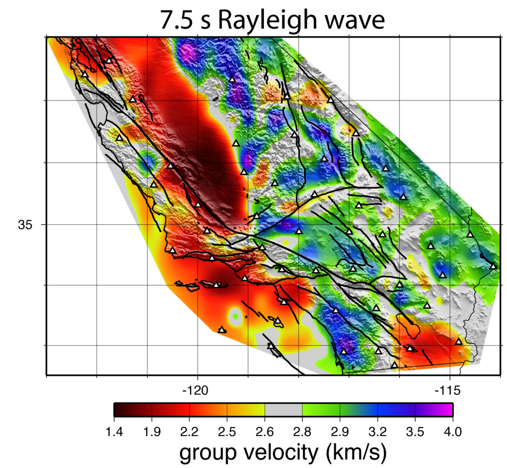
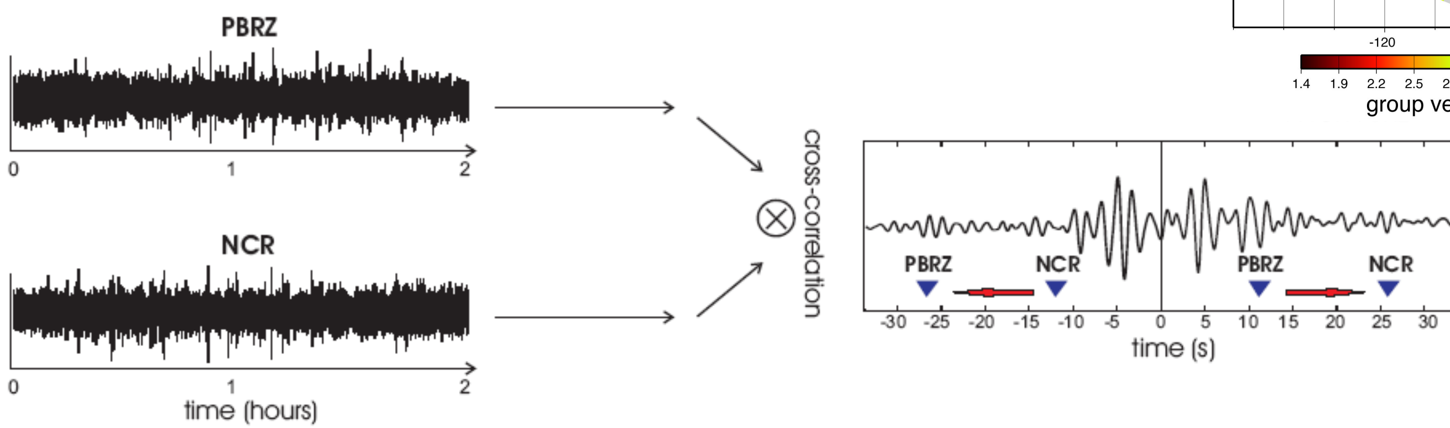
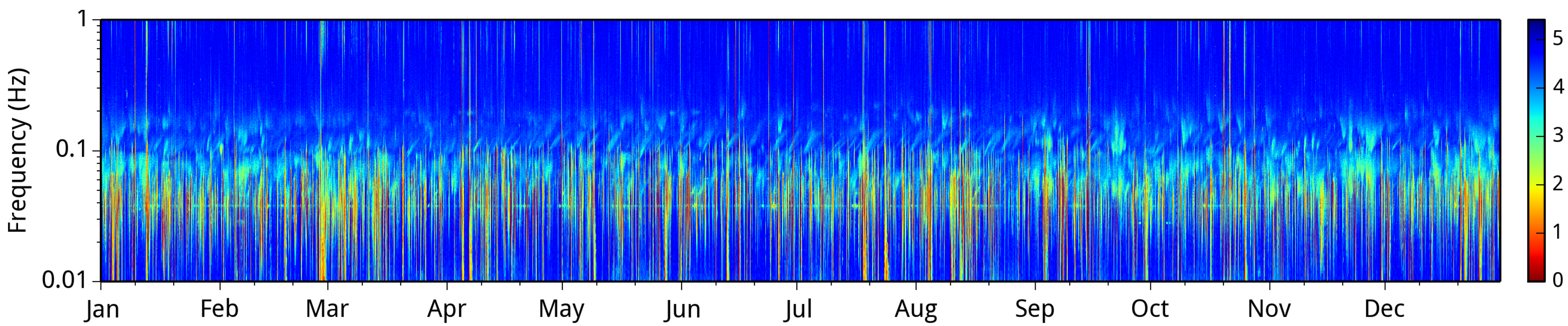
Tomography based on cross-correlations of the ambient seismic noise: accounting for inhomogeneous source distribution

Nikolai Shapiro, Institut de Physique du Globe de Paris

with contributions from: Léonard Seydoux, Julien de Rosny, Aurelien Mordret, Florent Brenguier, Michel Campillo, …





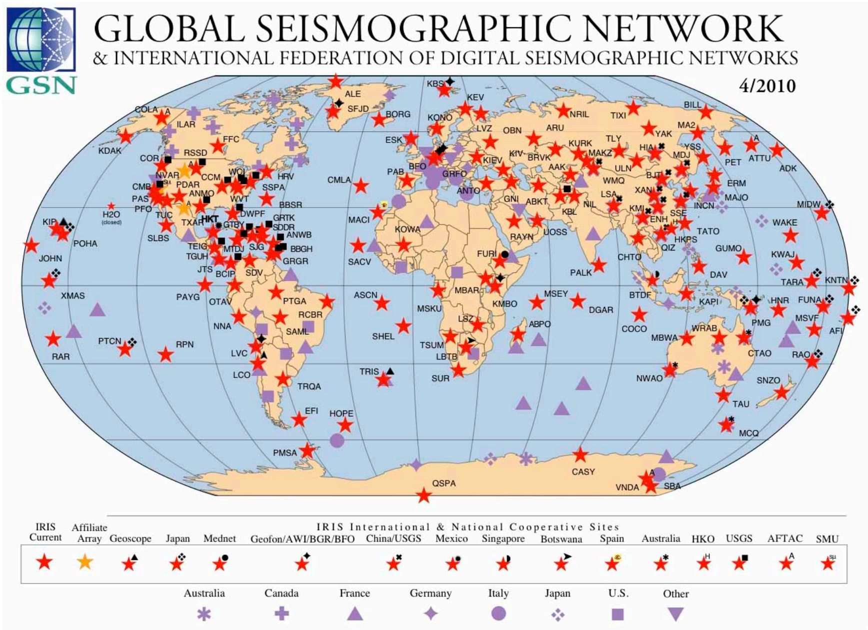


Outline

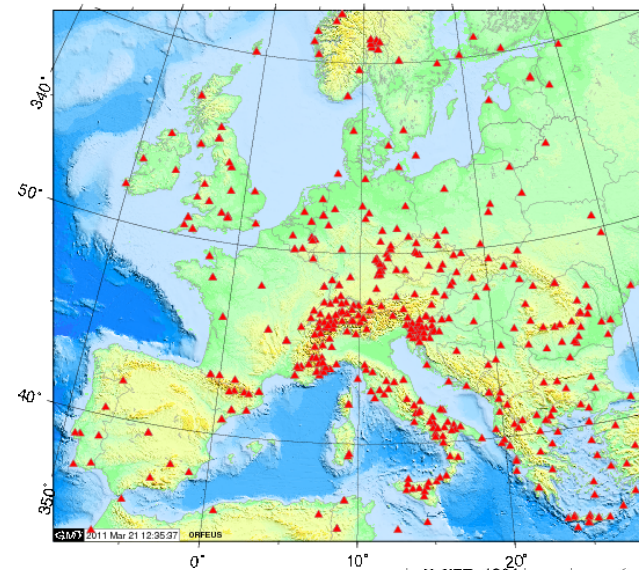
- Brief overview of passive seismic imaging
- “Noise correlation theorem” and the seismic imaging
- Noise-based seismic monitoring

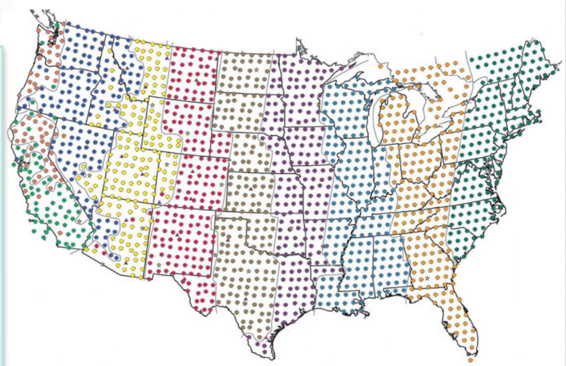
- Signal pre-processing to correct for inhomogeneity of the wavefield
- Using seismic arrays to characterize and to correct the wavefield anisotropy
- A large-scale example: seismic wavefield seen by USArray

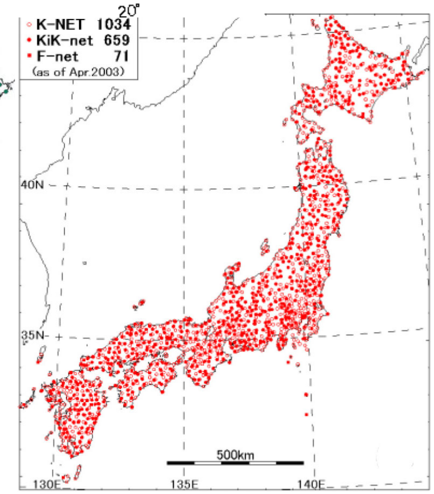
Modern seismological networks



dense regional networks





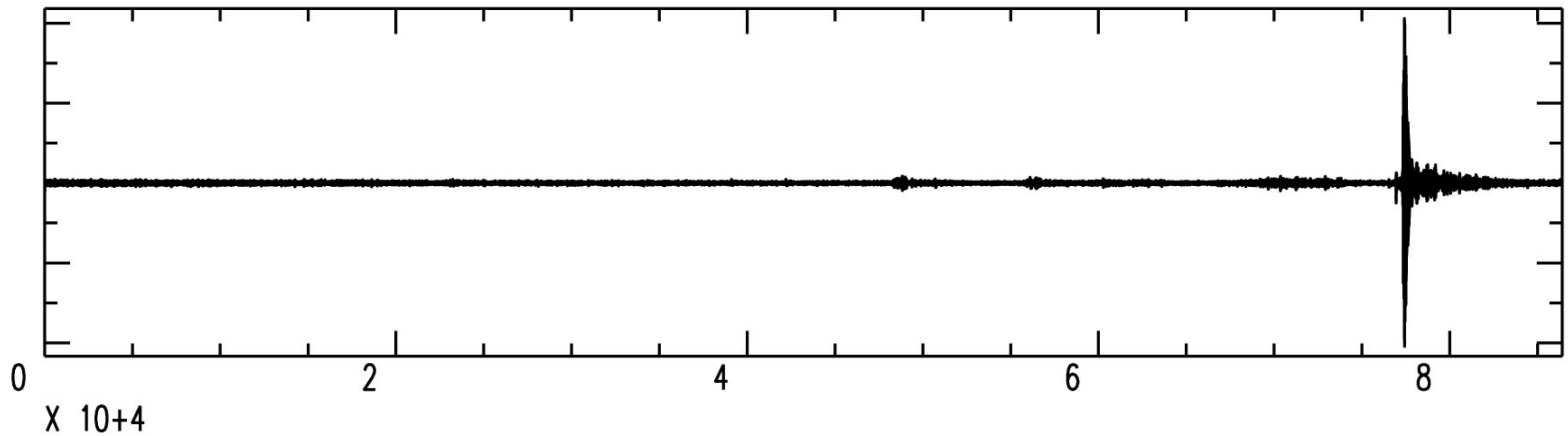


- **Thousands of permanent seismometers are operating continuously**
- **Some temporary networks regroup tens and hundreds of thousands of instruments**
- **Installed on or close to the Earth's surface**
- **Recorded frequencies: 0.001 – 100 Hz**

Seismological observations

records of ground motion (displacement, velocity, or accelerations) by seismographs

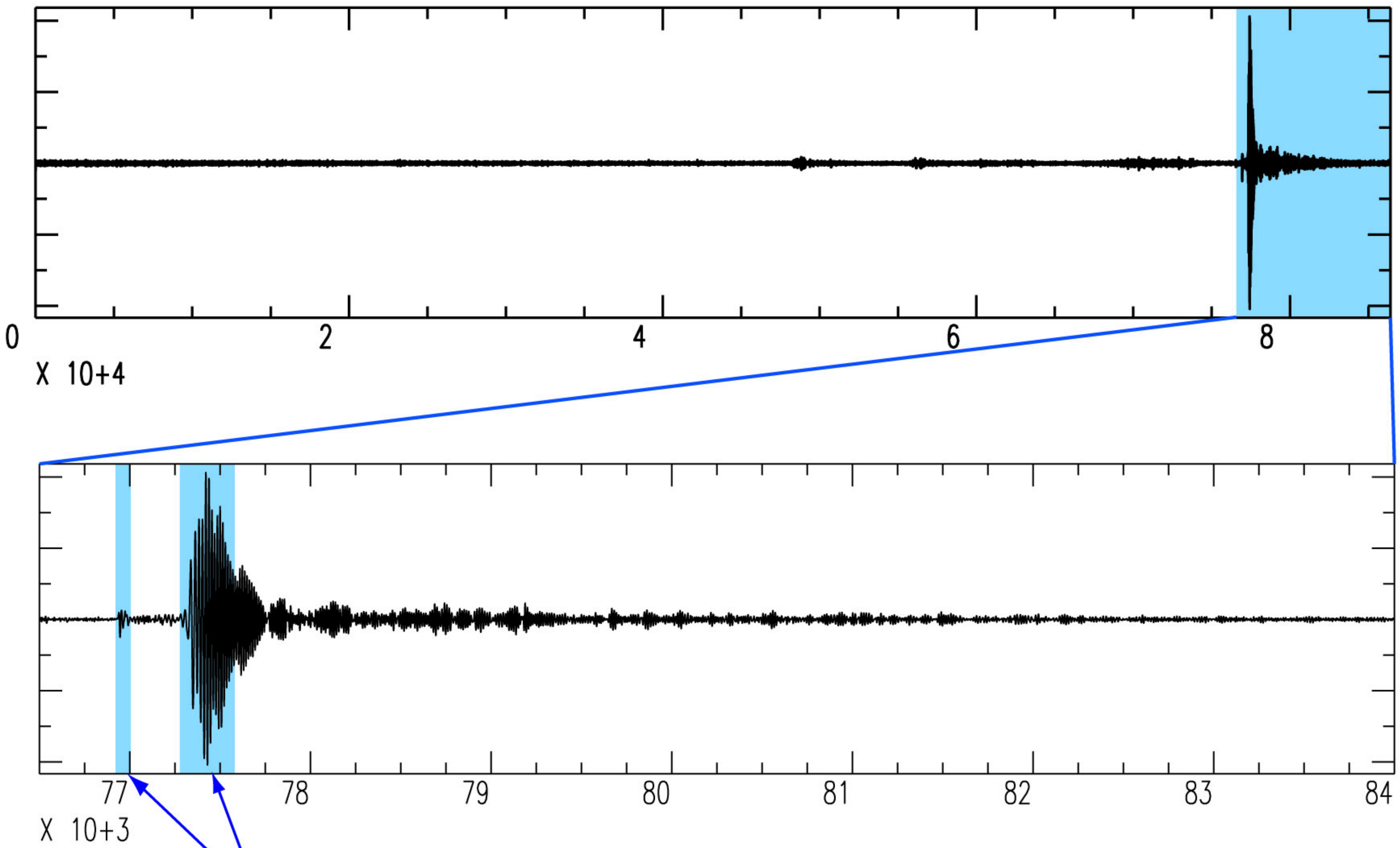
one day of seismic record



Seismological observations

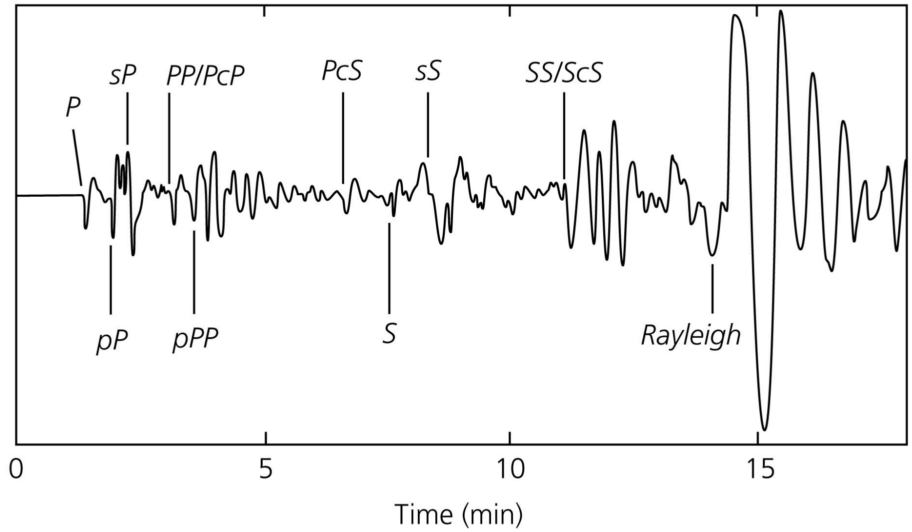
records of ground motion (displacement, velocity, or accelerations) by seismographs

