CCF* coda can be used for an iterative *CCF* reconstruction [*Froment et al.*, 2011; *Stehly et al.*, 2008], as mentioned in Section 4.1. Furthermore, *CCF* coda contains multiple-reflected P waves, although amplitudes of these signals are usually much weaker as compared to the dominant signal of *CCFs* [e.g., *Lin et al.*, 2013; *Poli et al.*, 2012a; *Poli et al.*, 2012b]. To ensure the data quality in the applications using the *CCF* tailing waves, we need an alternative measuring approach other than the one defined in Eq. (4.4). The $\widehat{ONL}(\omega)$ we discussed in Section 4.2 can be used for such a task.

With the obtained $\widehat{ONL}(\omega)$, the true noise level of the reference $CCF(\omega, T_{ref})$ can be estimated precisely by $\frac{\widehat{ONL}(\omega)}{\sqrt{T_{ref}}}$. Similar to Eq. (4.1), we may define the signal to noise ratio by

$$\widehat{SNR}(T_{ref}, \omega) \equiv \left\| CCF(\omega, T_{ref}) \right\| \cdot \sqrt{T_{ref}} / \widehat{ONL}(\omega). \tag{4.5}$$

Notice that the source property is still preserved in such expression. As shown in Figure 4-4a, the distance-dependence of *SNR* defined in Eq. (4.4) is much stronger than the ONL-based $\widehat{SNR}$. In

59
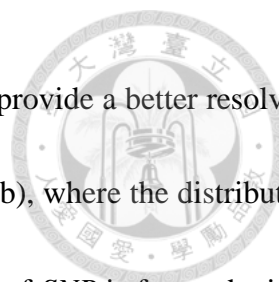
addition, compared to the empirically defined *SNR*, the ONL-based $\widehat{SNR}$ provide a better resolving

power for *CCF* quality. This is shown in their distribution map (Figure 4.4b), where the distribution

of the ONL-based $\widehat{SNR}$ spreads over a wide range, while the distribution of *SNR* is focused within

a narrow band.

ONL-based estimates are not restricted to the most predominant signal in the *CCF*, i.e., the

fundamental mode surface wave, but can be easily applied to any portion of the *EGF*. Using a

window function $W(T_w)$ prior to *ONL* measurement, the noise level of any segment can be

evaluated.

Moreover, since the $\widehat{ONL}$ can be estimated in the frequency domain, we may evaluate the

data quality at any given single frequency. This unique property is potentially important for

measurement using Aki's approach (see Chapter 2.4 for more details), in which the measurements

for both attenuation and velocity are conducted in the frequency domain.

Here, we present *CCF* convergence within four lag time windows, in which the major signal

and the corresponding tailing coda waves of causal and acausal *CCF* are considered separately

(Figure 4-7). Note that the decay of noises within coda windows also agrees with the theoretical

expectation. In addition, despite a strong amplitude asymmetry, the noise decays within two signal

windows are quite similar. As a matter of fact, there is nearly no direction dependence for the

obtained $\widehat{ONL}$ in Taiwan (Figure 4-7 & Figure 4-11), suggesting that the noise level of the causal

and acausal parts of *CCFs* are almost the same.
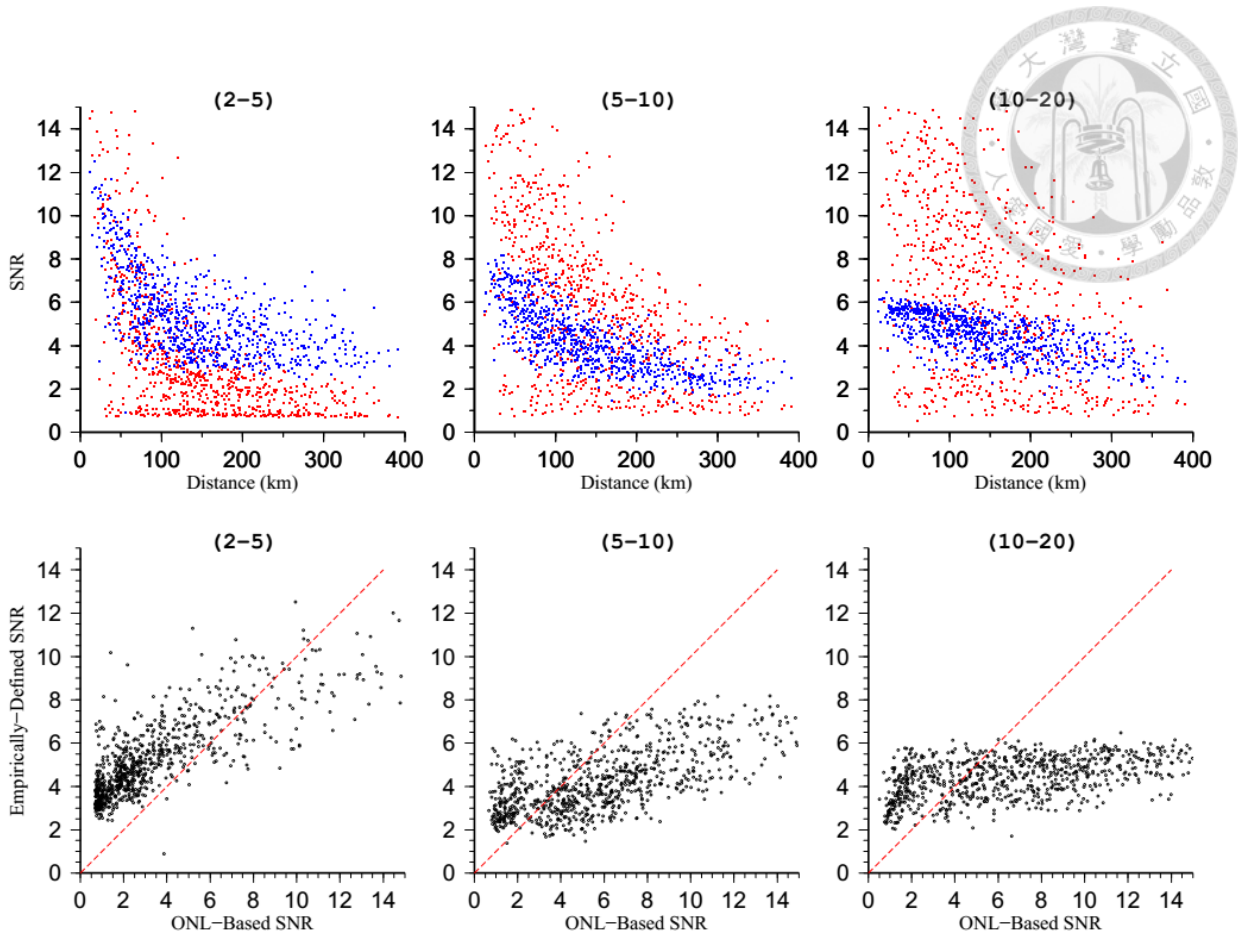
Figure 4-6: Comparisons of the empirical *SNR* measurements (blue) and the ONL-based *SNR* given in this study (red). The number in the parenthesis represents the corresponding frequency band. Here, we only present results obtained from *CCFs* in Taiwan. (a) Data quality estimates as a function of the interstation distance. (b) In general, empirical *SNR* and ONL-based *SNR* have a clear linear relationship. It is noticed that ONL-based *SNR* can provide a higher resolution for data selection, especially for the longer period cases. In this analysis, the empirically defined *SNR* is a ratio of the maxima amplitude within the signal window to the rms of the whole trace.

ANPB−YHNB 2006101 (2−5 sec; 59 km)

(a)

Signal

Ref(24)

18

12

6

1

lag time

(b)

Normalized residual

Causal Signal
Acausal Signal
Causal Coda
Acausal Coda

Hr
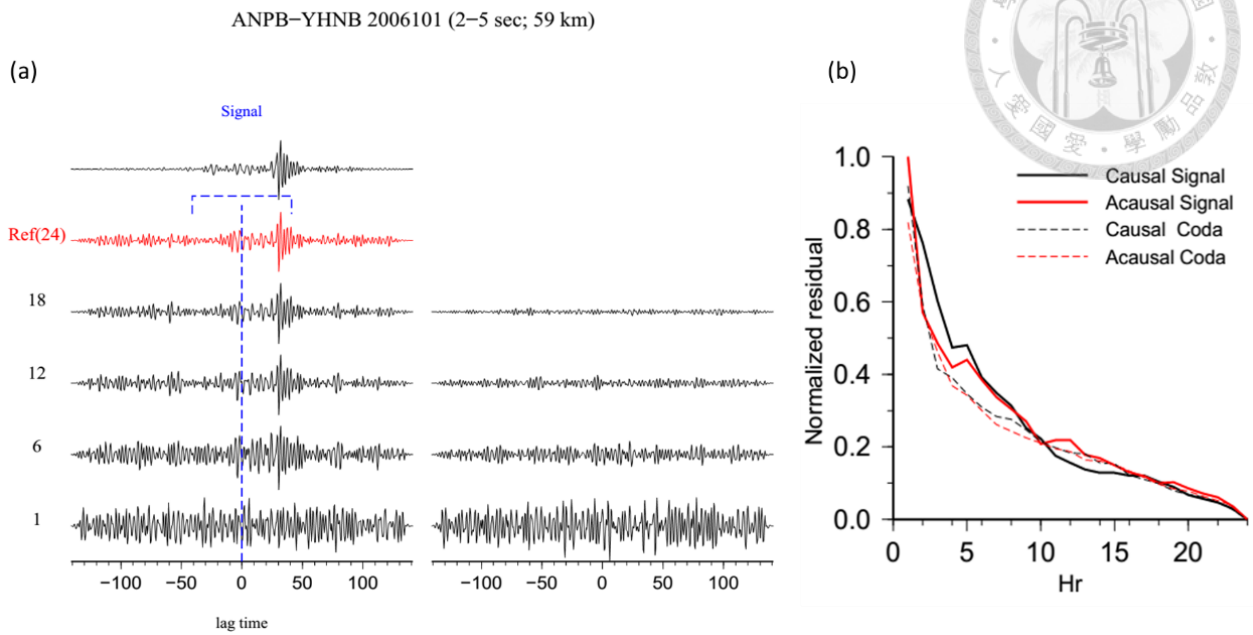
Figure 4-7: An example of *CCF* convergence within a day (2006,101). (a) The numbers next to *CCFs* are the total-correlation time. Using the *CCF* averaged over a day as a reference (marked in red), the corresponding waveform residual is shown in the right. The annual averaged *CCF* is also shown in the top for comparison. The blue dash lines denote the signal windows for the causal and acausal parts of *CCF* and the coda is defined as the rest of *CCF*. (b) Disregarding a miner temporal/spatial variation of source strength within a day, straight stacking is used in this analysis. Apparently, the noise level varies with total correlation time nicely.

4.3.2. On the Folding of *CCF*

In retrieving *EGFs* from noise cross-correlations, folding the causal and acausal signals of *CCF* is a common practice, and the idea is to improve the *SNR* of *EGFs*. Here we explore this intuitive processing from the point of view of ONL-based *SNR*, and show that folding does not necessarily improve the *EGF* quality.

Let A_1 and A_2 be the source strengths of causal and acausal direction, respectively, and $A_2/A_1 = n. (n > 1)$. Since the ONL is much more homogeneous than the background noise energy distribution, we can assume that the noise levels of the causal and acausal *CCF* are about the same, and are denoted by "Noise", thus, the original *SNR* of both causal and acausal *CCFs* can be written as

$$\begin{cases} SNR(A_1) = A_1/Noise \\ SNR(A_2) = A_2/Noise = nA_1/Noise. \end{cases} \tag{4.6}$$

The "folding" is equivalent to double the correlation time and to average the source strength from both directions, thus, the noise level and the source strength can be expressed as $Noise/\sqrt{2}$ and $(1 + n)A_1/2$, respectively, and the *SNR* for the folded *CCF* is simply $(\sqrt{2}(1 + n) \cdot A_1)/(2 \cdot Noise)$. With this, it's easy to show that *SNR* of the folded *CCF* is improved only when $n \leq 2.41$ (Figure 4-8).
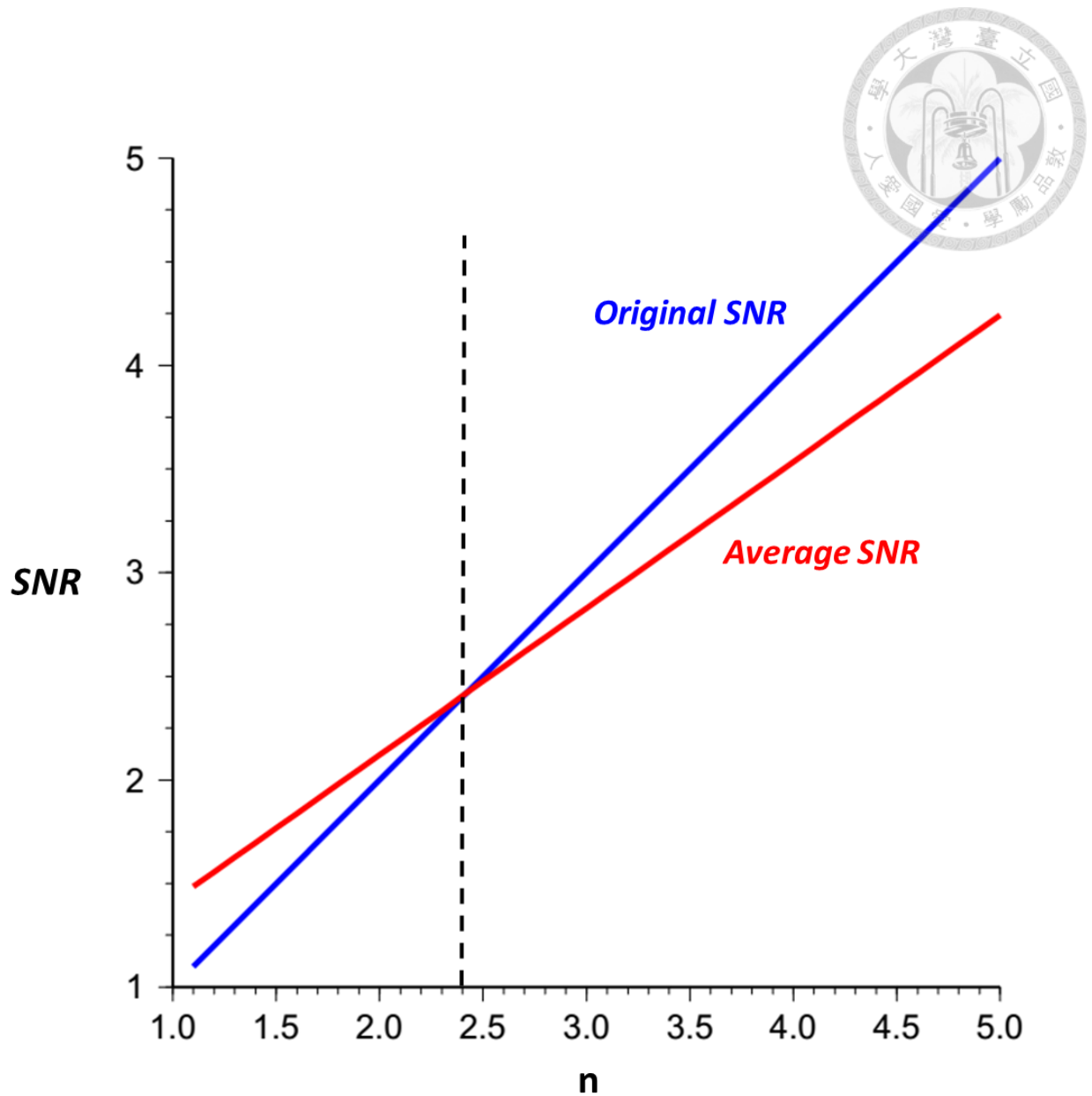
Figure 4-8: Comparison of the *SNR* of acausal *CCF* (blue), i.e., the one with better *SNR* before folding, and the *SNR* of folded *CCF* (red) as a function of relative strength n. Note that the SNR can be improved by folding only in cases when the relative strength (n) is smaller than 2.41.
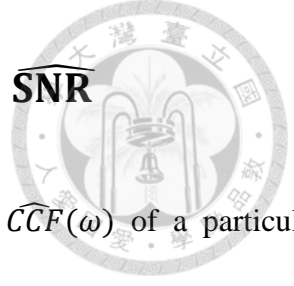
4.3.3. Expected Temporal Evolution of $\widehat{ONL}$-based $\widehat{SNR}$

Once the characteristic $\widehat{ONL}(\omega)$ and an expected CCF spectrum $\widehat{CCF}(\omega)$ of a particular area are given, we may evaluate the $\widehat{ONL}$-based $\widehat{SNR}$ as a function of cross-correlation time T. Here we compare results of such analysis for Taiwan and Korea and discuss their implications.

For simplicity, we only compare the average estimates for the resulting $CCFs$ in Taiwan and Korean respectively. To average station pairs with different inter-station distances, we have corrected for the geometrical spreading effects for the expected CCF spectrum. On the other hand, the $\widehat{ONL}(\omega)$ is only sensitive to the source population, thus, the characteristic $\widehat{ONL}(\omega)$ is directly evaluated by averaging all the individual $\widehat{ONL}(\omega)$. Consequently, rewriting Eq. (4.5) as

$$\widehat{ONL} - \widehat{based}\ SNR = \widehat{CCF}(\omega) \cdot \frac{\sqrt{T}}{\widehat{ONL}(\omega)}. \tag{4.7}$$

In figure 4-9, we compare the expected CCF spectrum, characteristic $\widehat{ONL}(\omega)$ spectrum and the corresponding $\widehat{SNR}$ in Taiwan and Korea. The expected CCF spectrums derived from these two area are rather similar; however, strength of the characteristic $\widehat{ONL}(\omega)$ are different, especially for shorter periods, in which the $\widehat{ONL}(\omega)$ in Korea is much smaller, and this is well demonstrated in the temporal evolution of SNR (right panel of Figure 4.9). Notice that the $\widehat{SNR}$ presented here is evaluated assuming inter-station distance is 1 km, therefore, the prediction for CCF with distance D should be corrected by $\widehat{SNR}(\omega, T, D) = \left(\frac{\widehat{CCF}(\omega)}{\sqrt{D}}\right) \cdot \left(\frac{\sqrt{T}}{\widehat{ONL}(\omega)}\right)$
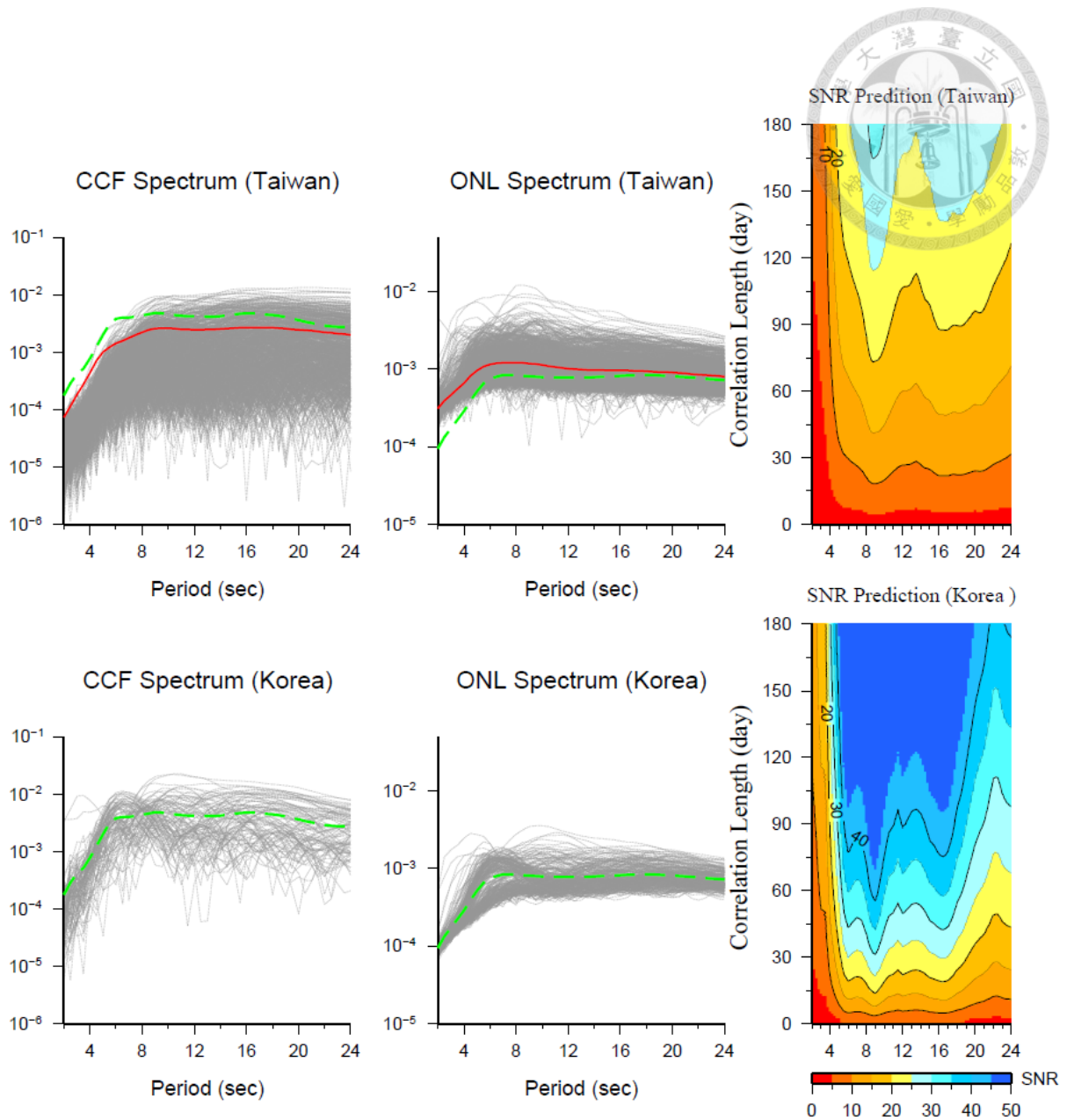
Figure 4-9: Spectrum for each vertical component *CCF* and their corresponding $\widehat{ONL}(\omega)$ in Taiwan and Korea. The red lines are mean spectrum, and the mean for the case in Korean, indicated by green dash lines, is also shown for a comparison. (Right column) The estimated *SNR* as a function of total correlation time in Taiwan and Korea. The effects of inter-station distance are removed by taking into account of the geometrical spreading effects.

4.4. A Constraining the noise source property using ONL

Since microseisms are the major sources of the noise-derived *CCFs*, it is expected that we may investigate the characteristics of the nearby or distant ocean waves by examining the noise-derived *CCFs*.

In Figure 4-10, we present the average spectrum of continuous data recorded by 37 broad-band stations in Taiwan and the derived *CCFs*, and compare them with the records of significant wave periods (Ts) of 5 buoys around Taiwan. To preserve the original strengths in the records, instead of spectrum whitening, 1-bit normalization is applied to the raw data in this comparison. Both the background seismic noises and *CCFs* demonstrate clearly that the dominant microseisms are distributed around the periods from 3-5 seconds, which is the well-known SPSM, as its energy peak (4 seconds) is about half of the period of the coastal ocean waves (Figure 4-10). The other significant energy of *CCFs* is located around 16-24 seconds, which is related to the distant primary microseisms (DPM, ~16-24 sec).

On the other hand, records from buoy indicate that the Ts of local ocean waves is about 8 seconds. Strangely, while the ocean waves with this characteristic period are widely distributed around the nearby oceans of Taiwan, the associated microseism, i.e., primary microseism (PM, ~7-8 sec), are not present in the *CCFs* or the ambient seismic noises.