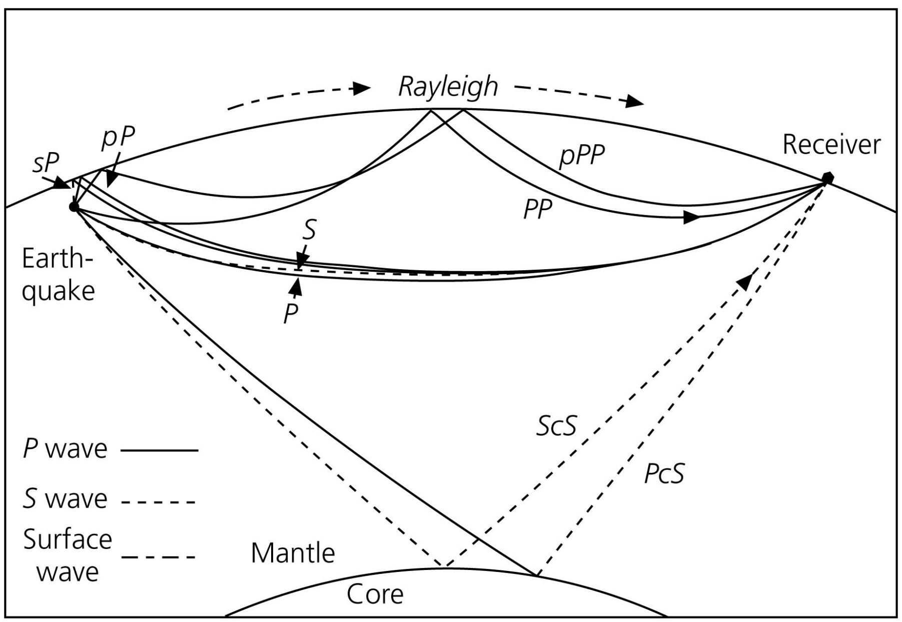
one day of seismic record



ballistic waves used in traditional tomography

Seismic waves emitted by an earthquake





Body waves

sample deep parts of the Earth

P and S

multiplicity of phases because of internal reflections

Surface waves

sample the crust and upper mantle

Rayleigh and Love

Traditional passive seismic imaging uses earthquakes

Strong signals

Sources localized in space and time

Many methods developed since 2-nd half of the 20th century

Inversion of:
- travel times
- amplitudes
- full waveforms

For:
- Vp
- Vs
- Q (attenuation)
- ρ
- anisotropy

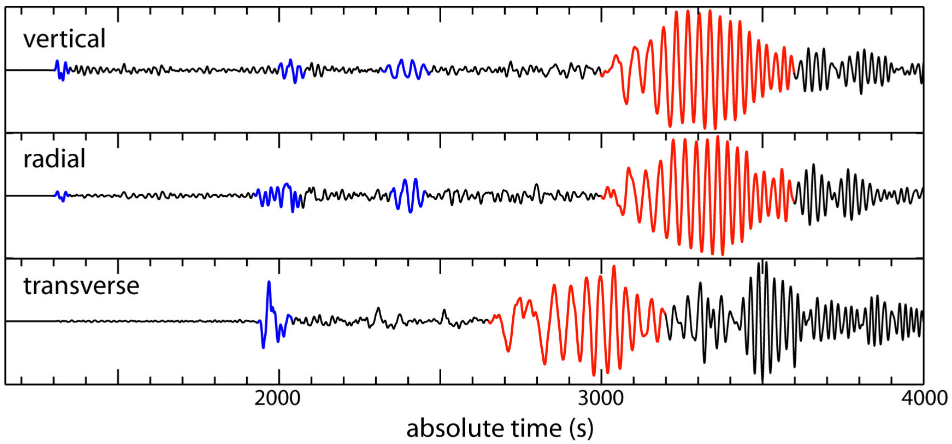
Body-wave tomography

Surface-wave tomography

…

Seismological Inverse problem

earthquake record



$$D = S \otimes M$$

D - seismic data
S - seismic source
M - media (Earth)

imaging, monitoring

we need to know **S** to find **M**

earthquakes

S - location, focal mechanism, time function

Shortcomings of the earthquakes-based methods

- earthquakes do not occur everywhere: limited resolution of resulted images
- earthquakes do not occur continuously: no continuous monitoring possible
- earthquakes rarely occur at the same place: difficult to make repeatable measurements

Preliminary Determination of Epicenters
358,214 Events, 1963 - 1998

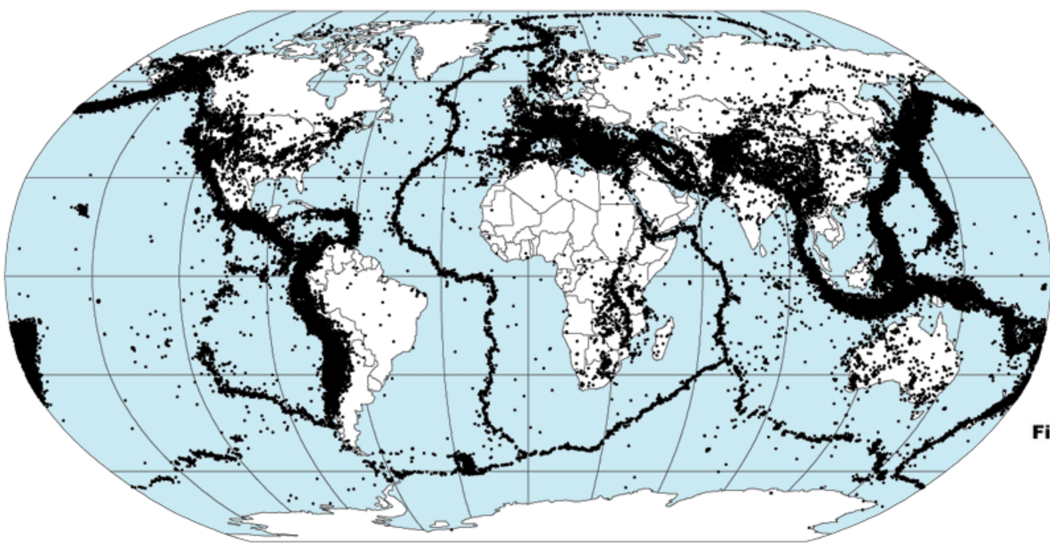
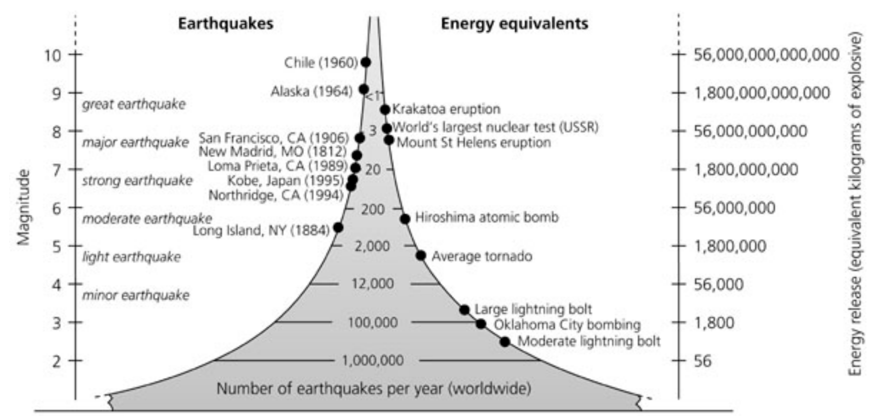
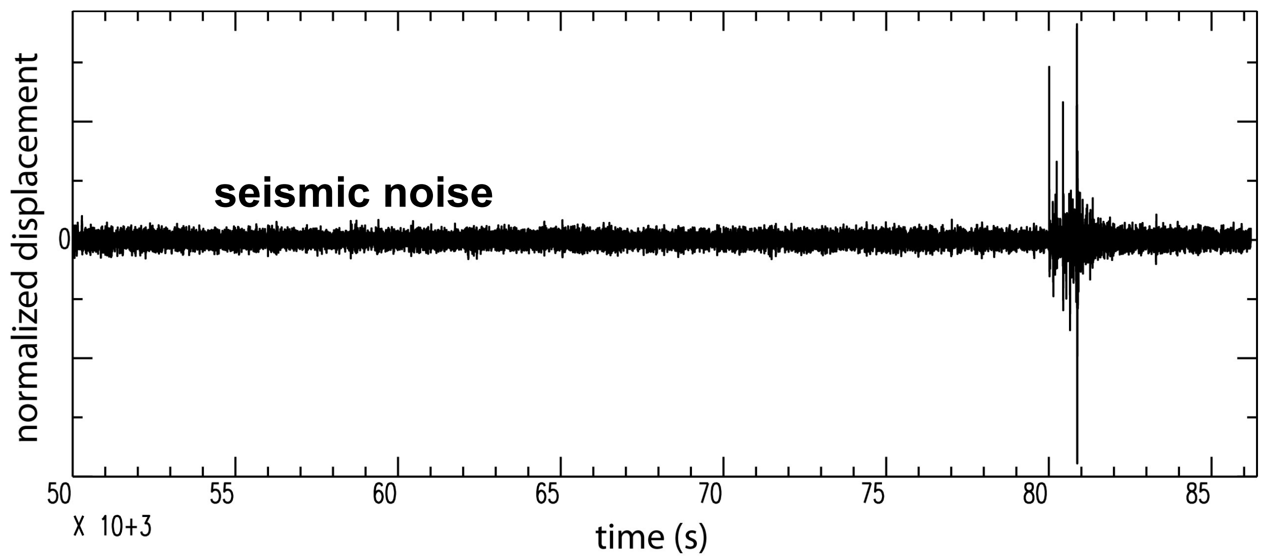


Figure 1.2-2: Comparison of frequency, magnitude, and energy release.



Seismology "without source" : noise based methods

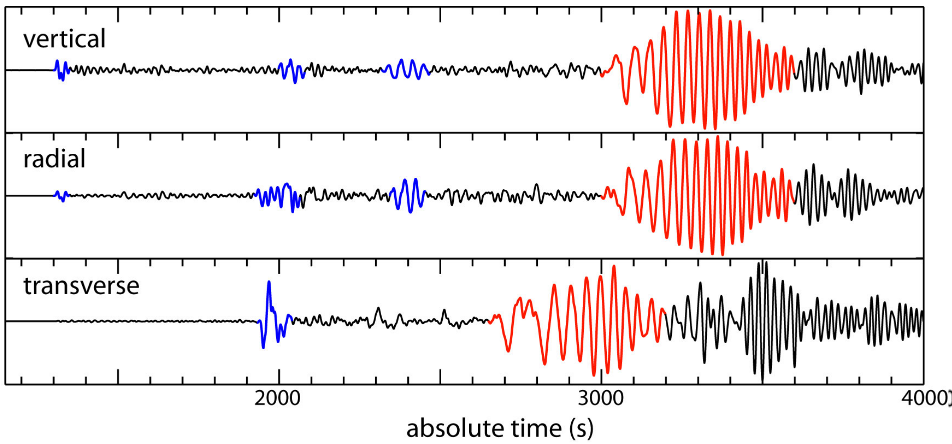


> 95% of seismograms are records of "seismic noise": waves continuously excited by the coupling between the ocean (atmosphere) and the Solid Earth

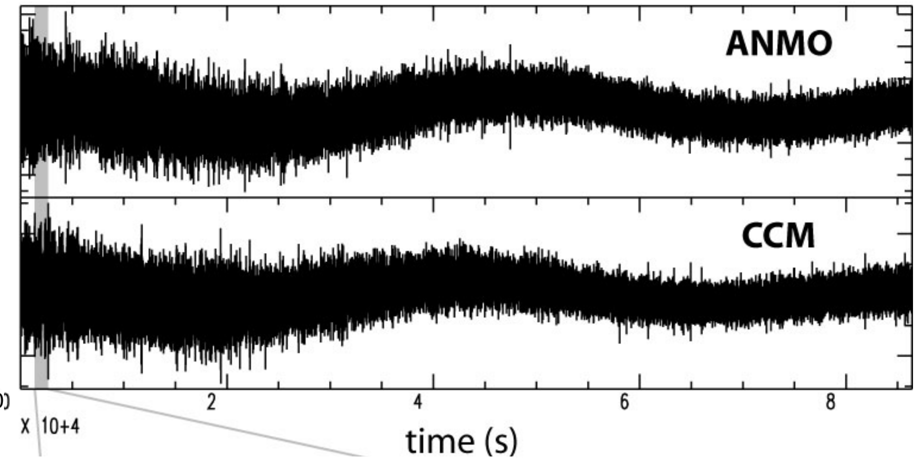
Seismological Inverse problem

Advantage of seismic noise:
Can be recorded anywhere and at any time

earthquake record



noise



$$\mathbf{D} = \mathbf{S} \otimes \mathbf{M}$$

D - seismic data
S - seismic source
M - media (Earth)

imaging, monitoring

we need to know **S** to find **M**

earthquakes

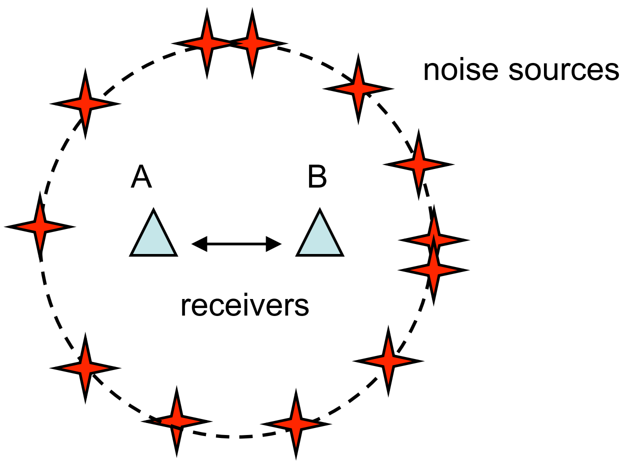
S - location, focal mechanism, time function

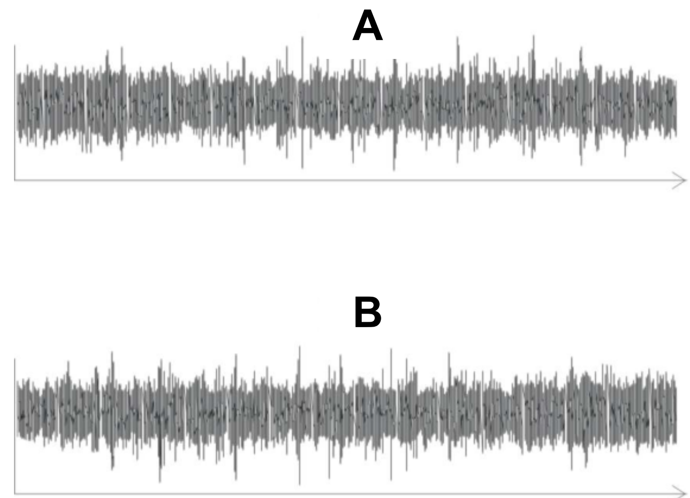
background noise

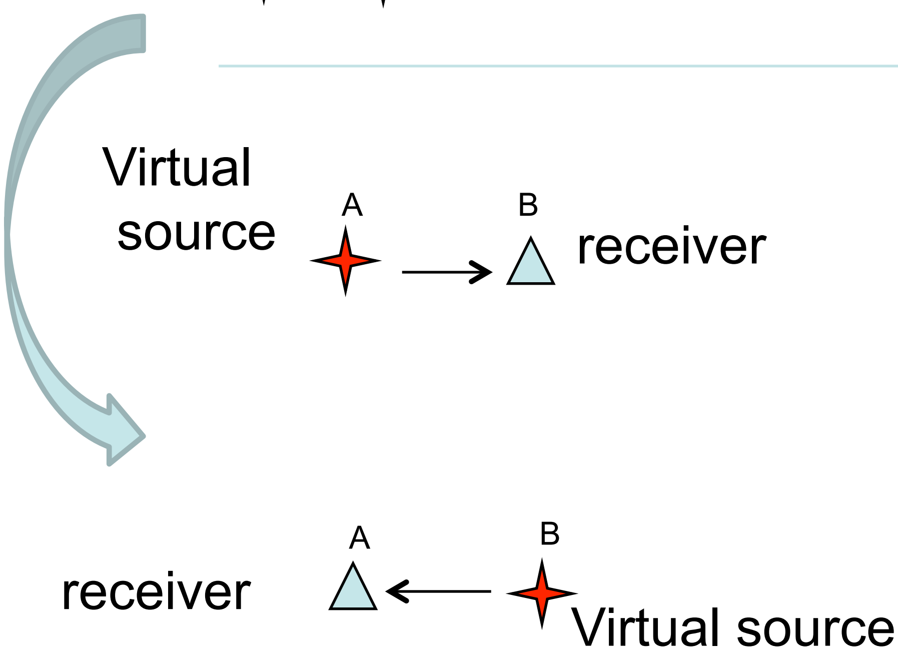
S - complex function

Using seismic noise records for imaging and monitoring

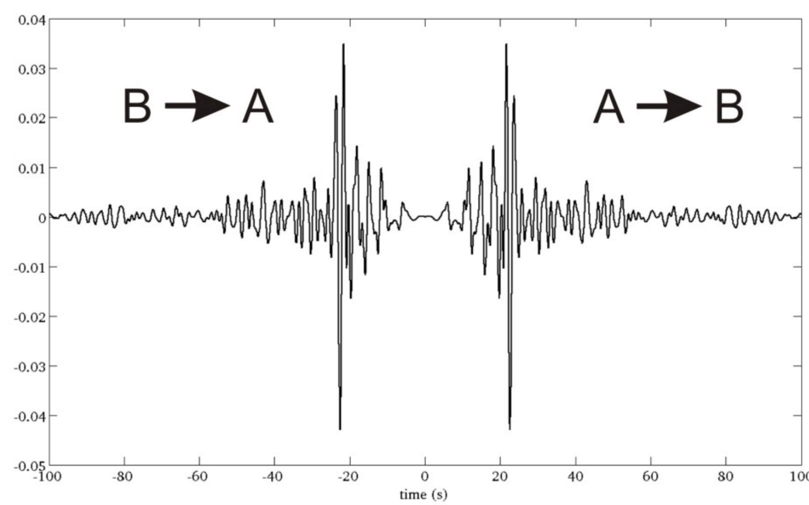
Main idea:
reconstructing impulsive response of the media from noise cross-correlations







Correlation function = Green's function



Noise Correlation Theorem

For a **random** wavefield with sources distributed **homogeneously** everywhere in the medium it can been shown that:

$$\frac{d}{d\tau} C_{A,B}(\tau) = \frac{-\sigma^2}{4\,a}\left(G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)\right)$$

noise cross-correlation **Green function**

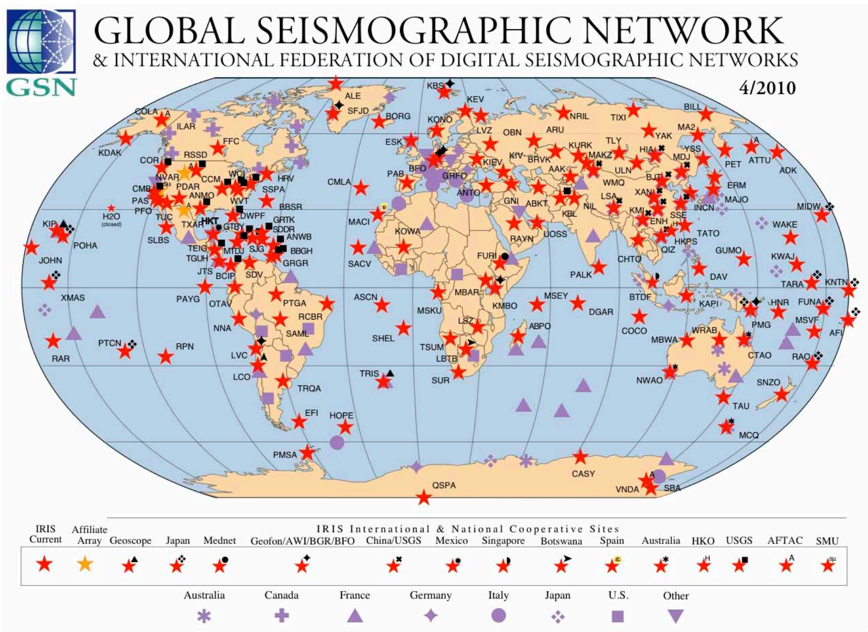
Computing noise cross-correlations between A and B is equivalent to an event occurred at A and recorded at B

$\mathbf{D} = \mathbf{S} \otimes \mathbf{M}$ ⟶ $\mathbf{C(D)} \approx \mathbf{M}$

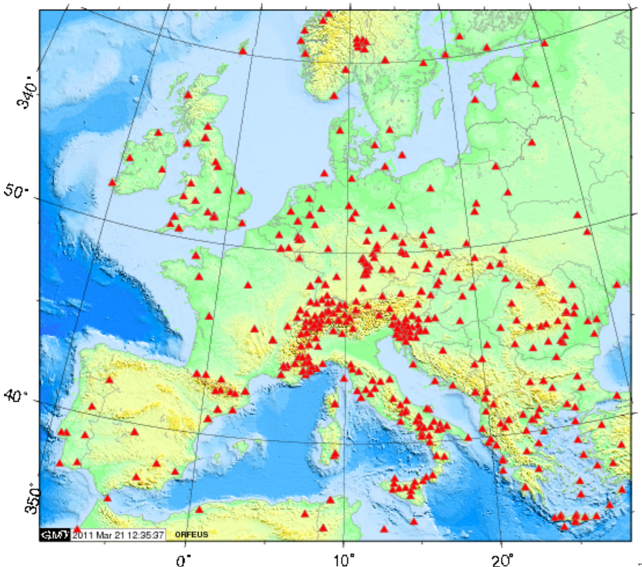
D - seismic data
S - seismic source
M - media (Earth)

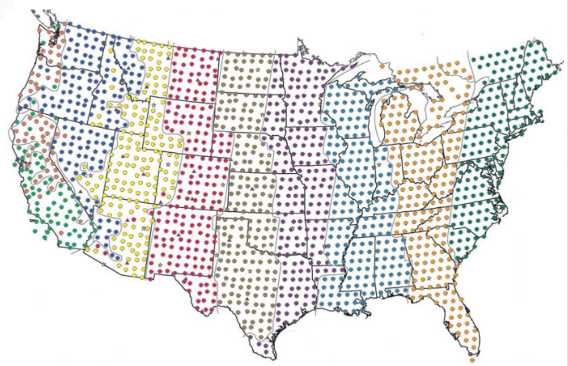
correlation eliminates source complexity

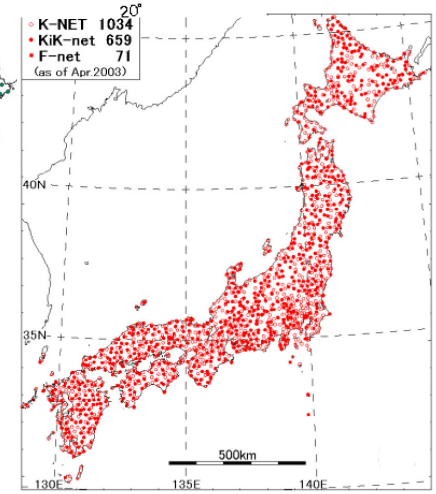
Application of the ‘noise correlation theorem’ to large seismological networks



dense regional networks





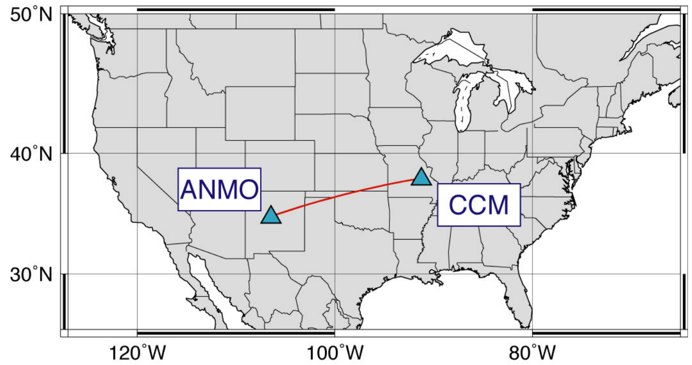


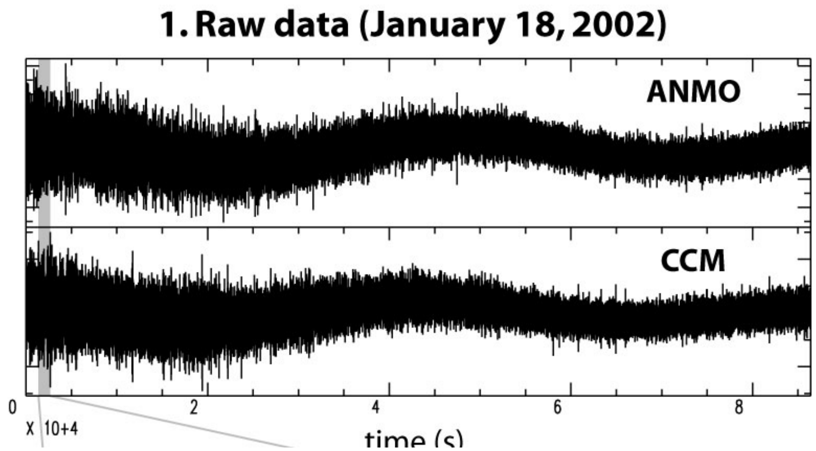
Every receiver acts as a virtual source recorded by all other receivers

N(N-1)/2 virtual seismograms

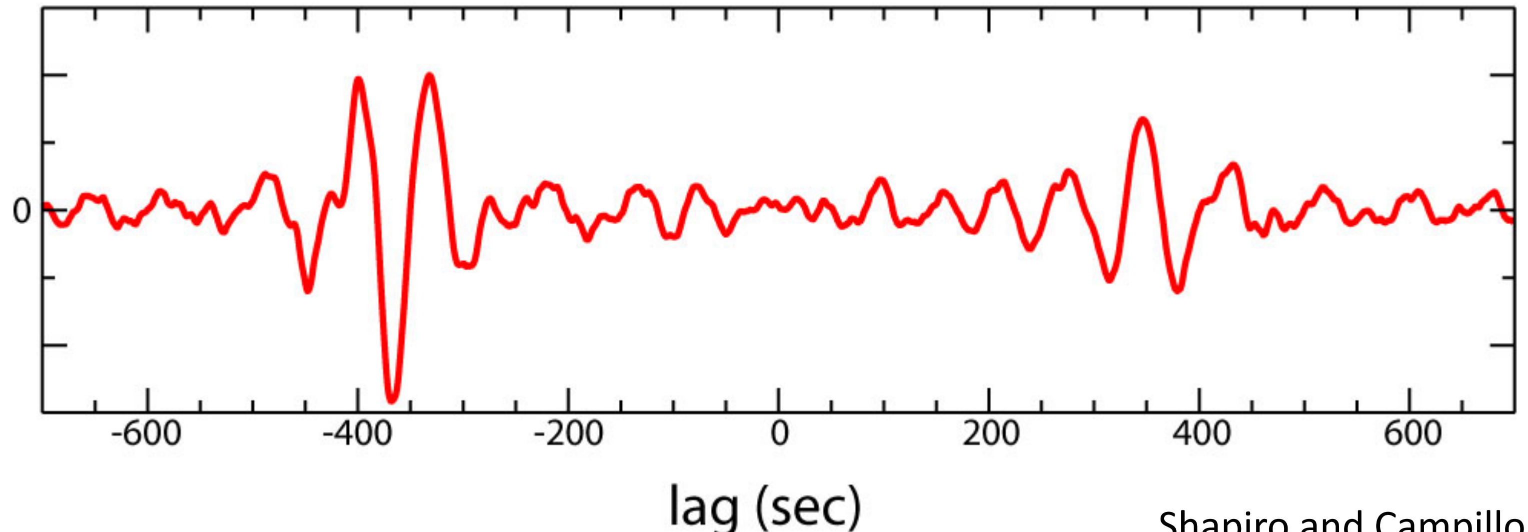
Imaging methods developed for earthquake-generated signals can be applied to virtual seismograms

Extraction of surface waves from noise cross-correlations





cross-correlations from 30 days of continuous vertical component records (2002/01/10-2002/02/08)