In this section, we show that, the PM signature can be detected in *ONL* spectrum for its high
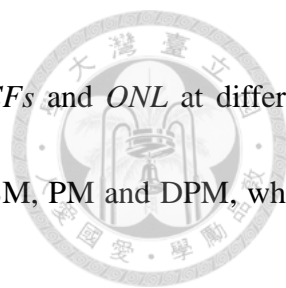
source population. Moreover, by comparing the relative strengths of *CCFs* and *ONL* at different

period range, we may clearly distinguish the source characteristics of SPSM, PM and DPM, which

are not accessible by simply analyzing *CCF* signals alone.
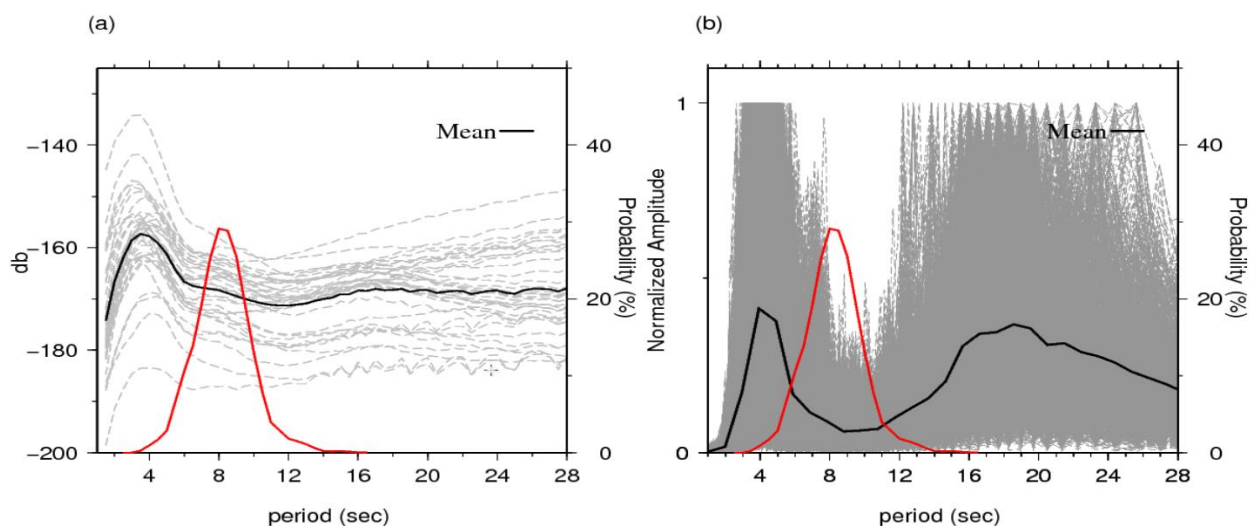
Figure 4-10: (a) Comparison of the spectrum of background seismic noises and the distribution of the significant ocean wave periods (Ts) around Taiwan in 2007. The gray dash lines are the annual average spectra of each seismic station used in this study, and the mean spectrum of all stations is shown by black line. The red line is the distribution curve of average Ts of 5 buoys deployed in international harbors of Taiwan (http://isohe.ihmt.gov.tw/index_eng.aspx), and the associated probabilities are shown in the left y-axis. (b) Same as Figure 4-10(a), but the comparison is made for the *CCF* spectra. All the spectra are normalized to emphasize their frequency contents. The gray lines show the spectra of annual *CCF* stacks for all the available pairs, and black curve represents the corresponding mean spectra. To preserve the original strengths in seismic records, instead of spectrum whitening, 1-bit normalization is applied to the raw data in this comparison.

4.4.1. Signatures of PM, SPSM and DPM Revealed by *ONL* and *CCF* Spectra

According to our earlier discussion, the strength of *CCF* signals are sensitive to source excitation and the source population within the effective sensitivity zone; on the other hand, *ONL* is primarily sensitive to the overall source population (Figure 4-11). Although the coupling strength from ocean waves to solid earth is highly influenced by the water depths, the source population of PM is expected to be high around Taiwan and such phenomena should be manifested in the $\widehat{ONL}$.

Indeed, there is a clear peak around 8 sec in the average $\widehat{ONL}$ spectrum, and it coincides with the most dominant ocean wave period observed around Taiwan (Figure 4-11). From the spectrum of *CCF* and $\widehat{ONL}$, we learn that the source population of PM is high, but the wave energy is hardly coupled into seismic ambient noises. This is totally understandable, mainly because the wave-induced pressure on the ocean bottom decays very quickly with water depths (Figure 4-12).

Similar comparison can be made to analyze SPSM and DPM. The *CCF* spectrum (Figure 4-13) clearly show that there are two energy peaks centering on the significant periods of SPSM and DPM, respectively.

The strong excitation of SPSM in *CCF* is accompanied with a relative low *ONL*, indicating the source population of SPSM is low. In the case of Taiwan, while the water depths in the Taiwan Strait is much shallow, results of $\widehat{ONL}$ suggest that the occurrence of SPSM excitations are likely
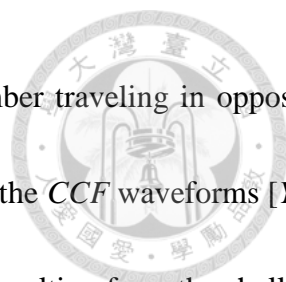
confined to the near-coast area where two waves with similar wave number traveling in opposite direction may meet each other. The strong amplitude asymmetry shown in the *CCF* waveforms [*Y-N Chen et al.*, 2011] is thus mainly related to the more energetic excitation resulting from the shallow water condition in the west coast.

Compared to PM, DPM is characterized by stronger excitation shown in *CCF* spectrum (Figure 4-13). We infer that this is because their effective wave-induced pressure on the ocean bottom may extend to much larger depths than PM, allowing a stronger coupling in the same coastal condition (Figure 4-12). The fact that signature of DPM is invisible is in the near-coast ocean wave observation confirms that these long periods microseisms are triggered in oceans far away from the island.
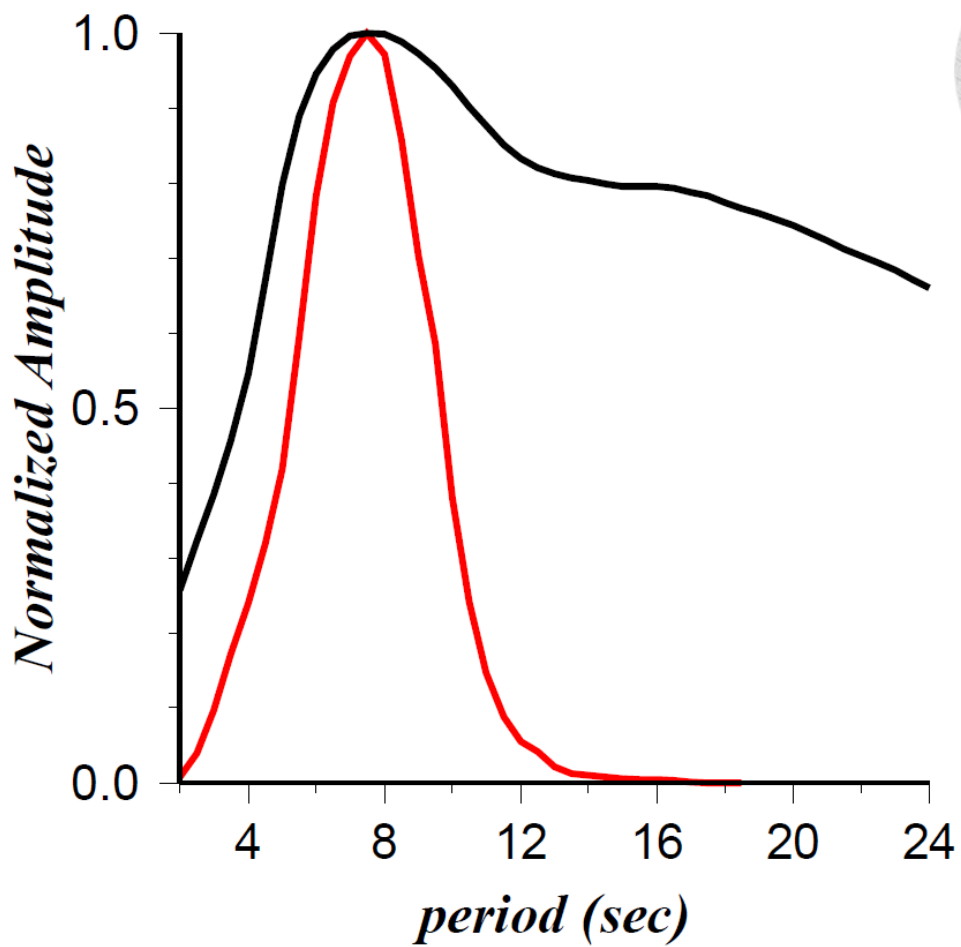
Figure 4-11: Comparison of the average *ONL* spectrum of Taiwan (black) and the distribution curve of average Ts of 5 buoy deployed in international harbors of Taiwan (the same as the one shown in figure 4-10b). The amplitudes are normalized for a better comparison.
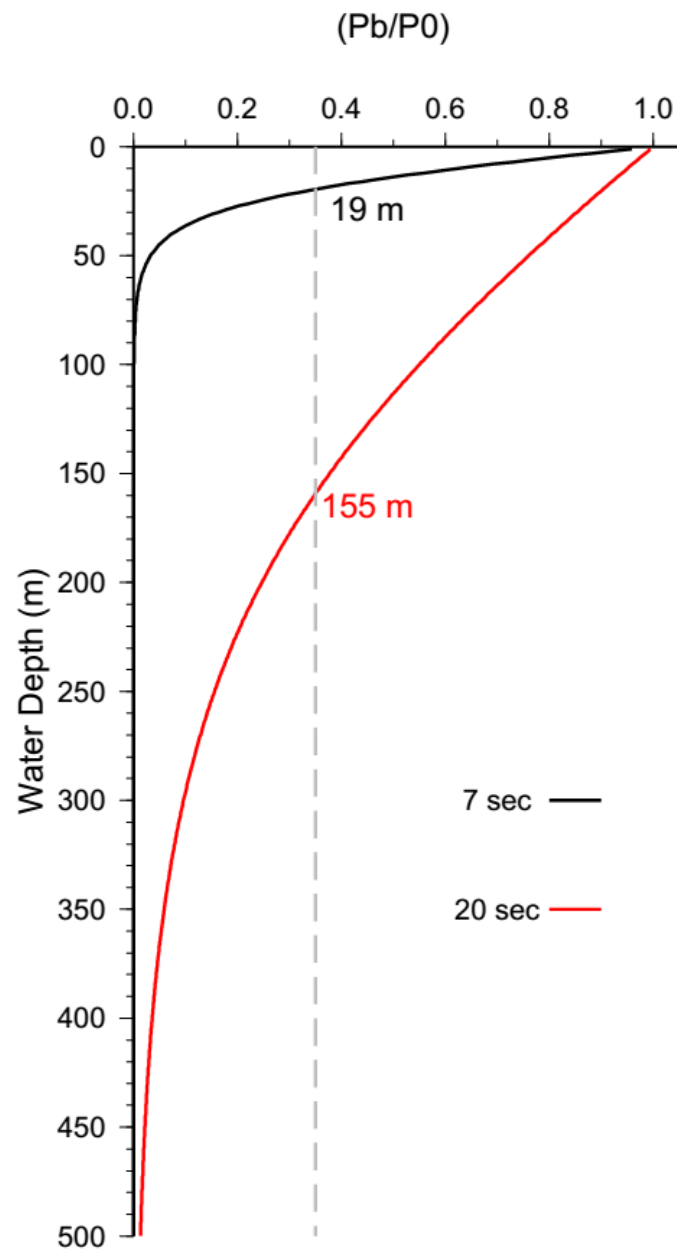
Figure 4-12: The pressure at the ocean bottom from hydrodynamic forcing, *Pb*, relative to that of

sea surface, *P0*, varies with depth h according to $\frac{Pb}{P0} = \frac{1}{cos(kh)}$, where k is the wave number of

the gravity wave. The efficient maximun excitation depth is defined by a 35% remanent pressure

at sea bottom (gray dash line), corresponding to the depths 19 and 155m for the 7 and 20 sec

microseisms, respectively [*Bromirski and Duennebier*, 2002].
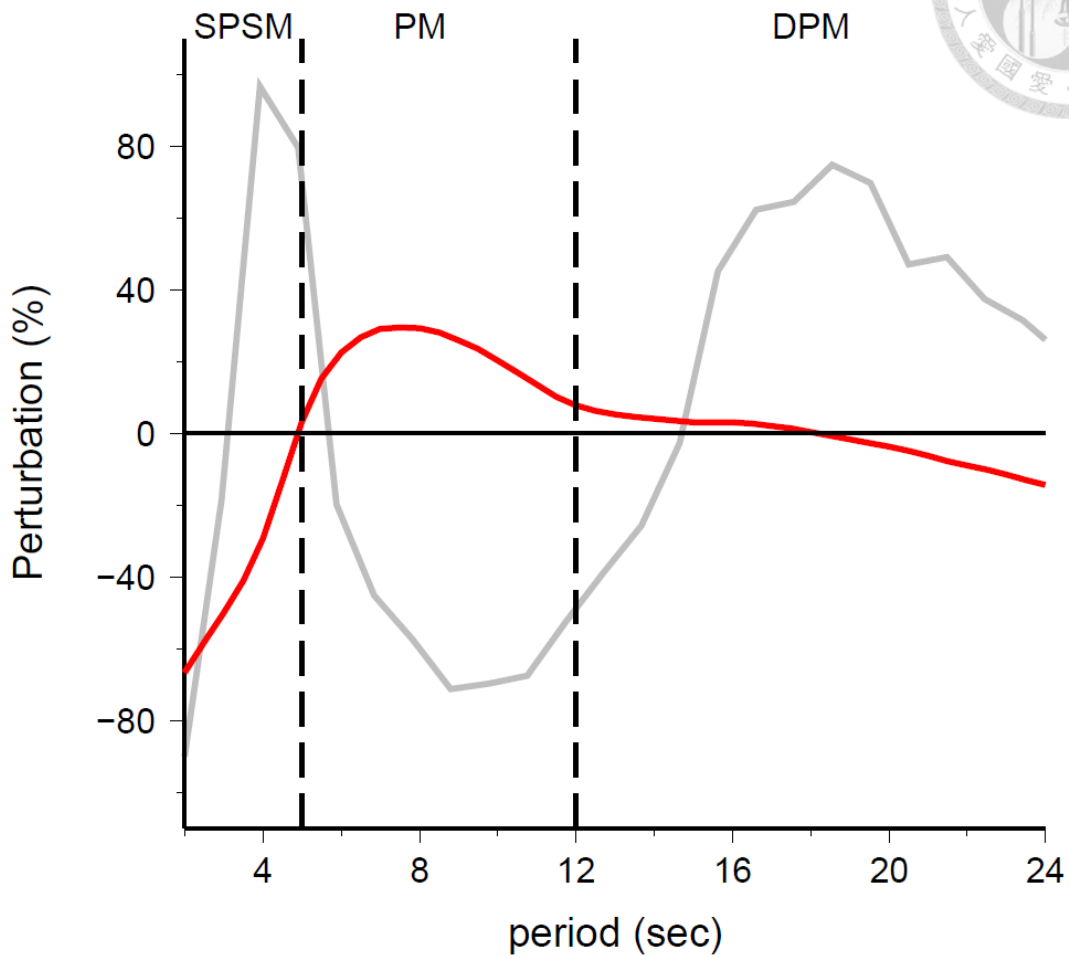
Figure 4-13: Comparison of *CCF* spectrum (gray line) and *ONL* spectrum (red). In both case, the

mean value is removed and the strength is shown by the perturbations w.r.t. the mean. Dashed lines

mark the representative period ranges of SPSM, PM and DPM.
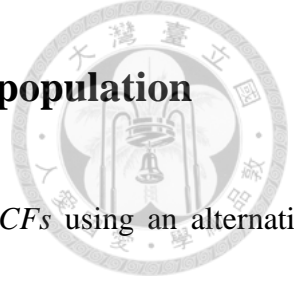
4.4.2. On the PM Energy and the associated source population

Here we further inspect the effects of PM source population on *CCFs* using an alternative approach. The weak PM energy in *CCF* can be amplified by taking the spectrum whitening (coherency) prior to the cross-correlation. In such operation, *CCFs* strengths are proportional to the ratio of the source power within sensitivity zone to the overall source power (equation 2.22); thus, we may better examine the energy variations of PM in the spectrum domain. To demonstrate the relationship between the source population and the associated PM strength, in what follows we present the coherency retrieved from Korea stations as an example.

The $\widehat{ONL}(\omega)$ observed in Korea (Figure 4-9) can be separated into two groups according to the characteristic behavior (Figure 4-14), with (group 2) or without (group 1) a peak around the period ~6-8 sec, which is exactly the energy band of the near-coast PM. Interestingly, similar phenomena is found in their corresponding *CCF* spectrum. While $\widehat{ONL}$ is primarily sensitive to the source population, *CCF* is sensitive to both the source population and excitation strength which in turn is related to wave height and water depths. The observed significant positive correlation between $\widehat{ONL}$ and *CCF* spectrum around the PM band suggests that the dominant factor for *CCF* PM energy is the source population, not the source excitations. In other words, the bathymetry of surrounding ocean has little influence on the variations of *CCF* amplitude around the PM band.
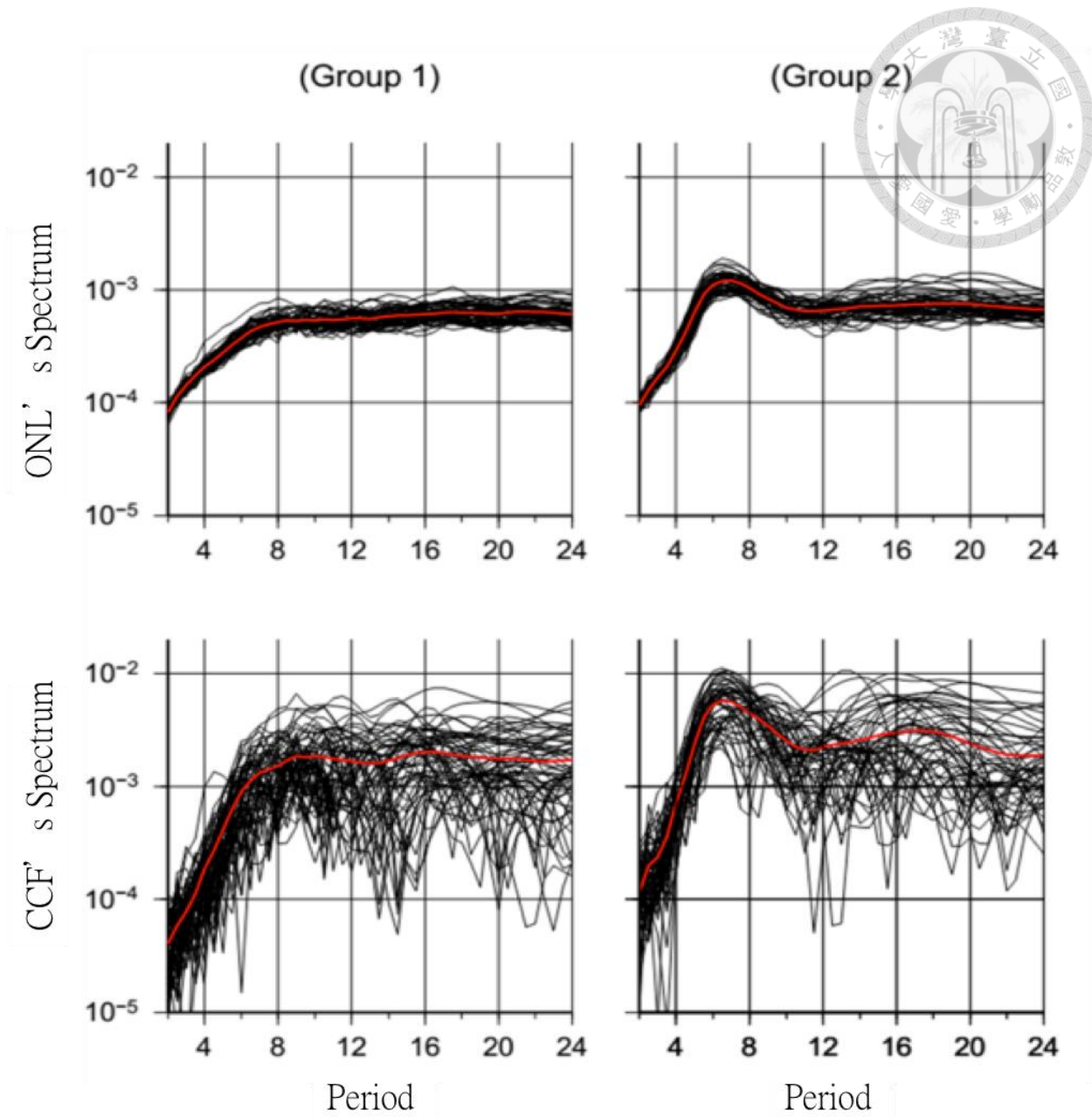
Figure 4-14: *CCF's* spectrum derived from spectrum whitening and the corresponding $\widehat{ONL}$ in Korea, where the red lines are mean. Apparently, $\widehat{ONL}$ spectrum can be separated into two groups according to the characteristic behavior (Figure 4-14), with (group 2) or without (group 1) a peak around the period ~6-8 sec, which is exactly the energy band of the near-coast PM.
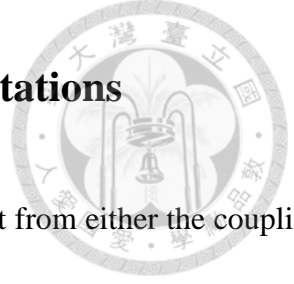
4.4.3. The origin of microseism observed by island stations

It is noticed that the microseisms observed around 5-10 sec may result from either the coupling of local ocean waves (5-10 sec) with the ocean bottom in the shallow water environment directly (PM), or the second-order pressure fluctuations from a distant swell of longer period 10-20 sec, namely the LPSM [e.g., *Bromirski and Duennebier*, 2002; *Bromirski et al.*, 2005]. *Bromirski et al.* [2005] show that a combination of both seismic data recorded by ocean bottom seismometers (OBS) and ocean wave observation (or wave modeling) may provide an explicit answer for the microseism origination. Without OBS records, here we attempt to constraint the microseism origin of this puzzling frequency band by examining the correlations between $\widehat{ONL}$ and CCF amplitudes.

The dominant excitation mechanism of microseisms can be distinguished through the statistical distribution of the ratio between the $\widehat{ONL}$ and *CCF's* strength ($ratio(\omega) \equiv \frac{\widehat{ONL}(\omega)}{CCF(\omega)}$). If the source population is the dominant factor in the strength of CCF amplitudes, the high correlation between $\widehat{ONL}$ and *CCF's* strength can be identified by the resulting concentrated distribution of the ratio; otherwise, the values of such ratio would be widely scattered.

The distribution of $\frac{\widehat{ONL}(\omega)}{CCF(\omega)}$ ratio at three representative frequency bands are shown in Figure 4-15. The selected frequency bands are properly chosen for SPSM (3~5 sec), PM (6~8 sec) and DPM (12~24 sec), respectively.
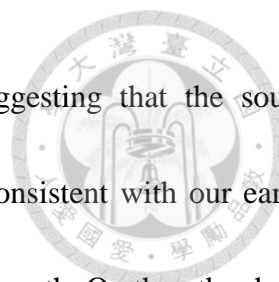
In the SPSM band, the ratio is relatively widely distributed, suggesting that the source population is not the dominant factor for the CCF amplitudes. This is consistent with our earlier conclusion that the bathymetry plays an important role in the excitation strength. On the other hand, for the PM and DPM, coherent behavior for CCF amplitudes and ONL are observed, suggesting that the source population is the main controlling factor for the CCF amplitudes. Moreover, the good correlation between CCF amplitudes and ONL in stations from Taiwan and Korea confirms that the corresponding microseism is PM rather than LPSM.

Additional constraint to locate the origin of the microseisms is their temporal variations w.r.t. the local weather condition. We compare the CCF energy in two frequency ranges, 2-8 seconds and 12-24 seconds, with the coastal wind speeds (Figure 4-16). Without much surprise, the expected high correlation between wind speeds and 2-8 seconds microseisms is confirmed in this comparison, and the he correlation is poor in the frequency band for the DPM, i.e., 12 – 24 seconds. The results further solidify the conclusions about their source origins.
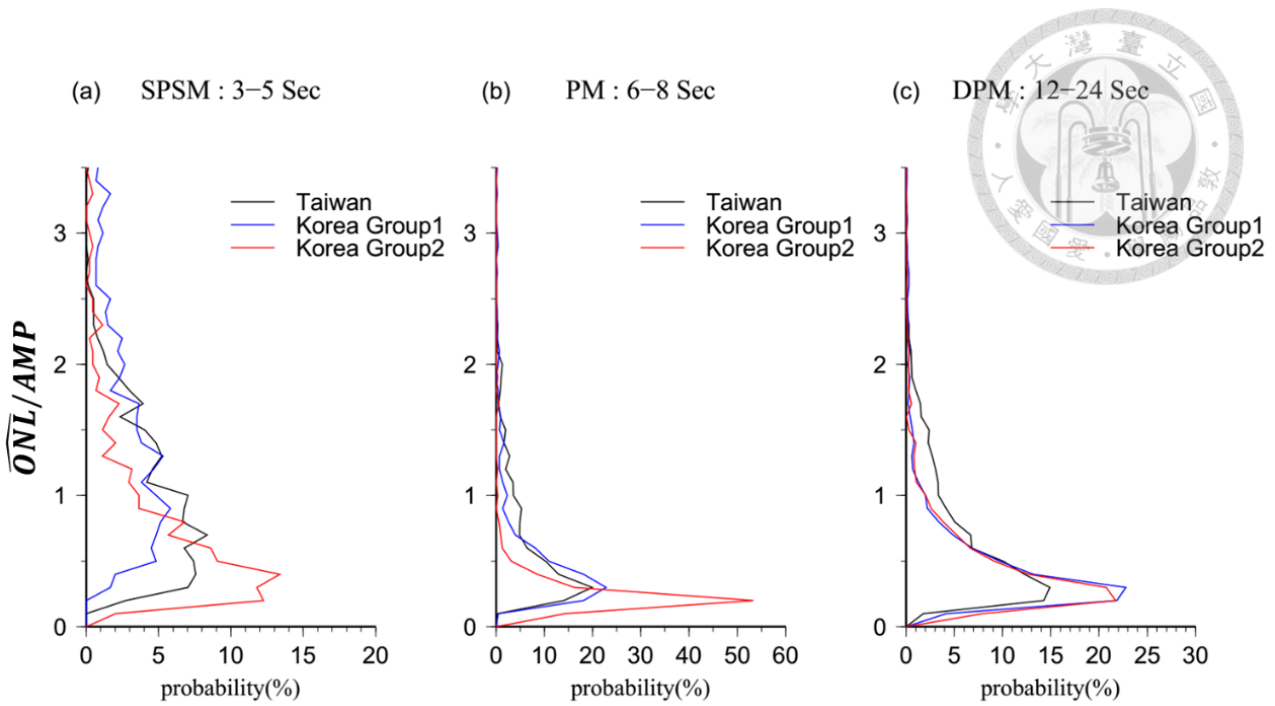
Figure 4-15: Statistical results of the ratio between the $\widehat{ONL}$ and the associated *CCF* strength.