Shapiro and Campillo, 2004

Extraction of surface waves from noise cross-correlations

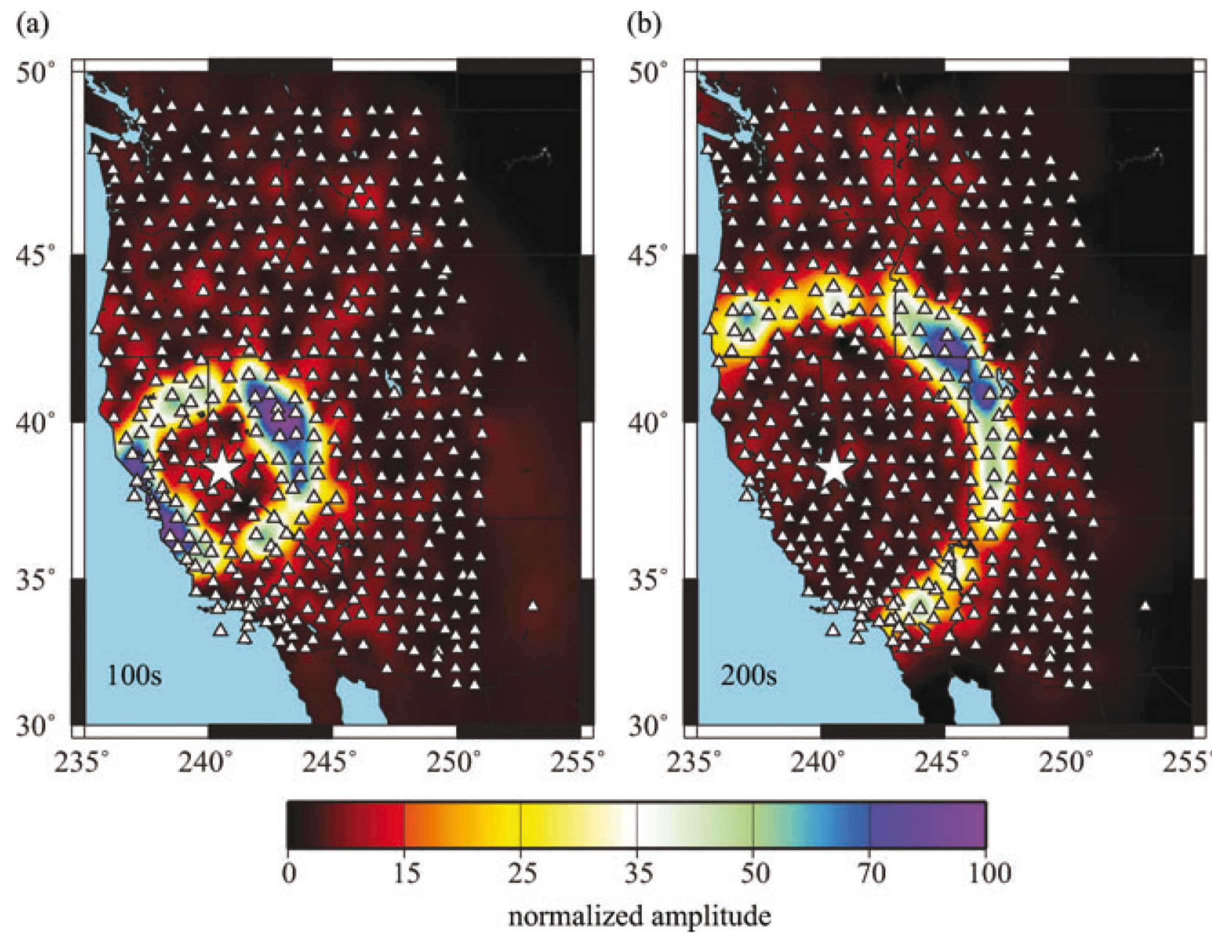
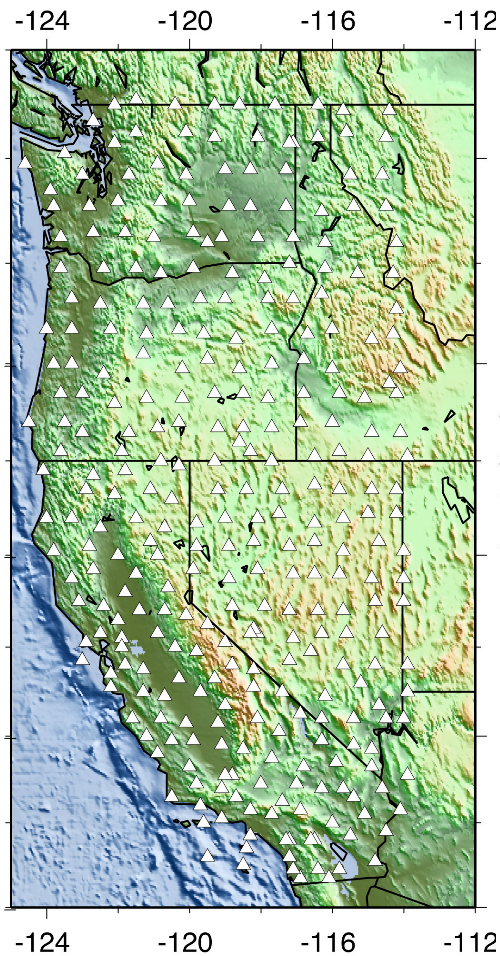
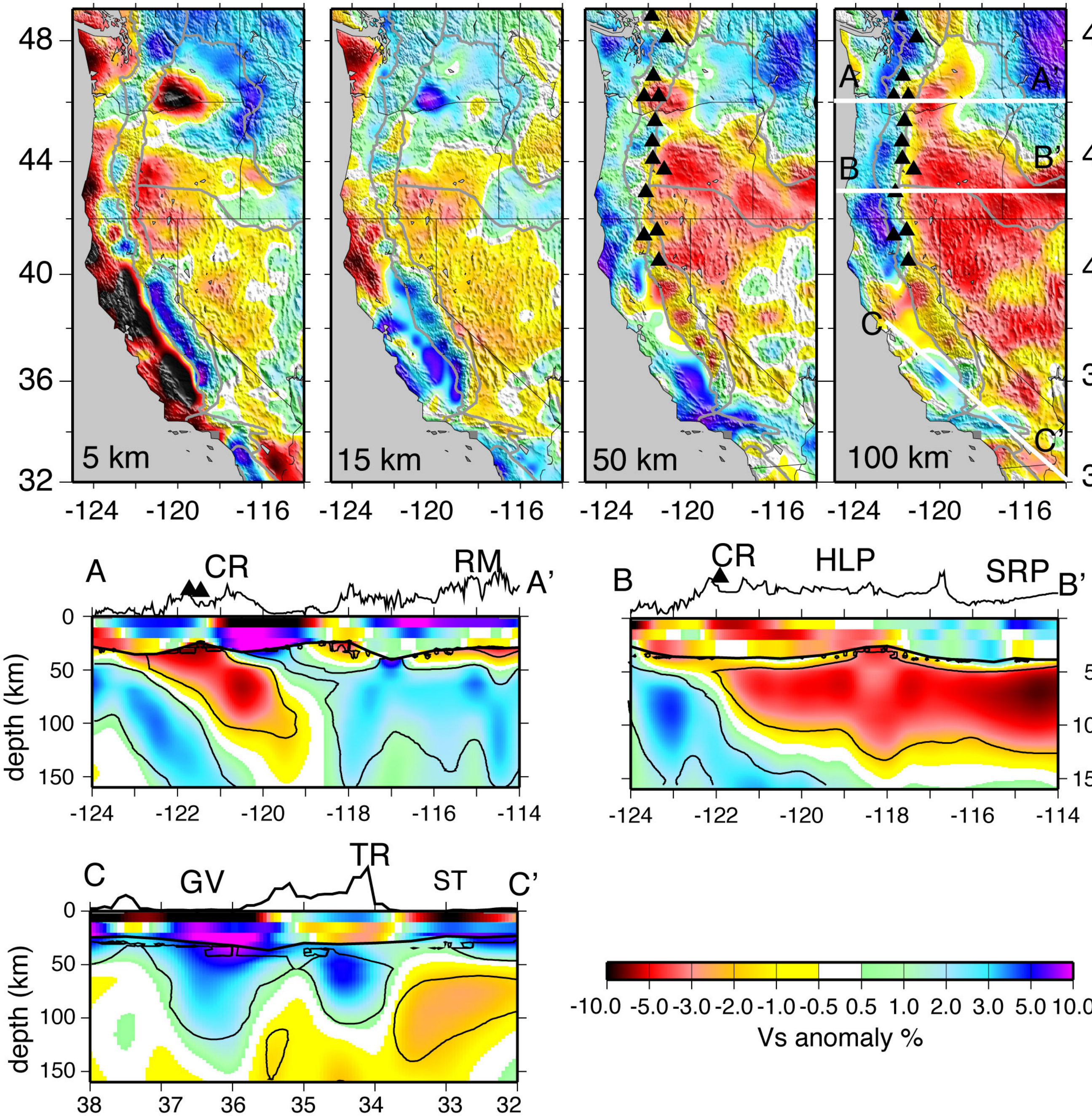


Figure 3 . Snapshots of the normalized amplitude of the ambient noise cross-correlation wavefield with TA station R06C (star) in common at the centre. Each of the 15–30 s band-passed cross-correlations is first normalized by the rms of the trailing noise (Lin et al. 2008) and fit with an envelope function in the time domain. The resulting normalized envelope functions amplitudes are then interpolated spatially. Two instants in time are shown, illustrating move-out and the unequal azimuthal distribution of amplitude.

Lin et al., 2009

Noise based seismic surface-wave tomography with the USArray data

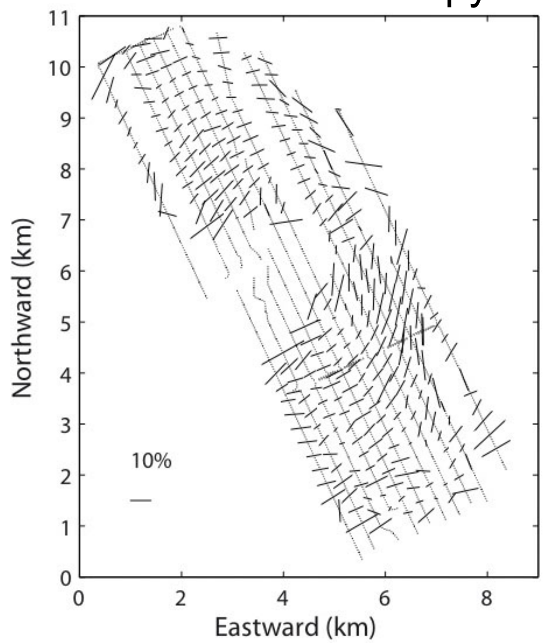


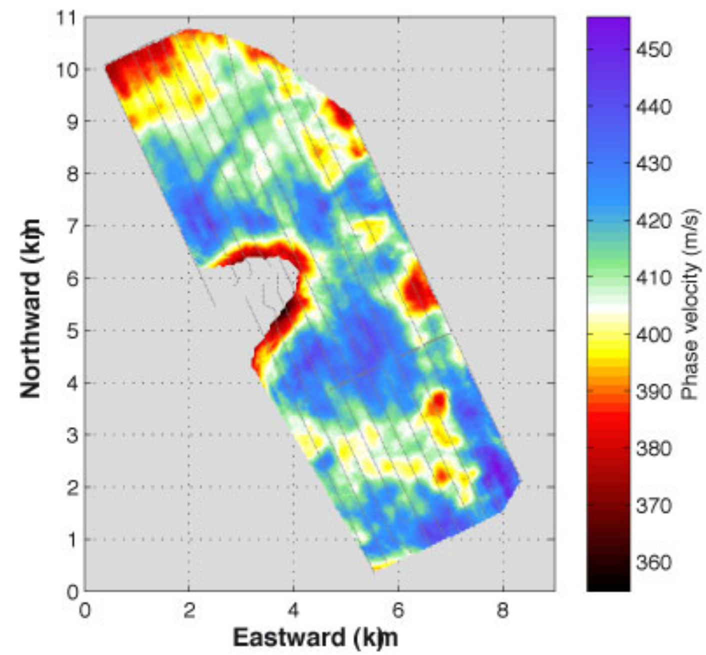


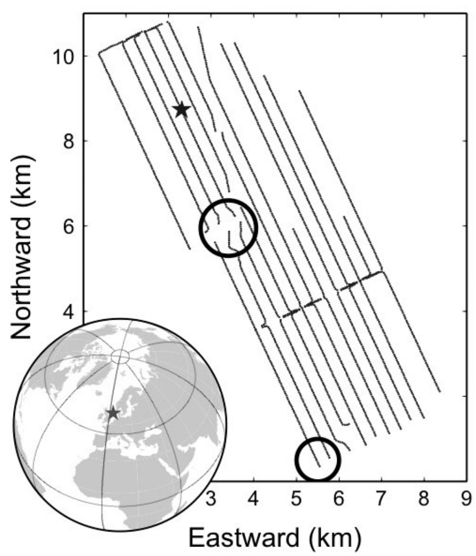
Yang et al. (2008)

Noise-based surface wave tomography of the subsurface above an oil reservoir

Seismic anisotropy







Valhall
Network
2400 receivers

Mordret et al., 2013

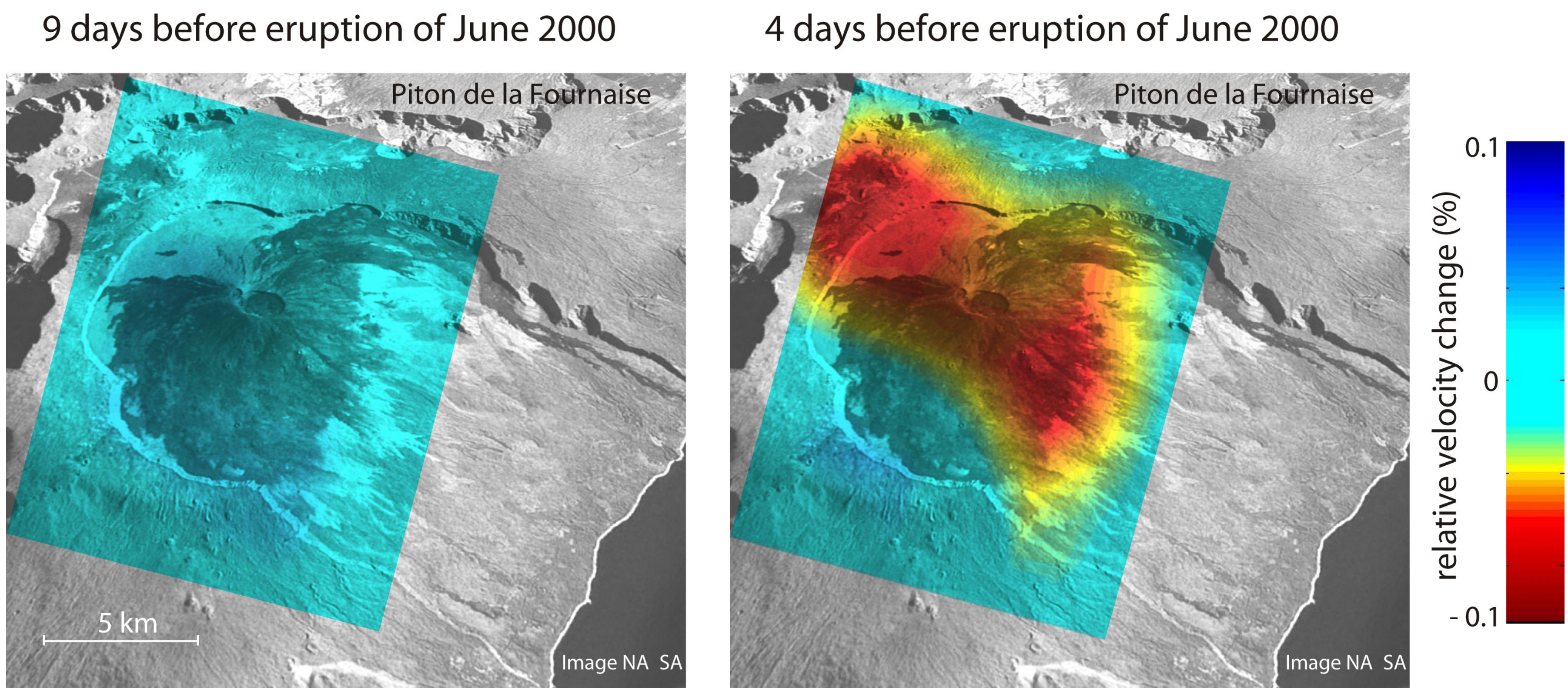
Noise-based monitoring

- **When media changes its Green functions change**

- **Green functions can be reconstructed from noise cross-correlations**

- **Noise cross-correlations can be computed in a nearly-continuous way providing a mean for a monitoring of the Earth's interior**