Notice that the $\widehat{ONL}$ and the associated *CCF* strength have a clear positive correlation in the PM

and DPM frequency band, suggesting that the microseism power is closely related to the source

population. On the contrary, such correlation does not exist in the SPSM frequency band, and

further confirms the conclusion made in Section (4.4.2) that the special property of SPSM around

Taiwan is highly correlated to the bathymetry.
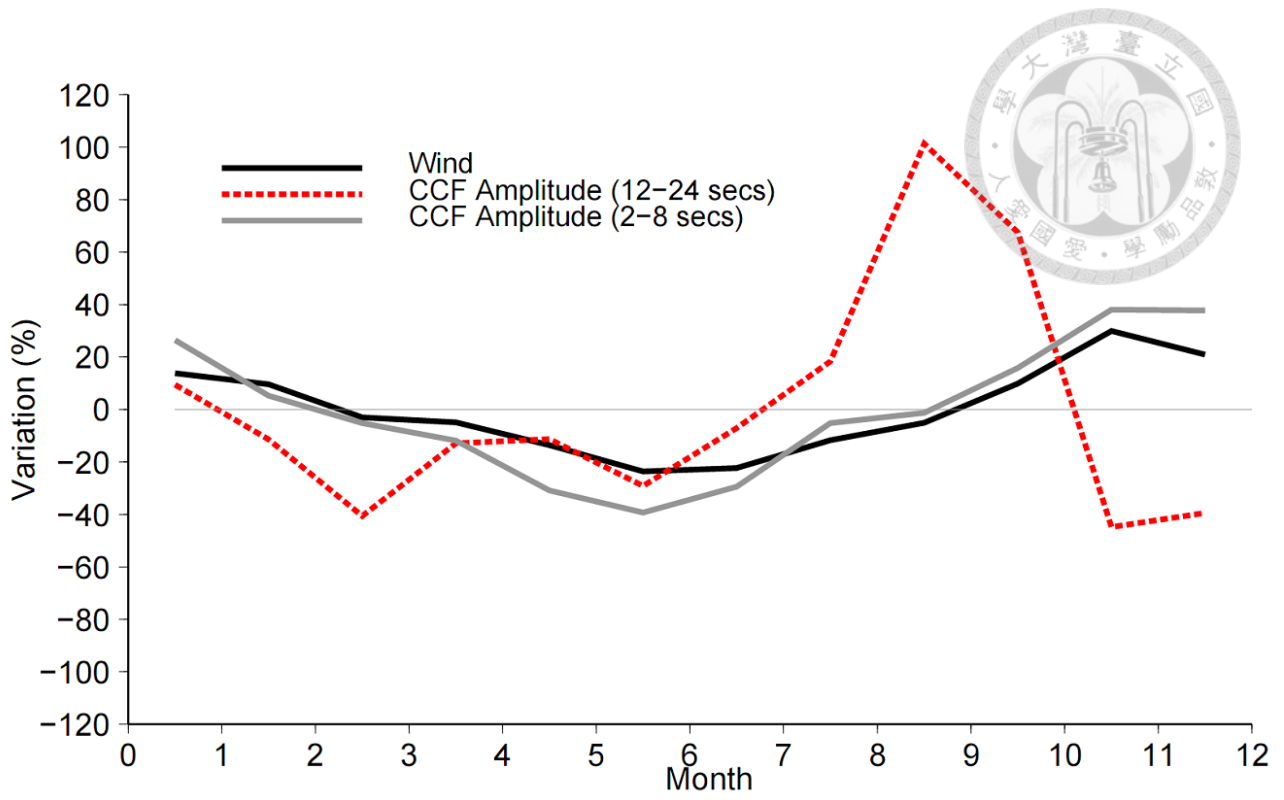
Figure 4-16: Same as Figure 3-4, except for the *CCFs* retrieved from broadband data in 2007.

Chapter 5

Conclusions

In this thesis, we have detailed formulations for relevant theories, devise a recipe to evaluate the noise level of *CCFs* quantitatively, and point out that the derived $\widehat{ONL}$ is related the source population. We then apply the method to realistic *CCFs* derived in Taiwan and Korea. The feasibility of our proposed method has been confirmed by the fact that the obtained temporal evolution of *CCF* noise level follows the theoretical prediction nicely.

We show that, unlike the empirically defined *SNR*, the $\widehat{ONL}$-based *SNR* estimates are much less dependent on inter-station distance, and may thus provide a better resolving power for *CCF* quality. Moreover, such measurement can be similarly applied to any time windows of the *CCF* trace to evaluate the quality of overtones or coda waves, which are potentially important when taking *CCFs* as seismic data for broader applications.

We should point out that the stable temporal source condition is assumed and empirically satisfied in the estimating procedure, thus, the resulting $\widehat{ONL}$ is insensitive to the source distribution, namely, the expected temporal evolution of noise level can be equally achieved even the source distribution is highly non-homogeneous. Example in Figure 5-1 represents such a particular case, in which the same operation is applied to a single day *CCF* except that the random stacking is not used. The resulting temporal evolution of noise level is similar with those shown in
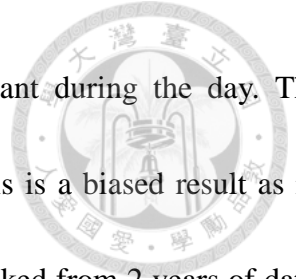
Chapter 4, implying that the source condition remains relatively constant during the day. The estimated *SNR* (2.92) also suggests that it is a quality *CCF*, however, this is a biased result as its major signals are very different from those in the final reference *CCF* stacked from 2 years of data. Obviously, the discrepancy is caused by strong inhomogeneous source distribution. To guarantee a reliable measurement using this approach, a reliable reference *CCF* is important, i.e., a significant long correlation time is necessary to achieve a quasi-diffused wave field.

The strength of *CCF* signals are sensitive to source excitation and the source population within the effective sensitivity zone; on the other hand, *ONL* is primarily sensitive to the overall source population. Combination of both measurements may help us to better constrain the properties of microseisms. In this thesis, we have reported new constraints on the microseisms around Taiwan and Korea using this method: (1) The positive correlation between *CCF* strength and $\widehat{ONL}$ in the period 5~10 sec suggests that these dominant oscillations observed in Taiwan and Korea are mostly contributed by PM, rather than LPSM proposed by previous studies; (2) The correlation between *CCF* strength and $\widehat{ONL}$ is much weaker in the SPSM frequency band, suggesting that the high SPSM level in Taiwan Strait is mainly caused by the bathymetry effect; (3) The low $\widehat{ONL}$ in the SPSM band implies that sources for these dominant signals in *CCFs* are likely confined in the near-coast region; (4) Records from buoys show that the most significant period of local ocean waves is about 8 seconds around Taiwan. The expected high source population of PM is well demonstrated by the strong $\widehat{ONL}$ signals in the period around 6-9 seconds, although the PM signals

are invisible in the *CCF* records or the background seismic noises.

There are other potential applications of $\widehat{ONL}$. For instance, we could improve the temporal resolution of the source strengths, and investigate the tidal effect on the source excitations. To begin with, the source strength as a function of time T is rewritten as a perturbation ($\alpha(T)$) with respect to the reference one $(\overline{A(T_{ref})})$: $A(T) = (1 + \alpha(T))\overline{A(T_{ref})}$. Notice that the discrepancy ($Diff(T)$) between the Noise level derived from the $\widehat{ONL}$ and that from the straight stacked *CCFs* represents the temporal variation of sources (Figure 4-3), and consequently it can be expressed as $Diff(T) = \frac{1}{T}\int \alpha(T))\overline{A(\theta, T_{ref})}dT$. Now we could obtain the detailed evolution of source strength through a simple relation: $\alpha(T))\overline{A(\theta, T_{ref})} = \frac{d(T*Diff(T))}{dT}\Big|_T$.

In addition to investigations of microseisms sources, we also notice that using $\widehat{ONL}$, the noise-derived *EGF*s can be further divided into two segments, the major signals and their tailing trains, coda waves. Unlike the $\widehat{ONL}$ of the signal window that is sensitive to the population of ocean microseisms source, the $\widehat{ONL}$ of the coda window remains nearly constant in direction and only simply decreases with frequency. We then suggest that the origins of the *CCF* coda are closely related to the scatter effect. In other words, structures are much heterogeneous at higher frequencies, and consequently the associated stronger scatter effect is beneficial to mature the *EGF*s coda. This indicates that there are two origins in one *CCF*, the oceanic microseisms sources and the "structure-induced" sources. As a result, we may probe the scatter properties from the coda of the noise-derived *EGF*s.

ANPB−YHNB 2006001 (2−5 sec; 59 km)
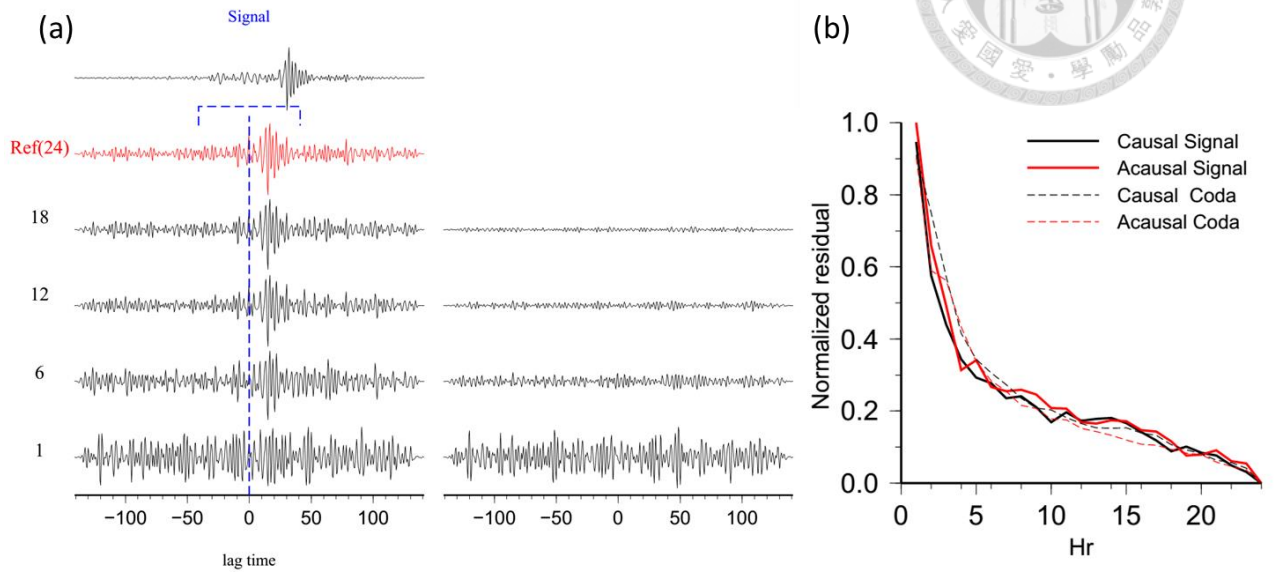
Figure 5-1: Same as Figure 4-7, except for the *CCF* convergence within a different date (2006,001).

The $\widehat{ONL}$-based *SNR* of the causal *CCF* is about 2.92.

Bibliography

Aki, K. (1957), Space and time spectra of stationary stochastic waves with special reference to microtremors, *Bull. Earthquake Res. Inst.Tokyo Univ*, *35*, 415-456.

Asten, M. W. (2006), On bias and noise in passive seismic data from finite circular array data processed using SPAC methods, *Geophysics*, *71*(6), V153-V162.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007), Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophysical Journal International*, *169*(3), 1239-1260.

Boschi, L., C. Weemstra, J. Verbeke, G. Ekström, A. Zunino, and D. Giardini (2013), On measuring surface wave phase velocity from station–station cross-correlation of ambient signal, *Geophysical Journal International*, *192*(1), 346-358.

Boué, P., P. Poli, M. Campillo, and P. Roux (2014), Reverberations, coda waves and ambient noise: Correlations at the global scale and retrieval of the deep phases, *Earth and Planetary Science Letters*, *391*(0).