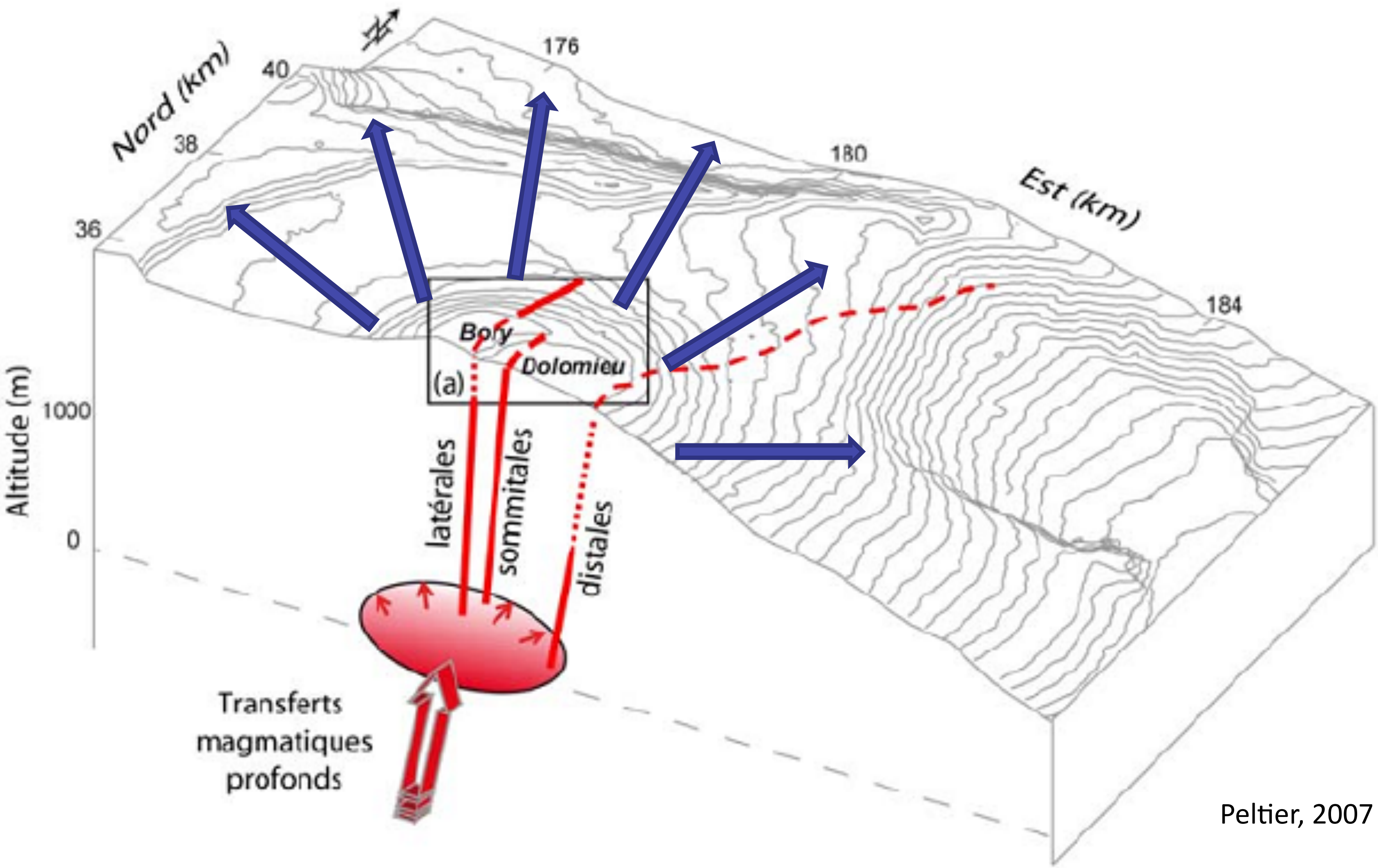
Monitoring Piton de la Fournaise volcano (La Reunion Island)



Detected velocity variations are localized in the vicinity of the main crater: consistent with a shallow source of deformation

Brenguier et al., 2008

Stress build-up within the reservoir "dilates" the edifice and opens cracks



Peltier, 2007

Crustal velocity changes during the 2011 Tohoku earthquake in Japan

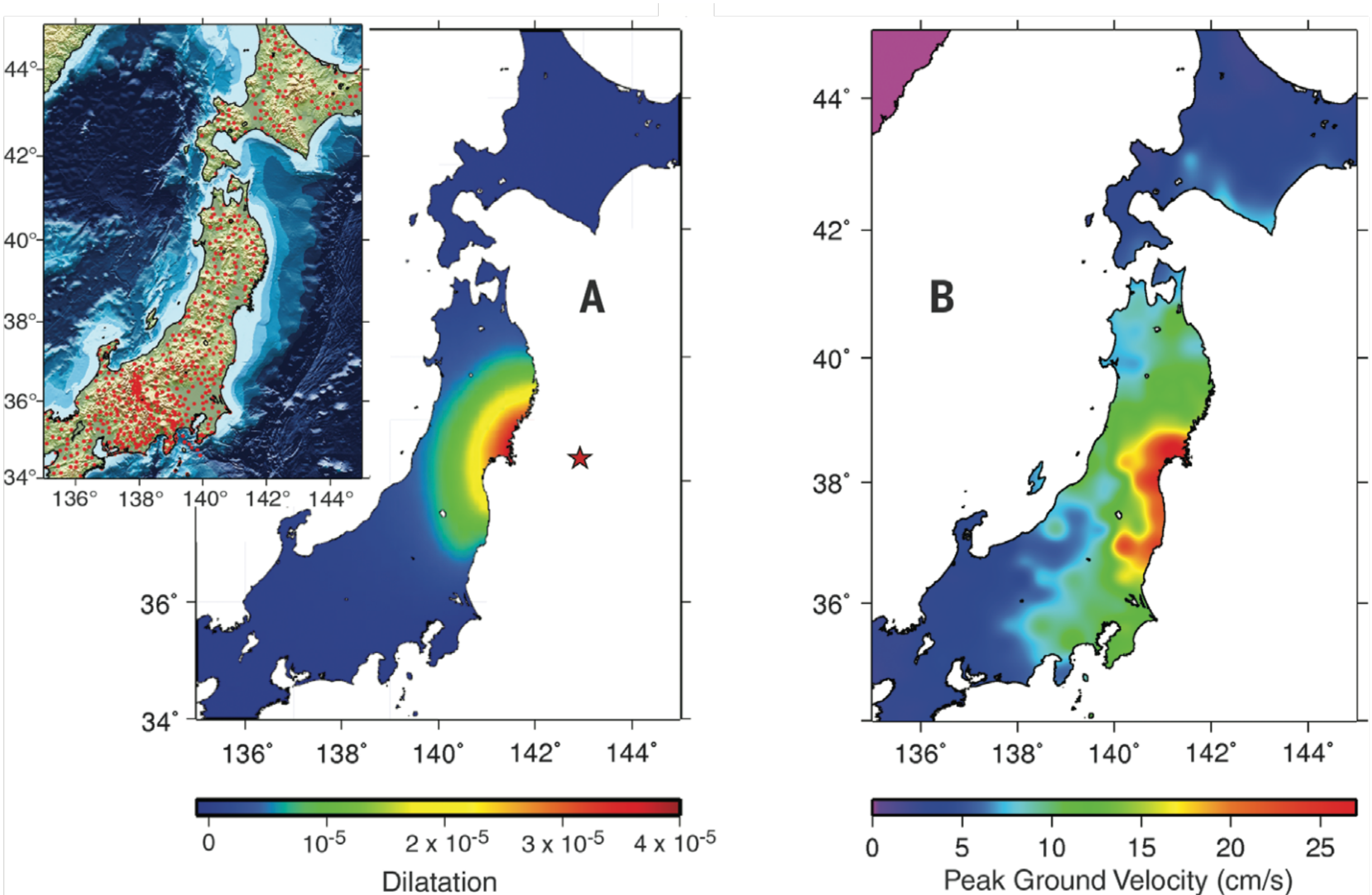
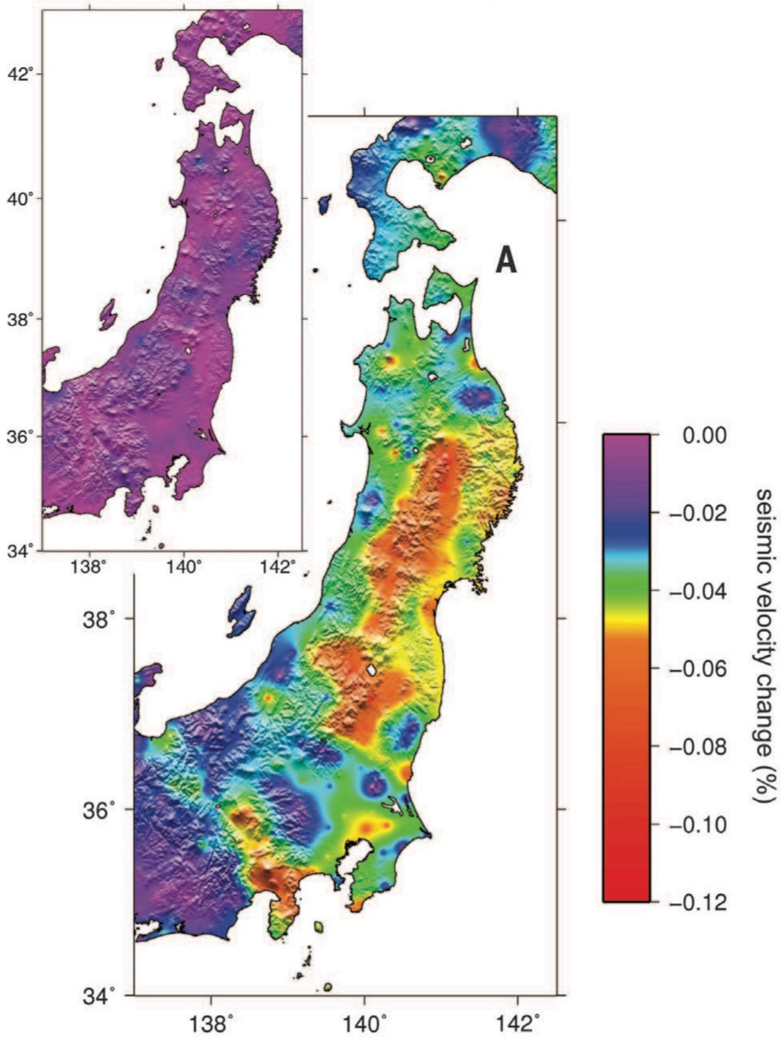


Fig. 1. Static strain and ground shaking caused by the Tohoku-Oki earthquake. (**A**) Modeled coseismic dilatation static strain at 5 km in depth (*7*). The red star shows the position of the epicenter of the Tohoku-Oki earthquake. (Inset) Positions of the Hi-net seismic stations (red points). (**B**) Averaged peak ground velocity measured using the KiK-net strong-motion network (*7*).

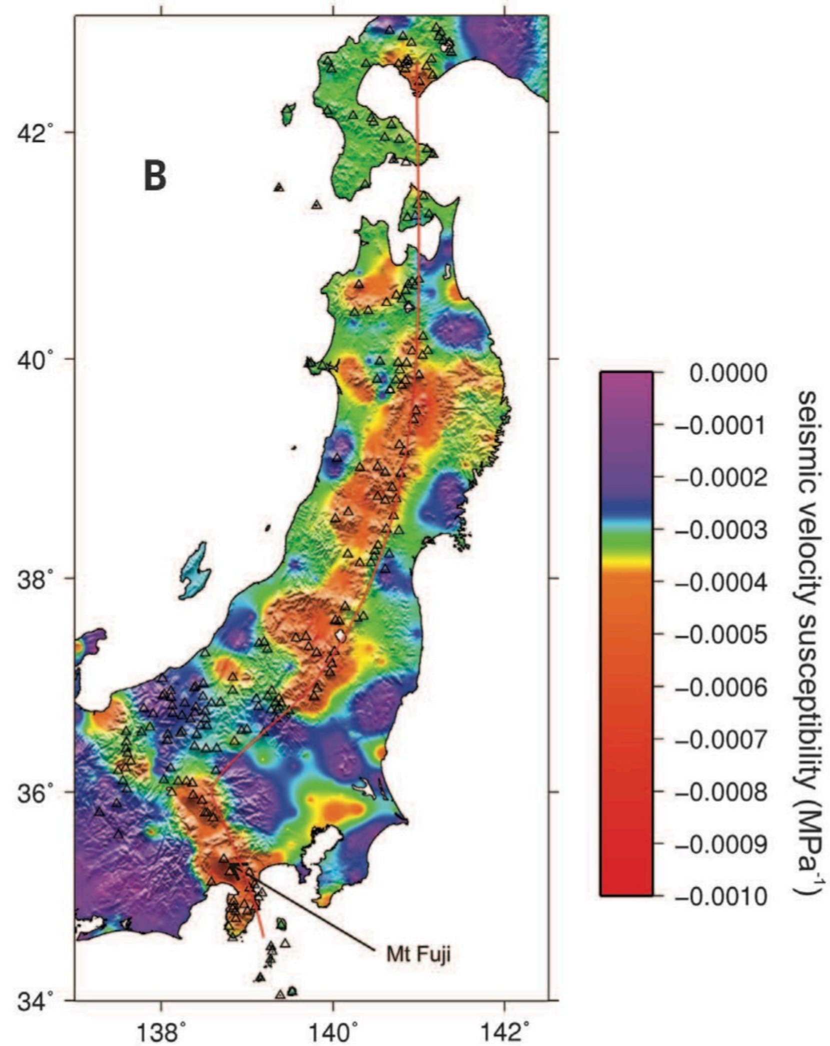
Crustal velocity changes during the 2011 Tohoku earthquake in Japan

5 days before the earthquake



5 days after the earthquake

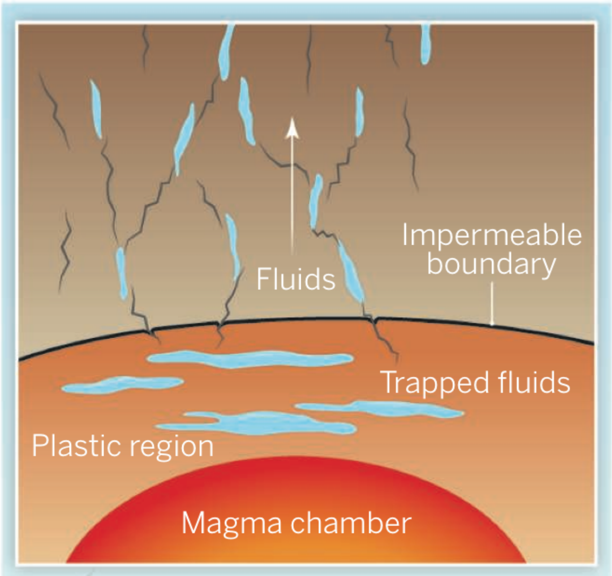
seismic velocity susceptibility: ratio between the observed reductions in the seismic velocity and the estimated dynamic stress.