Brenguier, F., N. M. Shapiro, M. Campillo, V. Ferrazzini, Z. Duputel, O. Coutant, and A. Nercessian (2008), Towards forecasting volcanic eruptions using seismic noise, *Nature Geosci*.

Bromirski, P. D. (2001), Vibrations from the "Perfect Storm", *Geochemistry, Geophysics, Geosystems*, *2*(7), 1030.

Bromirski, P. D., and F. K. Duennebier (2002), The near-coastal microseism spectrum: Spatial and temporal wave climate relationships, *Journal of Geophysical Research: Solid Earth*, *107*(B8), ESE 5-1-ESE 5-20.

Bromirski, P. D., R. E. Flick, and N. Graham (1999), Ocean wave height determined from inland seismometer data: Implications for investigating wave climate changes in the NE Pacific, *Journal of Geophysical Research: Oceans*, *104*(C9), 20753-20766.

Bromirski, P. D., F. K. Duennebier, and R. A. Stephen (2005), Mid-ocean microseisms, *Geochemistry, Geophysics, Geosystems*, *6*(4), Q04009.

Brzak, K., Y. J. Gu, A. Ö keler, M. Steckler, and A. Lerner-Lam (2009), Migration imaging and forward modeling of microseismic noise sources near southern Italy, *Geochemistry, Geophysics, Geosystems*, *10*(1), Q01012.

Campillo, M., and A. Paul (2003), Long-Range Correlations in the Diffuse Seismic Coda, *Science*, *299*(5606), 547-549.

Cassereau, D., and M. Fink (1992), Time-reversal of ultrasonic fields. III. Theory of the closed time-reversal cavity, *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, *39*(5), 579-592.

Cato, D. H., and S. Tavener (1997), Ambient Sea Noise Dependence on Local, Regional and

Geostrophic Wind Speeds: Implications for Forecasting Noise, *Applied Acoustics*, *51*(3), 317-338.

Cessaro, R. K. (1994), Sources of primary and secondary microseisms, *Bulletin of the Seismological Society of America*, *84*(1), 142-148.

Chen, Y.-N., Y. Gung, S.-H. You, S.-H. Hung, L.-Y. Chiao, T.-Y. Huang, Y.-L. Chen, W.-T. Liang, and S. Jan (2011), Characteristics of short period secondary microseisms (SPSM) in Taiwan: The influence of shallow ocean strait on SPSM, *Geophysical Research Letters*, *38*(4), L04305.

Chen, Y., Y. Gung, S. You, L. Chiao, W. Liang, and C. Lin (2009), On short period ambient noise of Taiwan (1) ambient noise tomography (2) probing source of ambient noise, *Eos Trans. AGU, 90(52), Fall Meet. Suppl., Abstract S53A-1460.*

Dahlen, F. A., and J. Tromp (1998), *Theoretical Global Seismology*, Princeton University Press.

Derode, A., E. Larose, M. Tanter, J. d. Rosny, A. Tourin, M. Campillo, and M. Fink (2003), Recovering the Green's function from field-field correlations in an open scattering medium (L), *The Journal of the Acoustical Society of America*, *113*(6), 2973-2976.

Donn, W. L. (1951), FRONTAL MICROSEISMS GENERATED IN THE WESTERN NORTH ATLANTIC OCEAN, *Journal of Meteorology*, *8*(6), 406-415.

Donn, W. L. (1952), CYCLONIC MICROSEISMS GENERATED IN THE WESTERN NORTH ATLANTIC OCEAN, *Journal of Meteorology*, *9*(1), 61-71.

Donn, W. L. (1966), Microseisms, *Earth-Science Reviews*, *1*(2–3), 213-230.

Froment, B., M. Campillo, and P. Roux (2011), Reconstructing the Green's function through iteration of correlations, *Comptes Rendus Geoscience*, *343*(8–9), 623-632.

Gerstoft, P., and T. Tanimoto (2007), A year of microseisms in southern California, *Geophysical Research Letters*, *34*(20), L20304.

Gerstoft, P., M. C. Fehler, and K. G. Sabra (2006), When Katrina hit California, *Geophysical Research Letters*, *33*(17), L17308.

Gerstoft, P., K. Sabra, P. Roux, W. Kuperman, and M. Fehler (2006), Green's functions extraction and surface-wave tomography from microseisms in southern California, *GEOPHYSICS*, *71*(4), SI23-SI31.

Godin, O. A. (1997), Reciprocity and energy theorems for waves in a compressible inhomogeneous moving fluid, *Wave Motion*, *25*(2), 143-167.

Godin, O. A. (2006), Recovering the Acoustic Green's Function from Ambient Noise Cross Correlation in an Inhomogeneous Moving Medium, *Physical Review Letters*, *97*(5), 054301.

Godin, O. A. (2007), Emergence of the acoustic Green's function from thermal noise, *The Journal of the Acoustical Society of America*, *121*(2), EL96-EL102.

Gu, Y. J., C. Dublanko, A. Lerner-Lam, K. Brzak, and M. Steckler (2007), Probing the sources of ambient seismic noise near the coasts of southern Italy, *Geophysical Research Letters*, *34*(22), L22315.

Harmon, N., M. S. D. L. Cruz, C. A. Rychert, G. Abers, and K. Fischer (2013), Crustal and mantle shear velocity structure of Costa Rica and Nicaragua from ambient noise and teleseismic Rayleigh

wave tomography, *Geophysical Journal International*, *195*(2), 1300-1313.

Hasselmann, K. (1963), A statistical analysis of the generation of microseisms, *Reviews of Geophysics*, *1*(2), 177-210.

Haubrich, R. A., W. H. Munk, and F. E. Snodgrass (1963), Comparative spectra of microseisms and swell, *Bulletin of the Seismological Society of America*, *53*(1), 27-37.

Henstridge, J. D. (1979), A signal processing method for circular arrays, *Geophysics*, *44*(2), 179-184.

Kedar, S., M. Longuet-Higgins, F. Webb, N. Graham, R. Clayton, and C. Jones (2008), The origin of deep ocean microseisms in the North Atlantic Ocean, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, *464*(2091), 777-793.

Kraeva, N., V. Pinsky, and A. Hofstetter (2009), Seasonal variations of cross correlations of seismic noise in Israel, *J Seismol*, *13*(1), 73-87.

Landès, M., F. Hubans, N. M. Shapiro, A. Paul, and M. Campillo (2010), Origin of deep ocean microseisms by using teleseismic body waves, *Journal of Geophysical Research: Solid Earth*, *115*(B5), B05302.

Larmat, C. S., R. A. Guyer, and P. A. Johnson (2010), Time-reversal methods in geophysics, *Physics Today*, *63*(8), 31-35.

Larose, E., P. Roux, and M. Campillo (2007), Reconstruction of Rayleigh--Lamb dispersion spectrum based on noise obtained from an air-jet forcing, *The Journal of the Acoustical Society of America*, *122*(6), 3437-3444.

Larose, E., P. Roux, M. Campillo, and A. Derode (2008), Fluctuations of correlations and Green's function reconstruction: Role of scattering, *Journal of Applied Physics*, *103*(11), 114907.

Lin, C.-R., B.-Y. Kuo, W.-T. Liang, W.-C. Chi, Y.-C. Huang, J. Collins, and C.-Y. Wang (2010), Ambient noise and teleseismic signals recorded by ocean-bottom seismometers offshore eastern Taiwan, *Terrestrial, Atmospheric and Oceanic Sciences*, *21*(5), 743-755.

Lin, F.-C., M. H. Ritzwoller, and R. Snieder (2009), Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array, *Geophysical Journal International*, *177*(3), 1091-1110.

Lin, F.-C., V. C. Tsai, B. Schmandt, Z. Duputel, and Z. Zhan (2013), Extracting seismic core phases with array interferometry, *Geophysical Research Letters*, *40*(6), 1049-1053.

Liu, C.-S., S.-Y. Liu, S. E. Lallemand, N. Lundberg, and D. L. Reed (1998), Digital elevation model offshore Taiwan and its tectonic implications, *Terrestrial, Atmospheric and Oceanic Sciences*, *9*(4), 705-738.

Lobkis, O. I., and R. L. Weaver (2001), On the emergence of the Green's function in the correlations of a diffuse field, *The Journal of the Acoustical Society of America*, *110*(6), 3011-3017.

Longuet-Higgins, M. S. (1950), A Theory of the Origin of Microseisms, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *243*(857), 1-35.

Lukac, M., P. Davis, R. Clayton, and D. Estrin (2009a), Recovering temporal integrity with Data
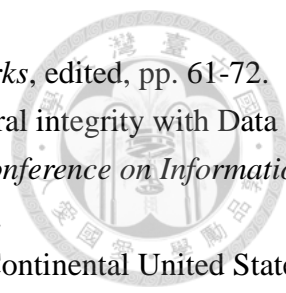
Driven Time Synchronization, in *Information Processing in Sensor Networks*, edited, pp. 61-72.

Lukac, M., P. Davis, R. Clayton, and D. Estrin (2009b), Recovering temporal integrity with Data Driven Time Synchronization, in *Proceedings of the 2009 International Conference on Information Processing in Sensor Networks*, edited, pp. 61-72, IEEE Computer Society.

McNamara, D. E., and R. P. Buland (2004), Ambient Noise Levels in the Continental United States, *Bulletin of the Seismological Society of America*, *94*(4), 1517-1527.

Nakahara, H. (2006), A systematic study of theoretical relations between spatial correlation and Green's function in one-, two- and three-dimensional random scalar wavefields, *Geophysical Journal International*, *167*(3), 1097-1105.

Okeke, E. O., and V. E. Asor (2000), On the microseisms associated with coastal sea waves, *Geophysical Journal International*, *141*(3), 672-678.

Poli, P., M. Campillo, and H. Pedersen (2012), *Body-Wave Imaging of Earths Mantle Discontinuities from Ambient Seismic Noise*, 1063-1065 pp.

Poli, P., M. Campillo, H. Pedersen, and L. W. Group (2012a), Body-Wave Imaging of Earth's Mantle Discontinuities from Ambient Seismic Noise, *Science*, *338*(6110), 1063-1065.

Poli, P., H. A. Pedersen, M. Campillo, and T. P. L. W. Group (2012b), Emergence of body waves from cross-correlation of short period seismic noise, *Geophysical Journal International*, *188*(2), 549-558.

Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009), Anelastic Earth structure from the coherency of the ambient seismic field, *Journal of Geophysical Research: Solid Earth*, *114*(B7), B07303.

Prieto, G. A., M. Denolle, J. F. Lawrence, and G. C. Beroza (2011), On amplitude information carried by the ambient seismic field, *Comptes Rendus Geoscience*, *343*(8–9), 600-614.

Sabra, K. G., P. Roux, and W. A. Kuperman (2005), Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function, *The Journal of the Acoustical Society of America*, *118*(6), 3524-3531.

Sabra, K. G., P. Roux, P. Gerstoft, W. A. Kuperman, and M. C. Fehler (2006), Extracting coherent coda arrivals from cross-correlations of long period seismic waves during the Mount St. Helens 2004 eruption, *Geophysical Research Letters*, *33*(6), L06313.

Seats, K. J., J. F. Lawrence, and G. A. Prieto (2012), Improved ambient noise correlation functions using Welch's method, *Geophysical Journal International*, *188*(2), 513-523.

Shapiro, N. M., and M. Campillo (2004), Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophysical Research Letters*, *31*(7), L07614.

Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller (2005), High-Resolution Surface-Wave Tomography from Ambient Seismic Noise, *Science*, *307*(5715), 1615-1618.

Snieder, R. (2004), Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Physical Review E*, *69*(4), 046610.

Stehly, L., M. Campillo, and N. M. Shapiro (2006), A study of the seismic noise from its long-range correlation properties, *Journal of Geophysical Research: Solid Earth*, *111*(B10), B10306.
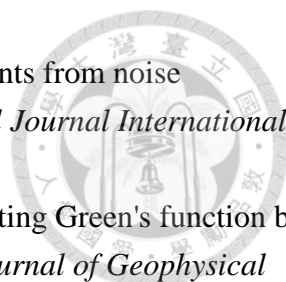
Stehly, L., M. Campillo, and N. M. Shapiro (2007), Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts, *Geophysical Journal International*, *171*(1), 223-230.

Stehly, L., M. Campillo, B. Froment, and R. L. Weaver (2008), Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise, *Journal of Geophysical Research: Solid Earth*, *113*(B11), B11306.

Stephen, R. A., F. N. Spiess, J. A. Collins, J. A. Hildebrand, J. A. Orcutt, K. R. Peal, F. L. Vernon, and F. B. Wooding (2003), Ocean Seismic Network Pilot Experiment, *Geochemistry, Geophysics, Geosystems*, *4*(10), 1092.

Tanimoto, T. (2008), Normal-mode solution for the seismic noise cross-correlation method, *Geophys J*, *175*(3), 1169-1175.