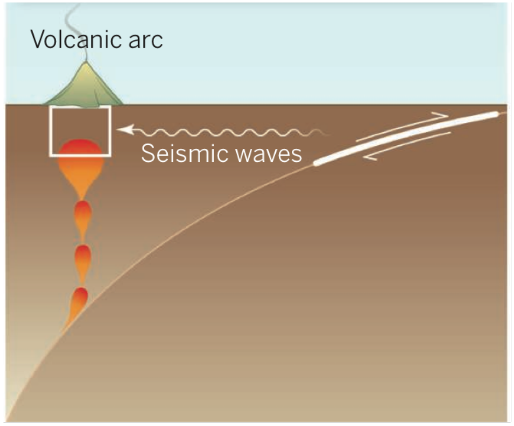


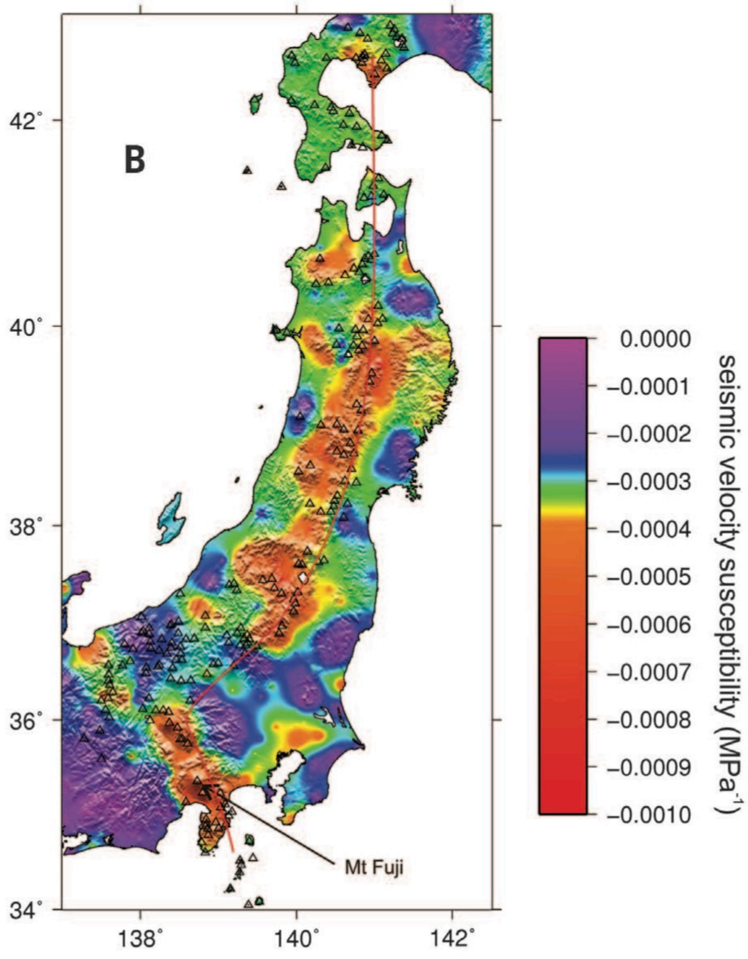
Brenguier et al., 2014

Possible explanations of the observations

Role of widespread hydrothermal fluids







Application of the 'noise correlation theorem' to seismological data

- Synthesis of virtual seismograms: ~N^2 – where N is number of used receivers

- Previously developed imaging methods applied to virtual seismograms

- Proliferation of applications at different scales since 2005

- Noise-based surface wave tomography become a 'standard' and very broadly used method

- Attenuation tomography

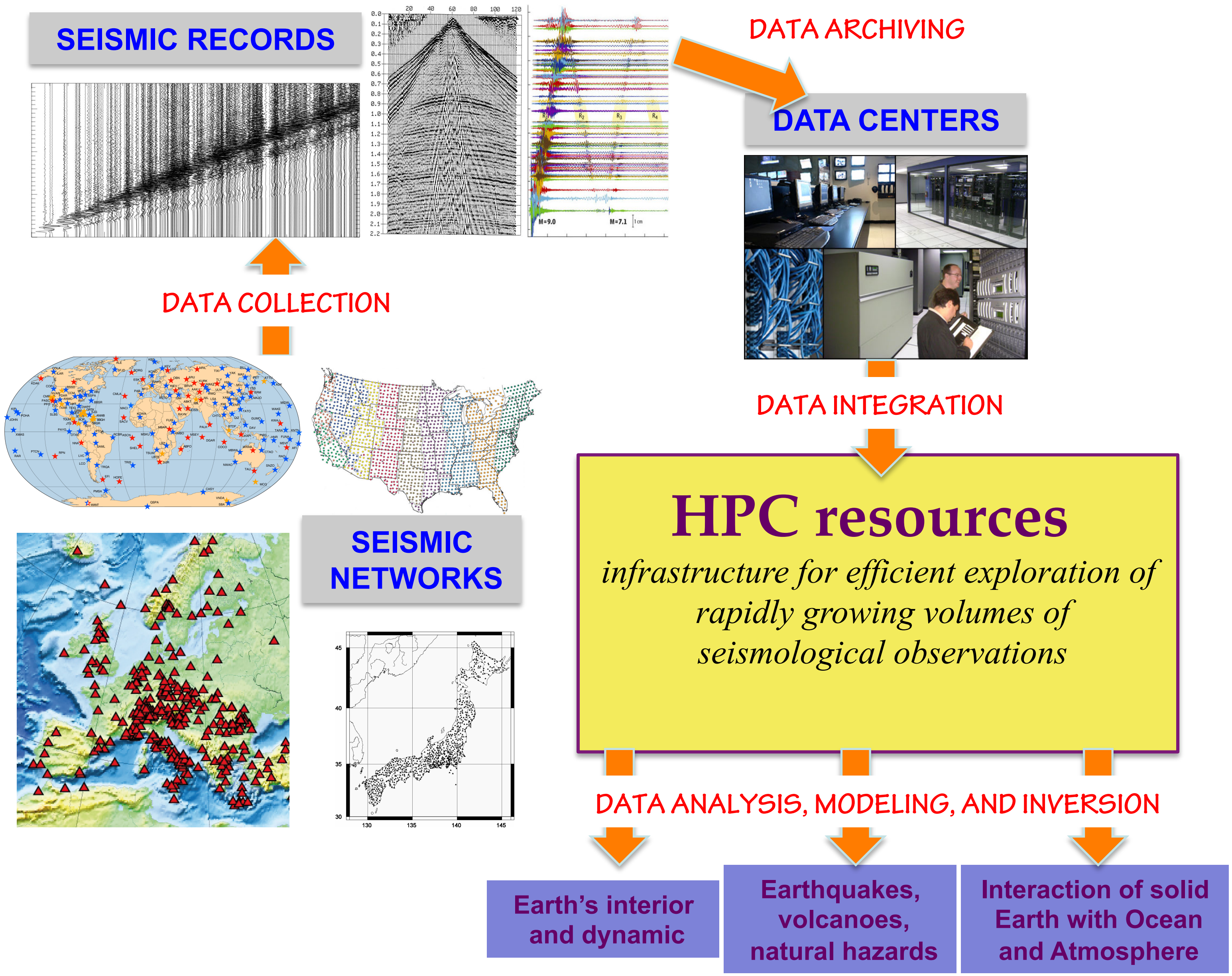
- First demonstrations of the feasibility of the noise-based body wave imaging

- Noise-based monitoring of volcanic and seismogenic areas and of industrial objects

- Empirical prediction of the ground motion from possible future earthquakes for the seismic hazard evaluation

seismic networks : large scale antennas

Data intensive seismology



Outline

- Brief overview of passive seismic imaging
- “Noise correlation theorem” and the seismic imaging
- Noise-based seismic monitoring

- Signal pre-processing to correct for inhomogeneity of the wavefield
- Using seismic arrays to characterize and to correct the wavefield anisotropy
- A large-scale example: seismic wavefield seen by USArray

Noise Correlation Theorem

For a **random** wavefield with sources distributed **homogeneously** everywhere in the medium it can been shown that:

$$\frac{d}{d\tau} C_{A,B}(\tau) \overset{?}{=} \frac{-\sigma^2}{4\,a}\left(G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)\right)$$

noise cross-correlation **Green function**

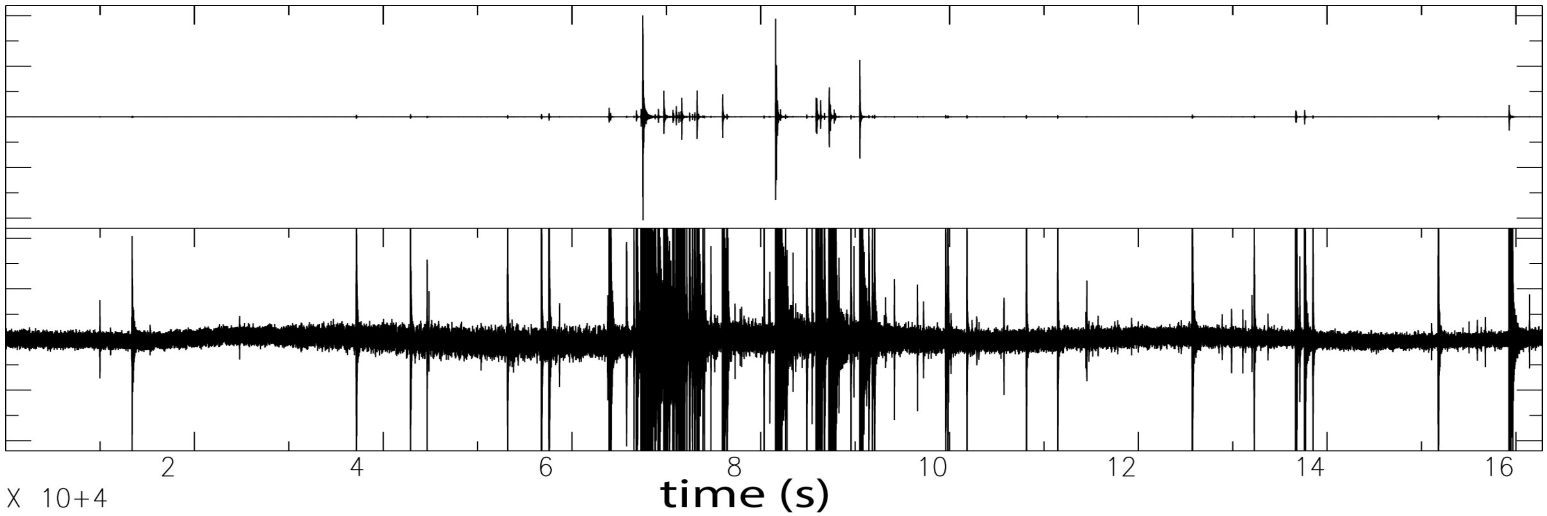
Computing noise cross-correlations between A and B is equivalent to an event occurred at A and recorded at B

To what extend the noise correlation theorem can be applied to real seismological data?

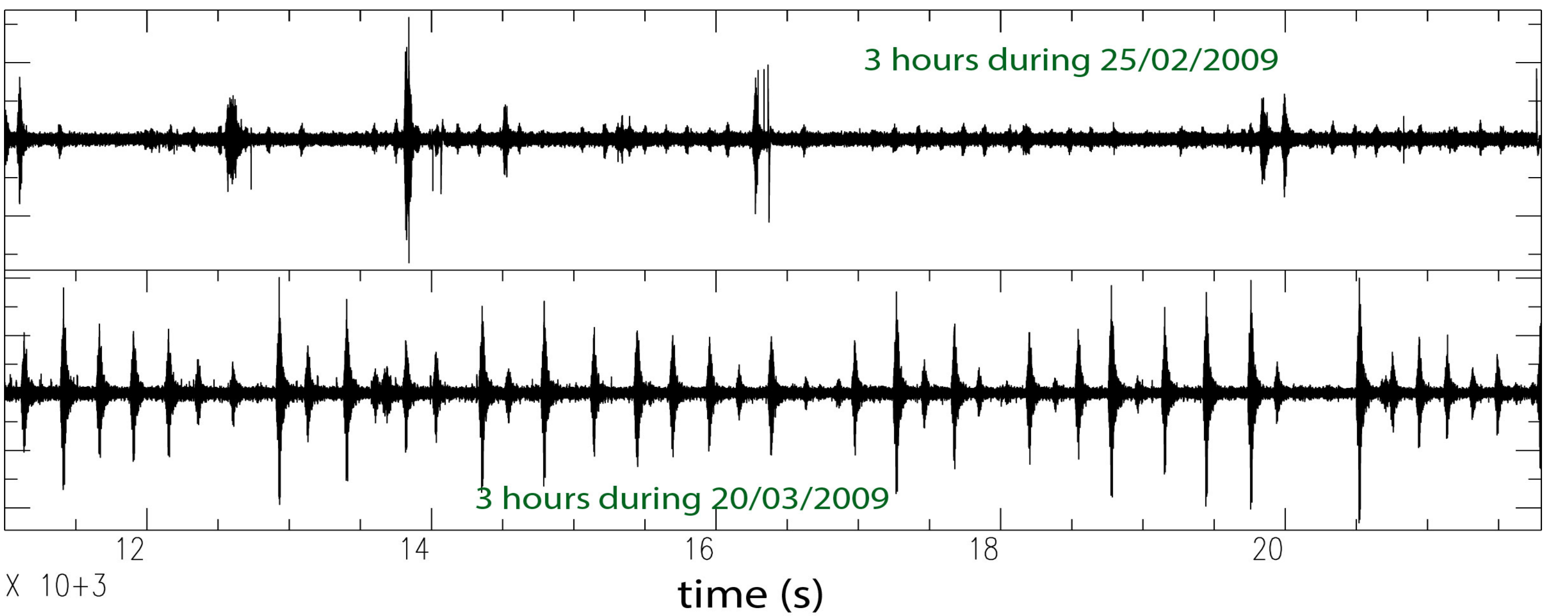
To what extend the real seismic records can be considered as a random diffuse noise?

Examples of seismic records

2 days of continuous record by a seismic station in a subduction zone

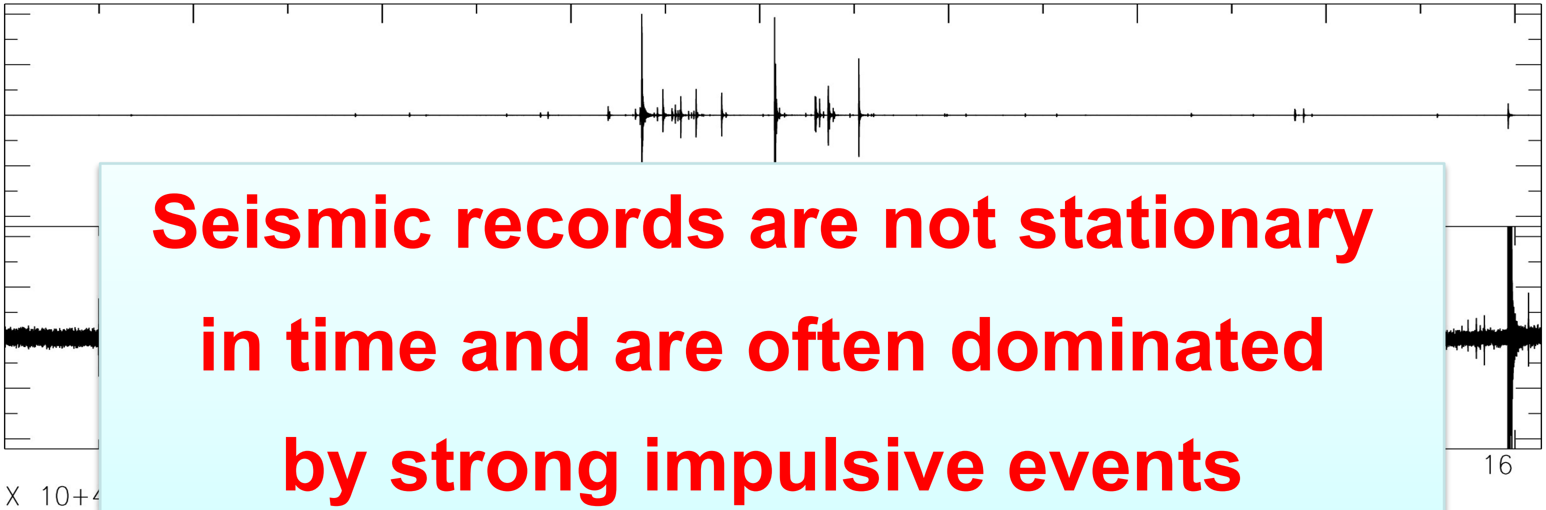


Signals recorded in vicinity of an active volcano



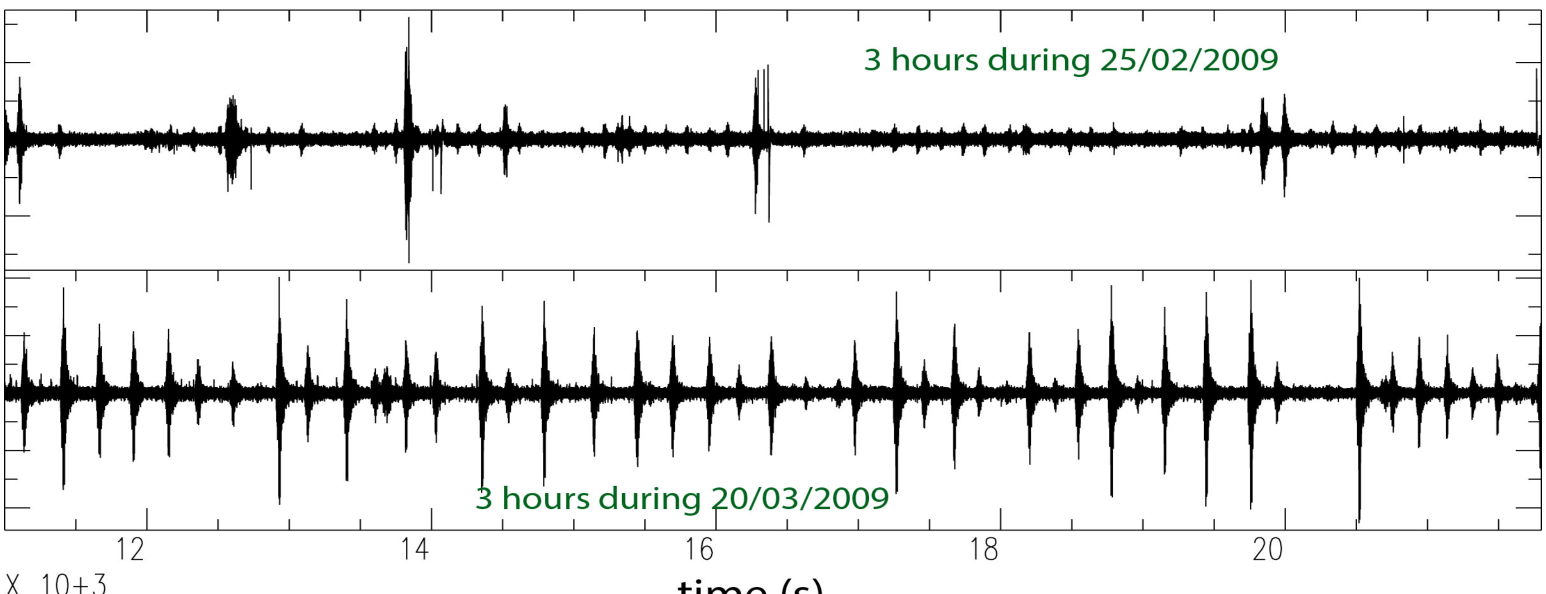
Examples of seismic records

2 days of continuous record by a seismic station in a subduction zone



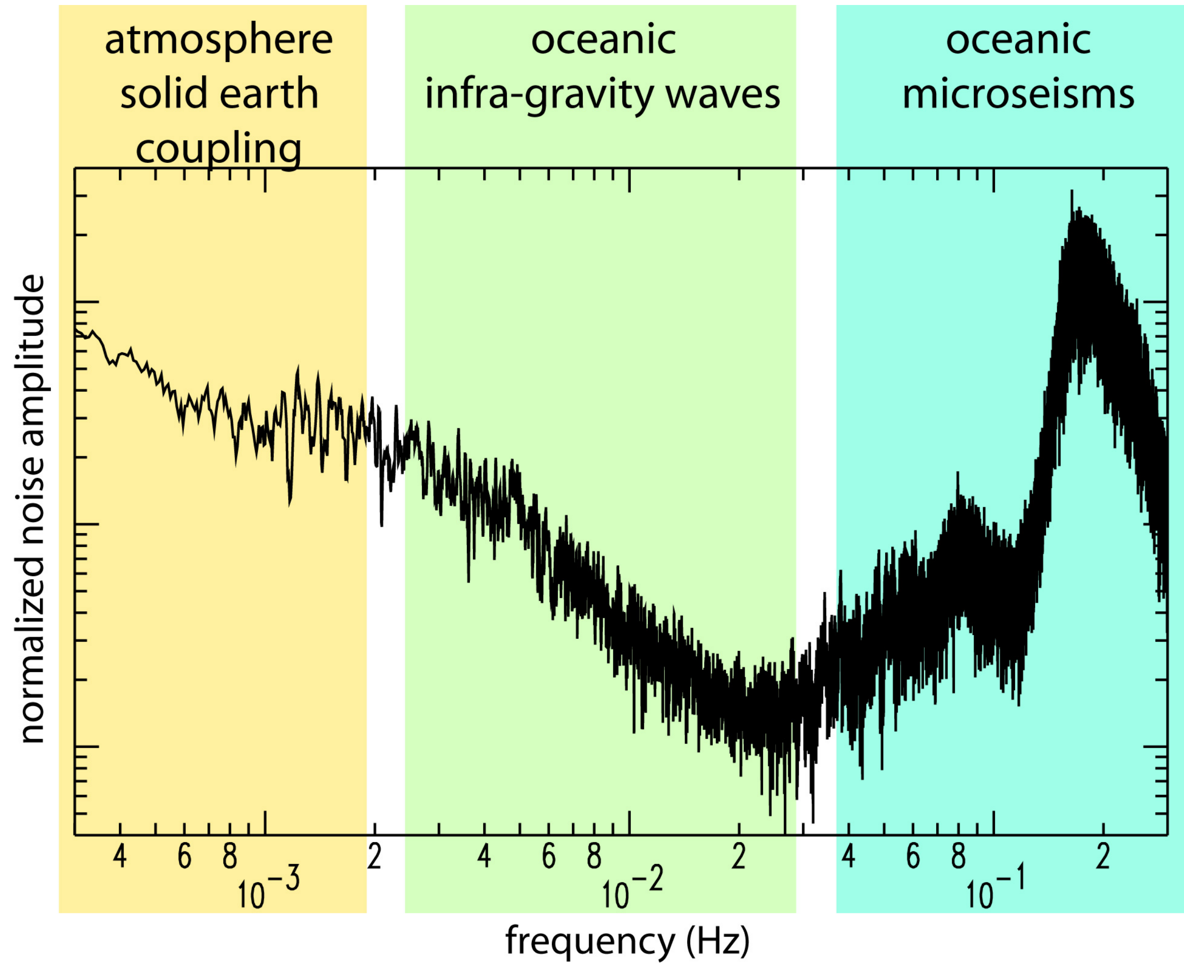
Seismic records are not stationary in time and are often dominated by strong impulsive events

Signals recorded in vicinity of an active volcano



Spectrum of the seismic noise

Fourier spectrum from one day of seismic noise (August 21, 2003; station OBN)



Need for the seismic records preprocessing

Seismic records are not stationary in time

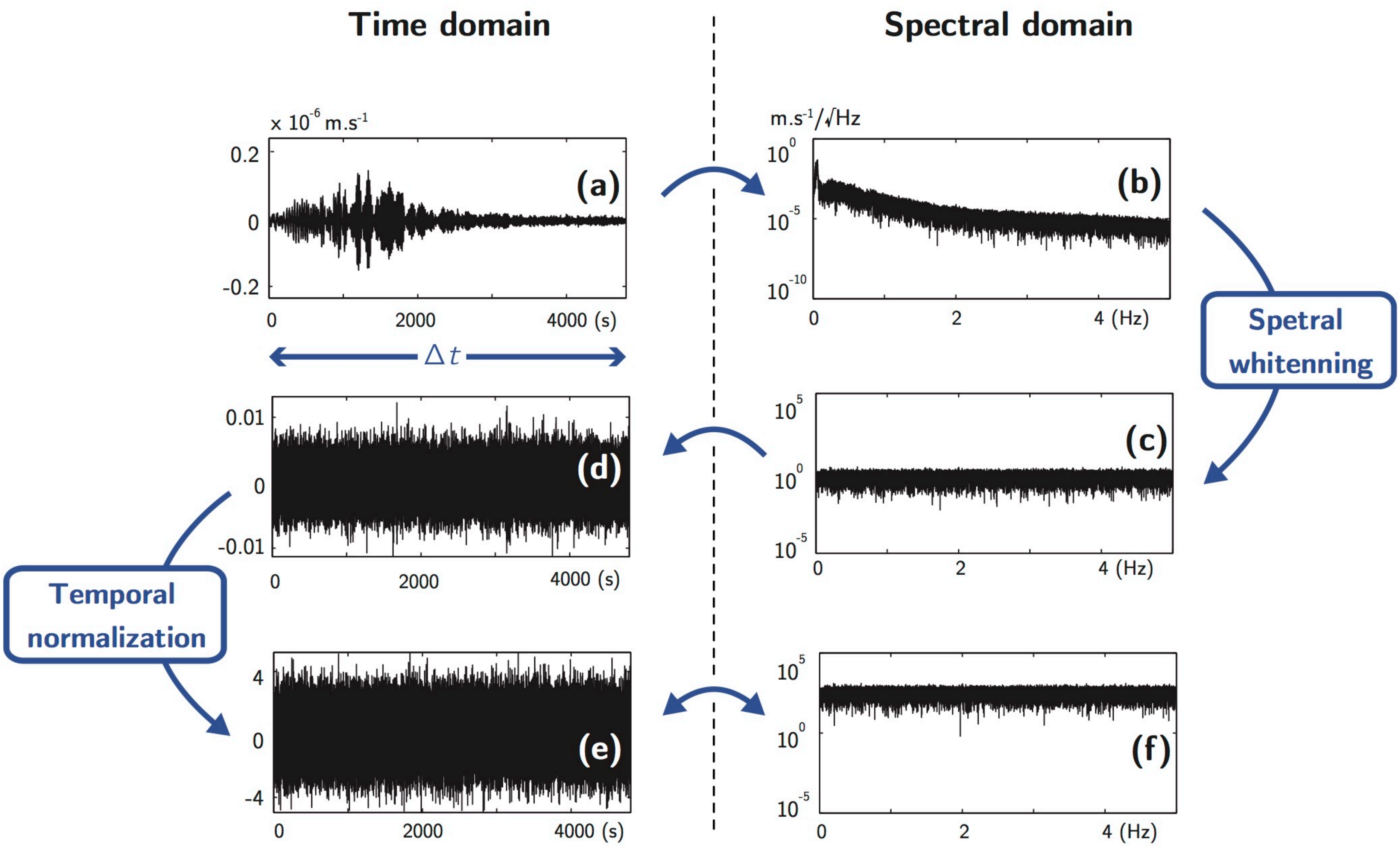
Seismic noise is dominated by strong spectral peaks

Before computing cross-correlations individual seismic records must be preprocessed

- **identification of windows containing strong events**
- **rejection of strong events**
- **equalization of amplitudes in time and spectral domains**

Preprocessing is a complex and often nonlinear set of operations

Noise records pre-processing



Noise Correlation Theorem

For a **random** wavefield with sources distributed **homogeneously** everywhere in the medium it can been shown that:

$$\boxed{\text{correlation of preprocessed records}} \stackrel{?}{=} \frac{-\sigma^2}{4\,a}\left(G_a(\tau,\vec{r}_A,\vec{r}_B) - G_a(-\tau,\vec{r}_A,\vec{r}_B)\right)$$

Green function

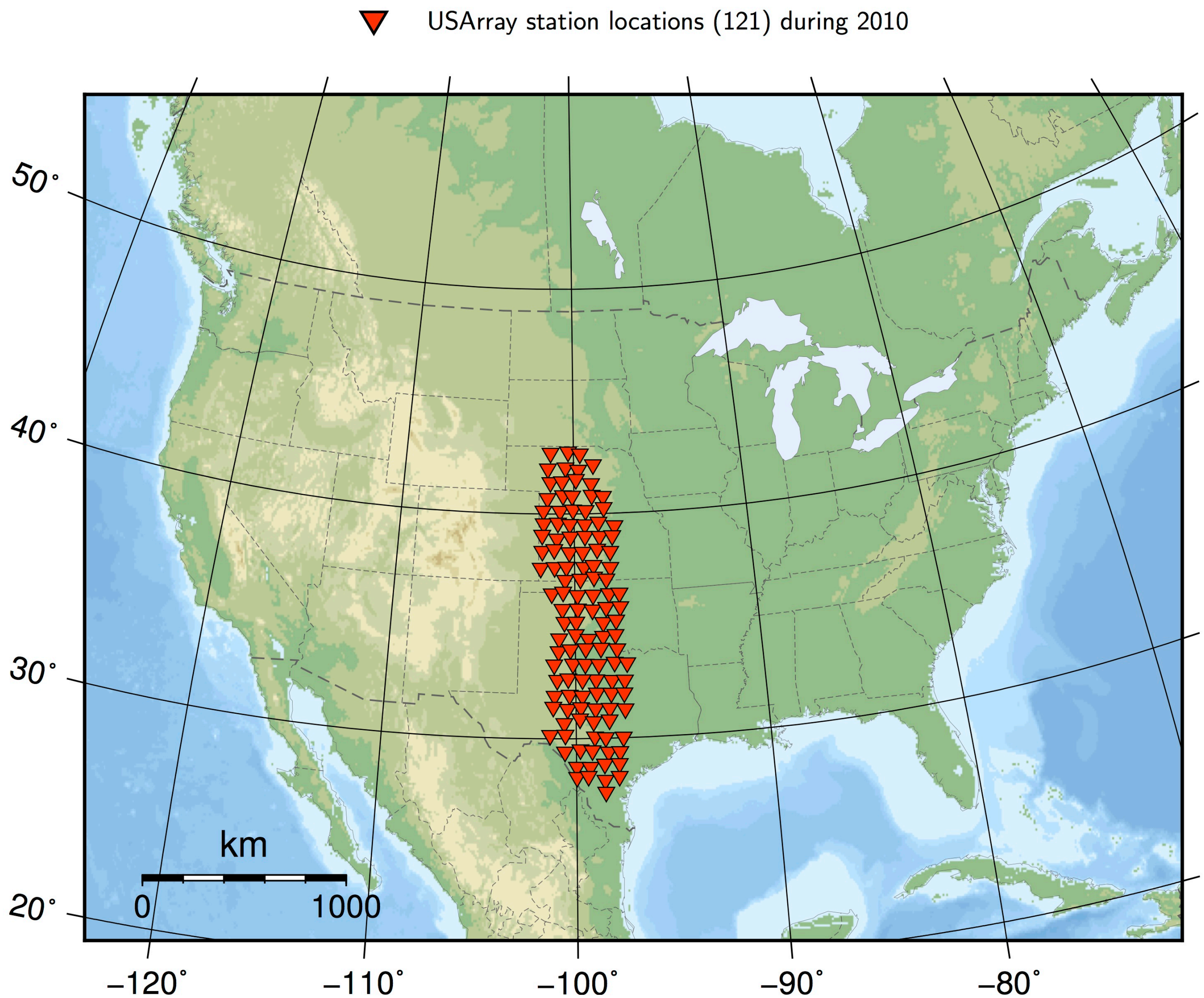
Computing noise cross-correlations between A and B is equivalent to an event occurred at A and recorded at B

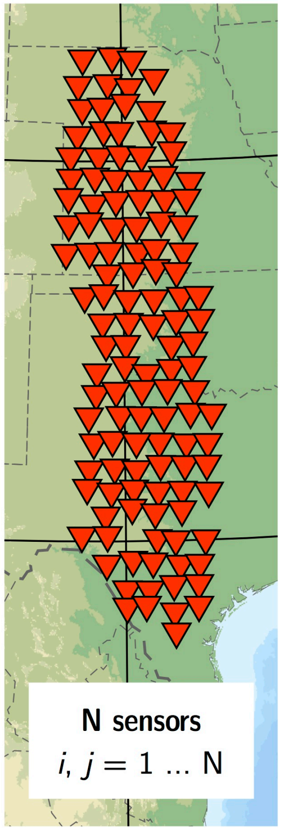
To what extend the preprocessing corrects for the noise time and spectral inhomogeneity?