ToksöZ, M. N., and R. T. Lacoss (1968), Microseisms: Mode Structure and Sources, *Science*, *159*(3817), 872-873.

Traer, J., P. Gerstoft, P. D. Bromirski, W. S. Hodgkiss, and L. A. Brooks (2008), Shallow-water seismoacoustic noise generated by tropical storms Ernesto and Florence, *The Journal of the Acoustical Society of America*, *124*(3), EL170-EL176.

Tsai, V. C. (2009), On establishing the accuracy of noise tomography travel-time measurements in a realistic medium, *Geophysical Journal International*, *178*(3), 1555-1564.

Tsai, V. C. (2010), The relationship between noise correlation and the Green's function in the presence of degeneracy and the absence of equipartition, *Geophysical Journal International*, *182*(3), 1509-1514.

Tsai, V. C. (2011), Understanding the amplitudes of noise correlation measurements, *Journal of Geophysical Research: Solid Earth*, *116*(B9), B09311.

Tsai, V. C., and M. P. Moschetti (2010), An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, *Geophysical Journal International*, *182*(1), 454-460.

Weaver, R., and O. Lobkis (2002), On the emergence of the Green's function in the correlations of a diffuse field: pulse-echo using thermal phonons, *Ultrasonics*, *40*(1–8), 435-439.

Weaver, R. L., and O. I. Lobkis (2001), Ultrasonics without a Source: Thermal Fluctuation Correlations at MHz Frequencies, *Physical Review Letters*, *87*(13), 134301.

Weaver, R. L., and O. I. Lobkis (2005), Fluctuations in diffuse field–field correlations and the emergence of the Green's function in open systems, *Journal of the Acoustical Society of America*, *117*(6), 3432-3439.

Welch, P. D. (1967), The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Transactions on audio and electroacoustics*, *15*(2), 70-73.

Wu, F. T., and R.-J. Rau (1998), Seismotectonics and identification of potential seismic source zones in Taiwan, *TAO*, *9*(4), 739-754.
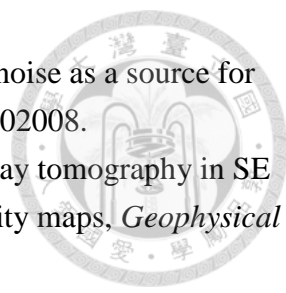
Yang, Y., and M. H. Ritzwoller (2008), Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochemistry, Geophysics, Geosystems*, *9*(2), Q02008.

Yao, H., R. D. Van Der Hilst, and M. V. De Hoop (2006), Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps, *Geophysical Journal International*, *166*(2), 732-744.

You, S.-H., Y. Gung, L.-Y. Chiao, Y.-N. Chen, C.-H. Lin, W.-T. Liang, and Y.-L. Chen (2010), Multiscale Ambient Noise Tomography of Short-Period Rayleigh Waves across Northern Taiwan, *Bulletin of the Seismological Society of America*, *100*(6), 3165-3173.

Yu, T.-C., and S.-H. Hung (2012), Temporal changes of seismic velocity associated with the 2006 Mw 6.1 Taitung earthquake in an arc-continent collision suture zone, *Geophysical Research Letters*, *39*(12), L12307.

Zhan, Z., S. Ni, D. V. Helmberger, and R. W. Clayton (2010), Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise, *Geophysical Journal International*, *182*(1), 408-420.

Zhang, J. (2010), Resolving p wave travel-time anomalies using seismic array observations of oceanic storms, *Earth and Planetary Science Letters*, *292*(3), 419.
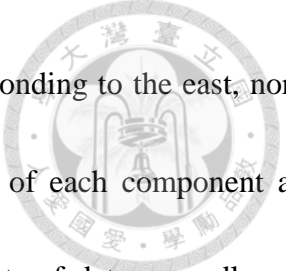
Auxiliary Material

1. Instrument response of short-period stations
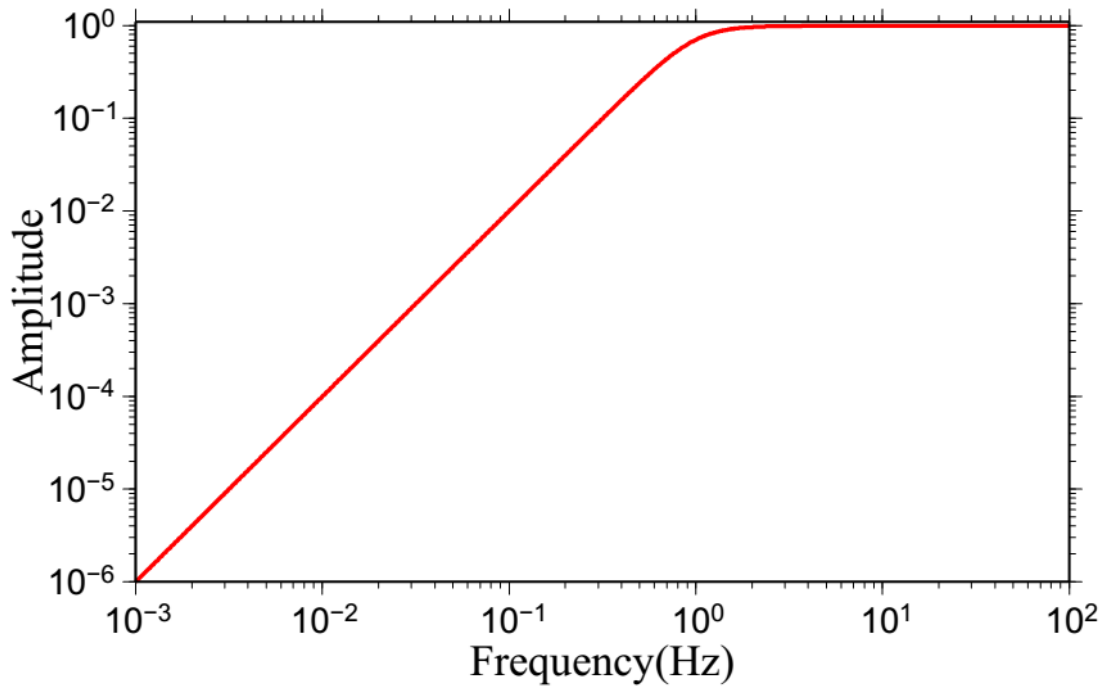
Figure S1: Instrument response of the short-period stations used in this study.

2. Data processing

The procedures of data processing prior to cross-correlation are as follows. First, the continuous raw data are chopped into daily records and decimated to 2 samples per second. Second, the trend, mean and instrument response are removed, and the daily data is then bandpass filtered between 0.1 and 1 Hz. The above steps are similar to the common practice in retrieving *CCFs* using seismic noises. However, we employ a slightly different normalization scheme, in which each component of the vector of filtered time series $S(t)$ is normalized by its

modulus $|S(t)|$, that is, $S(i, t) = \frac{S(i,t)}{|S(i,t)|}$, where i=1, 2, and 3, corresponding to the east, north and vertical component, respectively. As such, the relative strengths of each component are preserved, and the normalization is done once for all three components of data, regardless of which two rotated orthogonal components for each station pair are used in the *CCF* calculation. After the normalization, we compute the daily *CCFs* of the Z-Z, and T-T components up to ±300 seconds, and perform the stacking of the obtained *CCFs* on an annual and monthly basis for further analyses.
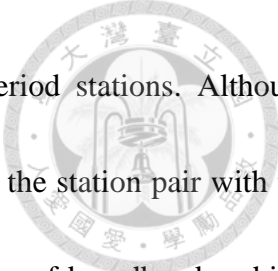
3.   Data selection criterion for the analysis of CCF amplitude asymmetry

We estimate signal-to-noise ratios (*SNR*) for the causal and acausal signals of *CCFs* separately as a measure to assess the effective emergence of the empirical Greens function (*EGF*). *SNR* is defined as the ratio of the peak amplitude in the signal window, bounded by the travel times of wave arriving with constant velocities of 0.5 km/s and 4.0 km/s, to the root-mean-square of the whole trace. The *CCF* traces with *SNR* greater than 5 from either the causal or acausal signals are used in the analysis of amplitude asymmetry.

4. Results from broad-band stations

To ensure our observations are not biased by the band-limited instrument response of short-period stations, we conducted a similar experiment using broad-band stations. Continuous data on the vertical component in the year 2006 from 10 representative coastal broad-band stations are used in this experiment (Figure S2 (a)). The same processing procedure for the short-period data is employed to the broad-band raw data, except that 1-bit normalization scheme is used here. To explore the east-west asymmetric excitations, only the pairs with one station from the west coast and the other from the east coast are analyzed, and the annualstacks of the resulting *CCFs* are shown in Figure S2 (b), in which the signals from the west and east coasts are shown in the acausal and part causal part, respectively. In Figure S2 (a), the relative strengths of the *CCFs* are shown as vectors at each station, with their lengths proportional to the peak amplitudes from the nearby coast (i.e., causal part for the east coast, and acausal for the west coast) and directions to the energy flow along the line of station pairs. It is clear that SPSM excitations from the west coast are stronger than those from the east coast. Besides the east-west asymmetric excitations, we also look into the spectral characteristics of the *CCFs* obtained from broad-band stations. We derive the mean spectrum of all the available monthly stacks, and compare it with the mean spectra from short-period stations. As shown in Figure S2 (c), the *CCFs* in the period range between 2 and 10 seconds derived from broad-band stations do have slightly stronger energy at periods greater than ~6 seconds, their peak energy essentially appears

at the same period around 4 second as those derived from short-period stations. Although observations of longer period microseisms (up to about 20 seconds for the station pair with the longest inter-station distance) could be obtained from cross correlations of broadband ambient noises, currently available broad-band stations along the coast of Taiwan are not dense enough to afford an equivalent investigation as provided by short-period stations.
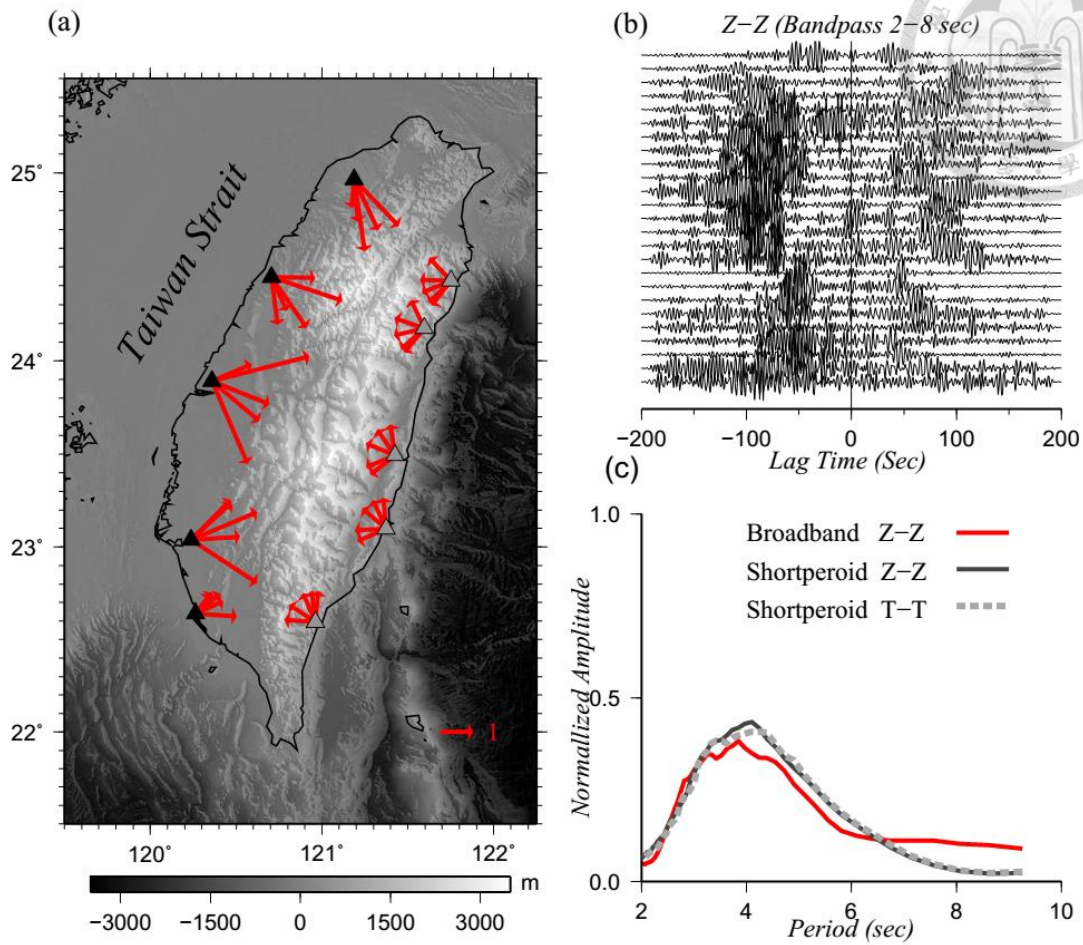
Figure S2: Same as Figure 3-1, except for broad-band stations. (a)The west and east coast stations are denoted by black and gray triangles, respectively. Vectors at each station represent the incoming directions of ambient noises and relative strengths of noise-derived *CCFs* in a unit scale shown in the lower-right corner. (b)*CCF* waveforms obtained from the east-west station pairs shown in Figure 1. The amplitude of each trace is normalized by the peak amplitude of causal signals excited from east coast. (c)Comparison of the mean spectra of *CCFs* derived from broad-band stations and short-period stations.
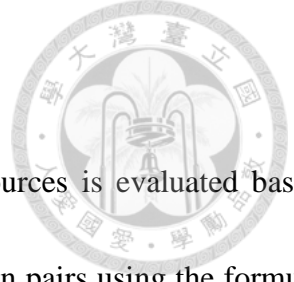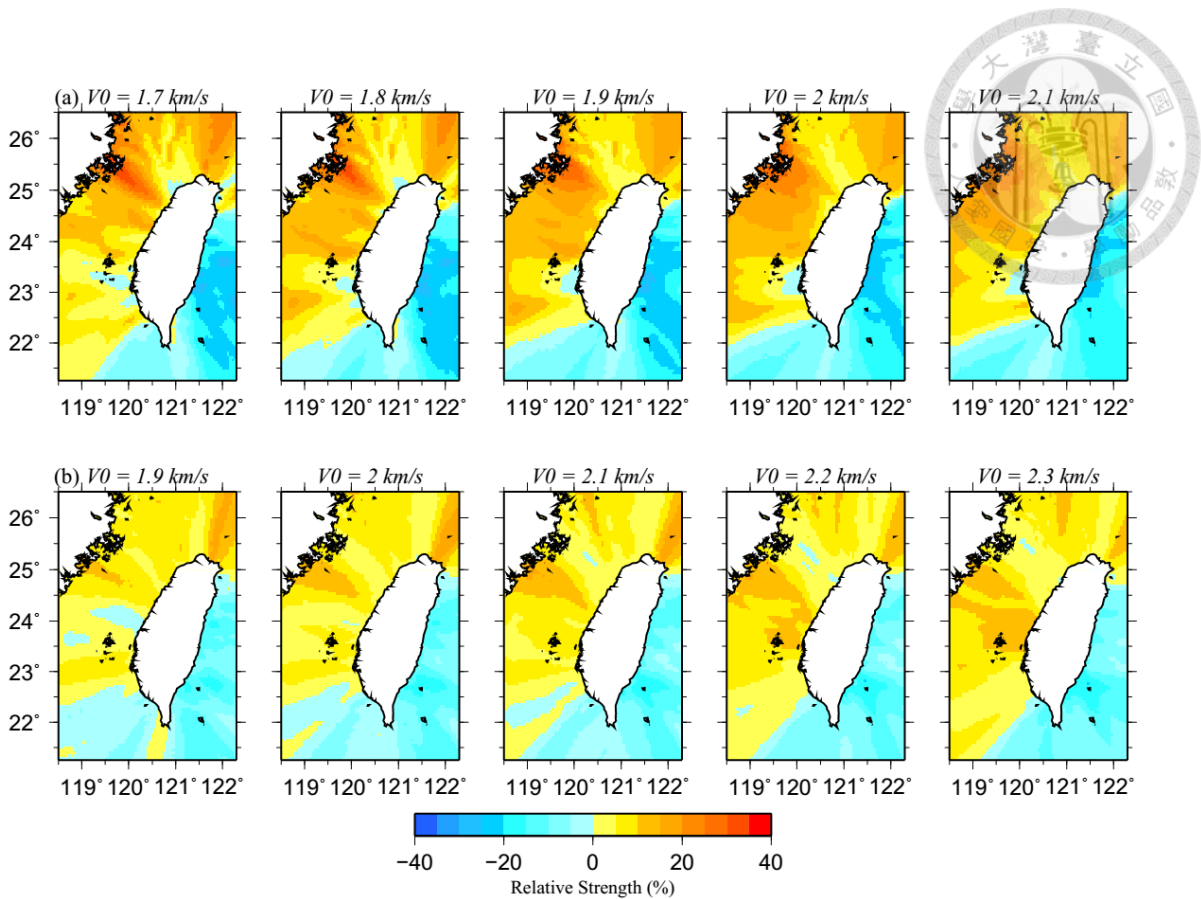
5. Migration imaging method

The excitation on the surveyed area of potential microseism sources is evaluated based upon the strength of cross-correlated signals among all available station pairs using the formula

$A(v, x, y) = \frac{1}{M*N} \sum_j^M \sum_i^N E\left(\frac{d_i(x,y)-d_j(x,y)}{v}\right)$, where A is the excitation intensity of the assumed source at location , M and N are the numbers of coastal and inland stations respectively, E is the envelope of the CCF for stations i and j, di and dj are the distances from the assumed noise source to stations i and j, respectively, and v is the average wave velocity for the area.

As the shallow velocity structure in the oceanic area is unknown, the island-averaged velocities, 1.9 km/s for Rayleigh waves and 2.1 km/s for Love waves at the period 4 seconds, are used for the oceanic area in the migration imaging. While using different oceanic velocities ranging from 1.7 to 2.1 km/s for Rayleigh waves and 1.9 to 2.3 km/s for Love wave s, the general pattern of the resulting images remains very similar (Figure S3).

Figure S3: Same as Figure 3-3, maps s of excitation intensity resulting from migration imaging, except that different migration velocities in oceanic area are tested. The top panels (a) show results of Rayleigh waves, and bottom panels (b) for Love waves.

6. Data sources of wind speeds and wave heights

6.1 wind speeds

The monthly means of 20 island-wide coastal observatories in 2006 are available in an on-line website (http://dbar.as.ntu.edu.tw/) of the Data Bank for Atmospheric Research (DBAR), maintained by the Department of Atmospheric Sciences, National Taiwan University.
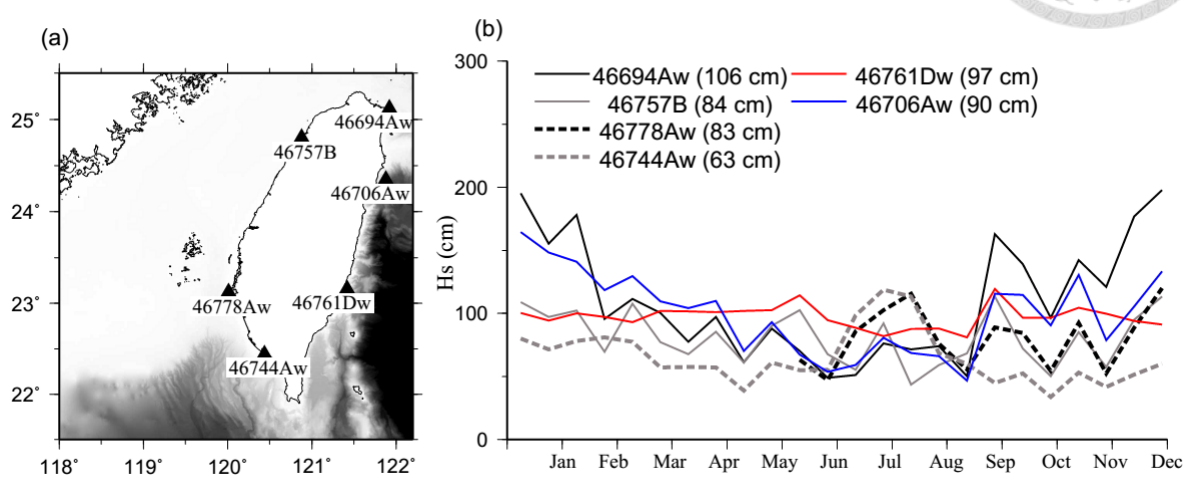
6.2 wave heights

The overall temporal variations of offshore ocean wave heights around Taiwan are obtained
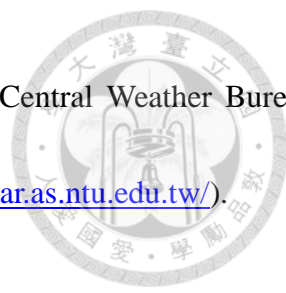
from 6 buoys shown in Figure S4. These buoys are maintained by Central Weather Bureau,

Taiwan, and the data are also available on the DBAR website (http://dbar.as.ntu.edu.tw/).


Figure S4: (a) Locations of 6 offshore buoys used in the estimations of temporal variations of wave heights. (b) The smoothed wave height variations in the year 2006. The mean height for each station is also shown in the parenthesis behind the station name.

7. Comparison of frequency content of CCF signals from the eastern coast and the western coast

The wave frequency in the Taiwan Strait and the eastern open sea is likely quite different, which might lead to the difference in the observed CCF strengths in the SPSM band. This could be examined by looking at the frequency content of CCF signals from the east coast and the west coast, since ocean waves of difference frequencies should have their signatures on the spectral properties of the resulting secondary microseisms. The results are shown in Figure S5, which confirms that the frequency content of the east and west offshore secondary waves are

very close to each other. Therefore, it rules out the possibility that ocean waves of different

frequencies generated in diverse offshore around Taiwan are responsible for the observed CCF

amplitude asymmetry.
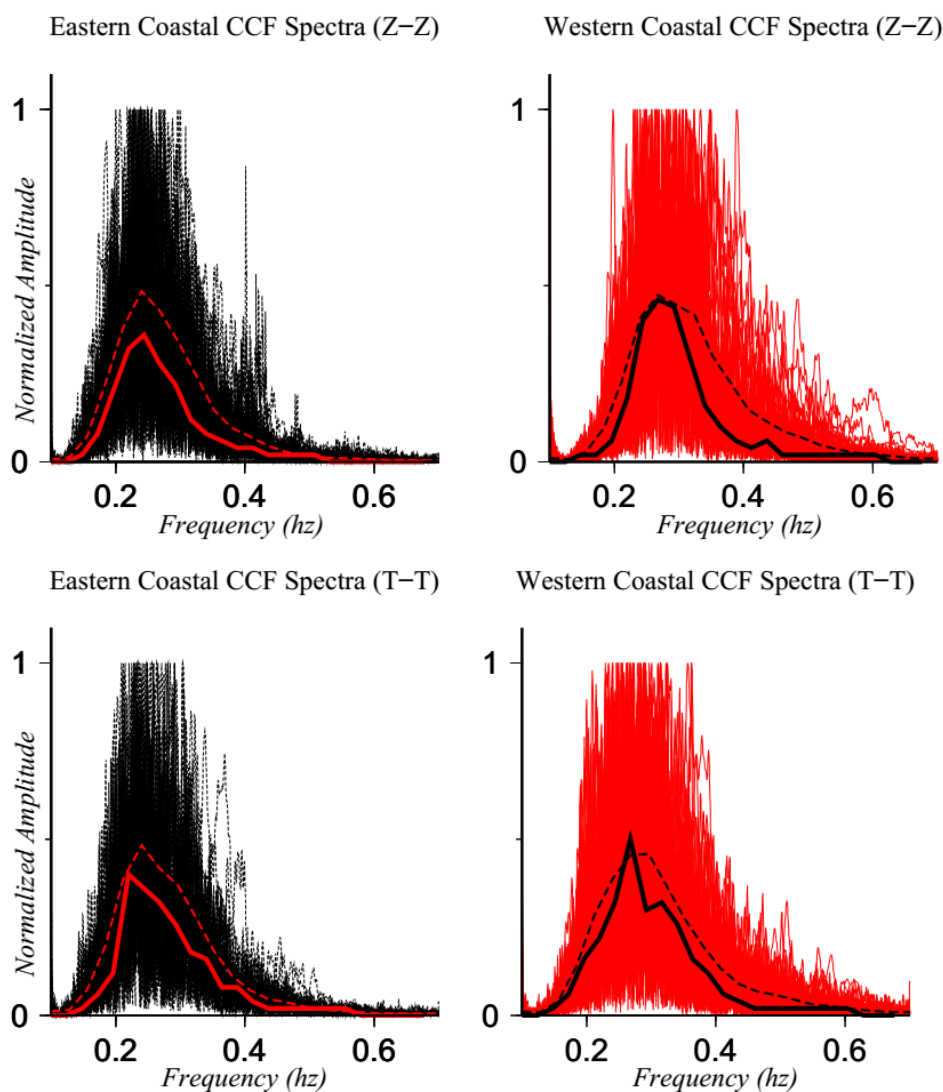
Figure S5: Comparison of normalized spectra of *CCF* signals from the east coast (left) and

the west coast (right) for both the Z-Z (top) and T-T components. The dashed lines and solid

lines are the corresponding mean and mode spectra.