How can we characterize the structure of the correlated wavefield?

Using network of N sensors as an antenna

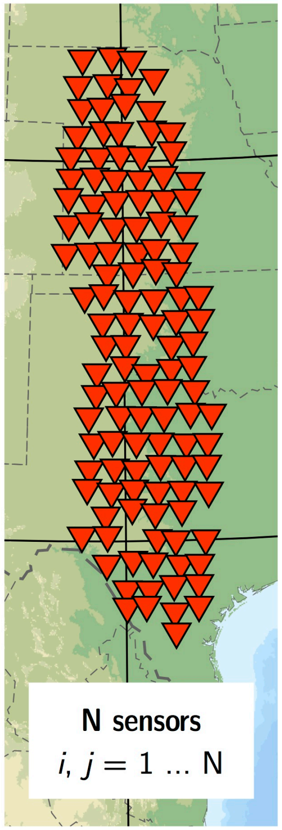
Ensemble of cross-correlations : N-size matrix





The cross-correlation $\boldsymbol{R}_{ij}(\tau)$ of signals recorded by stations i and j is often computed in the frequency domain (faster):

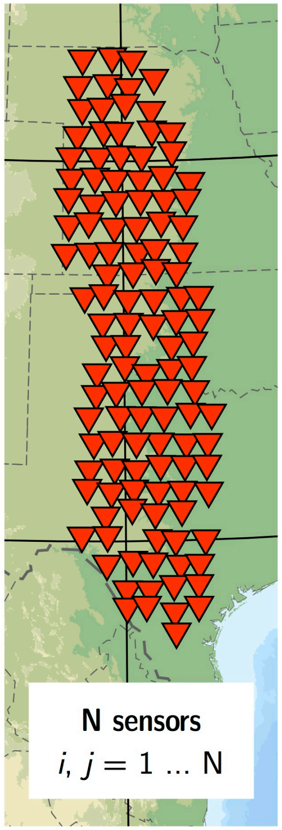
$$\mathbf{C}_{ij}(f) = \mathrm{FT}\left[\mathbf{R}_{ij}(\tau)\right] \quad = \quad \left\langle U_i(f) U_j^*(f) \right\rangle_{\Delta t}$$



The cross-correlation $\boldsymbol{R}_{ij}(\tau)$ of signals recorded by stations i and j is often computed in the frequency domain (faster):

$$\mathbf{C}_{ij}(f) = \mathrm{FT}\left[\mathbf{R}_{ij}(\tau)\right] = \left\langle U_i(f)U_j^*(f)\right\rangle_{\Delta t}$$

The array covariance matrix $\boldsymbol{C}_{ij}(f)$ between stations i and j is estimated from the time average over Δt of the product of $U_i(f)$ and $U_j^*(f)$:



The cross-correlation $\boldsymbol{R}_{ij}(\tau)$ of signals recorded by stations i and j is often computed in the frequency domain (faster):

$$\mathbf{C}_{ij}(f) = \mathrm{FT}\left[\mathbf{R}_{ij}(\tau)\right] = \left\langle U_i(f)\, U_j^*(f) \right\rangle_{\Delta t}$$

The array covariance matrix $\boldsymbol{C}_{ij}(f)$ between stations i and j is estimated from the time average over Δt of the product of $U_i(f)$ and $U_j^*(f)$:

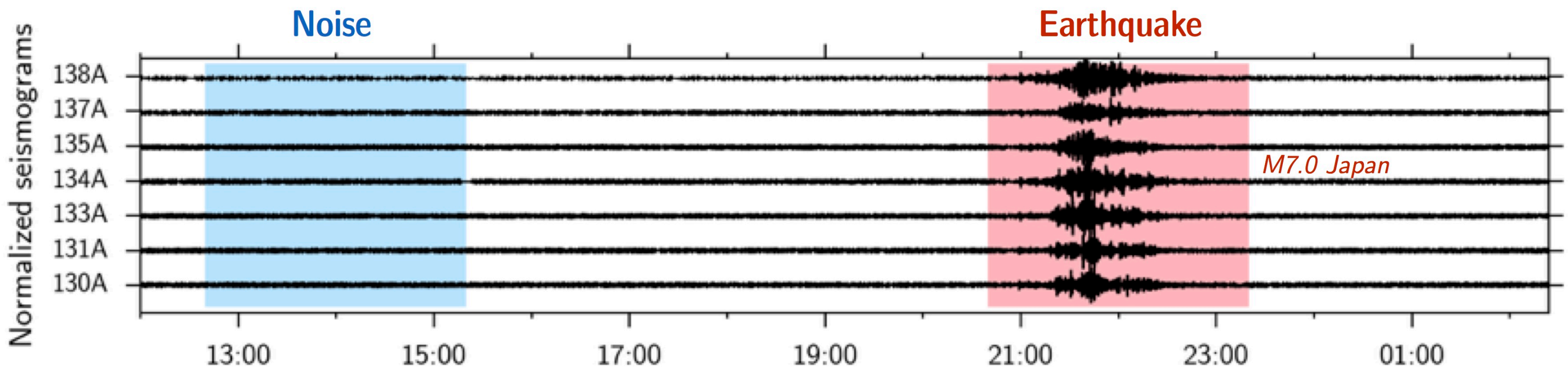
Covariance matrix spectrum: sorted eigenvalues

$$\lambda_n(f) \quad := \quad \lambda_0(f) > \lambda_1(f) > \dots > \lambda_N(f)$$

Relationship between the coherence of the wavefield and the covariance matrix eigenvalues ?

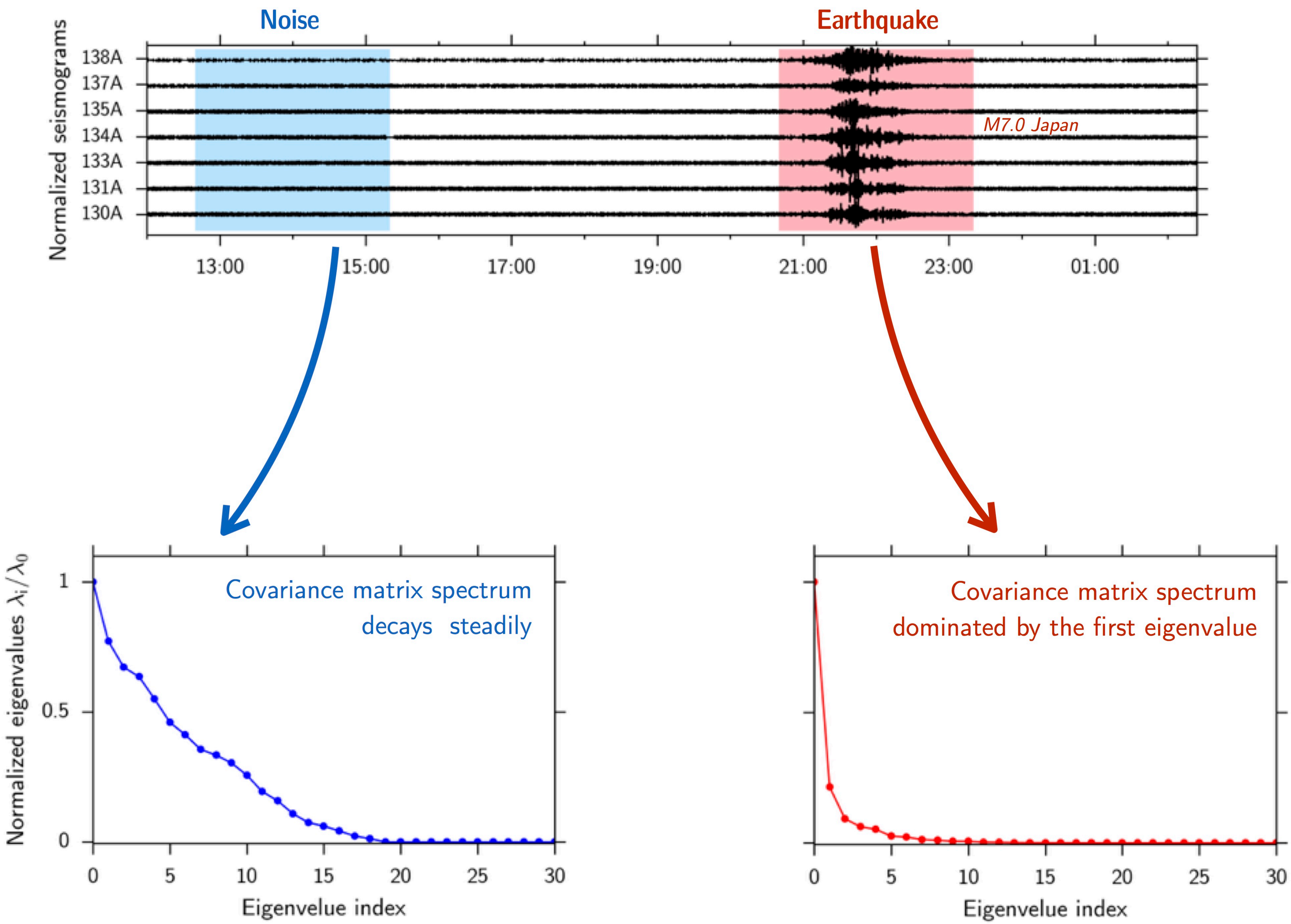
Array covariance matrix eigenvalues and wavefield coherence

Seydoux et al., GJI, 2016



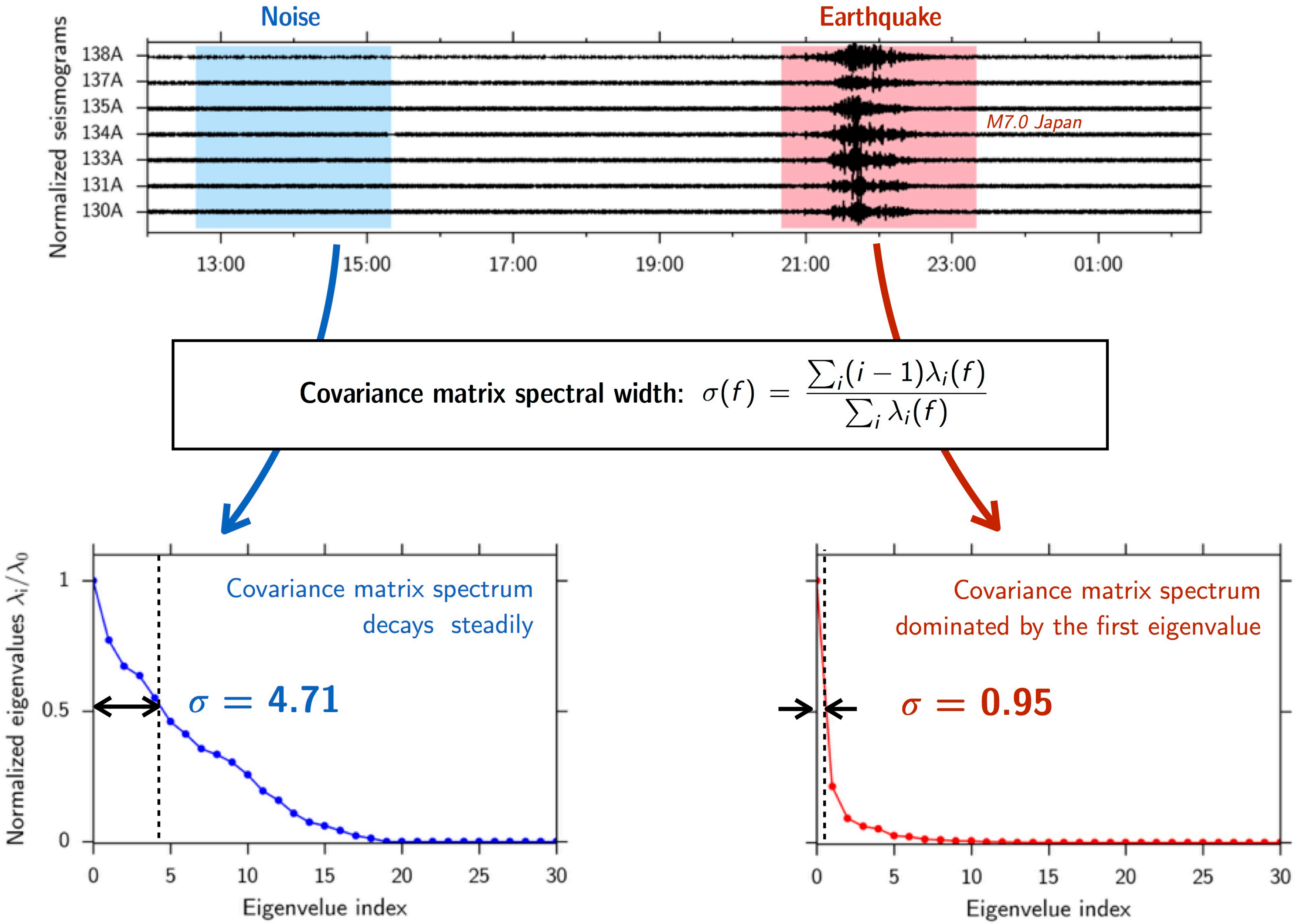
Array covariance matrix eigenvalues and wavefield coherence

Seydoux et al., GJI, 2016



Array covariance matrix eigenvalues and wavefield coherence

Seydoux et al., GJI, 2016

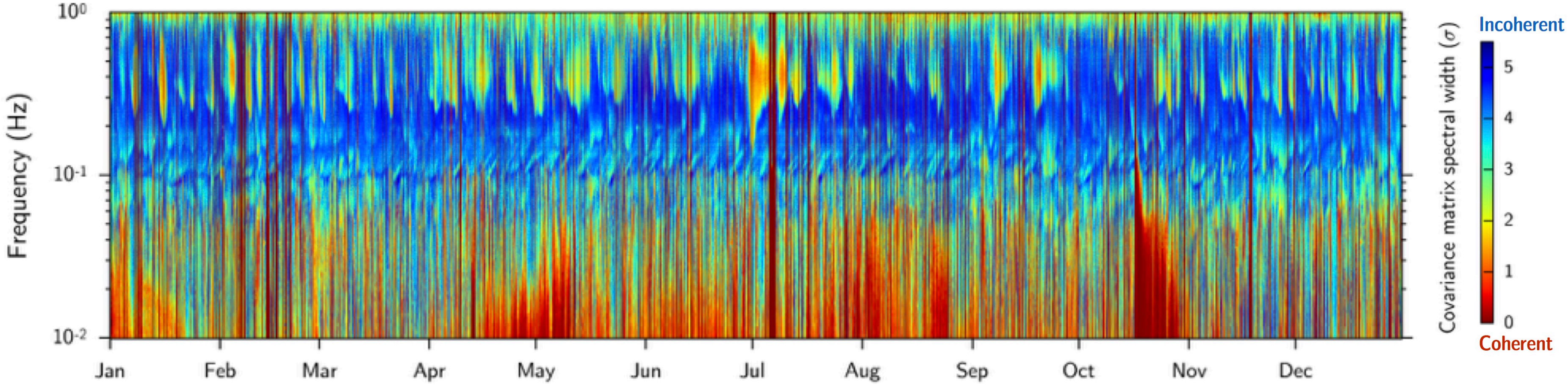


Covariance matrix spectral width: $\sigma(f) = \dfrac{\sum_i (i-1)\lambda_i(f)}{\sum_i \lambda_i(f)}$

Effect of the energy equalization on the covariance matrix spectral width

Seydoux et al., subm. to GRL

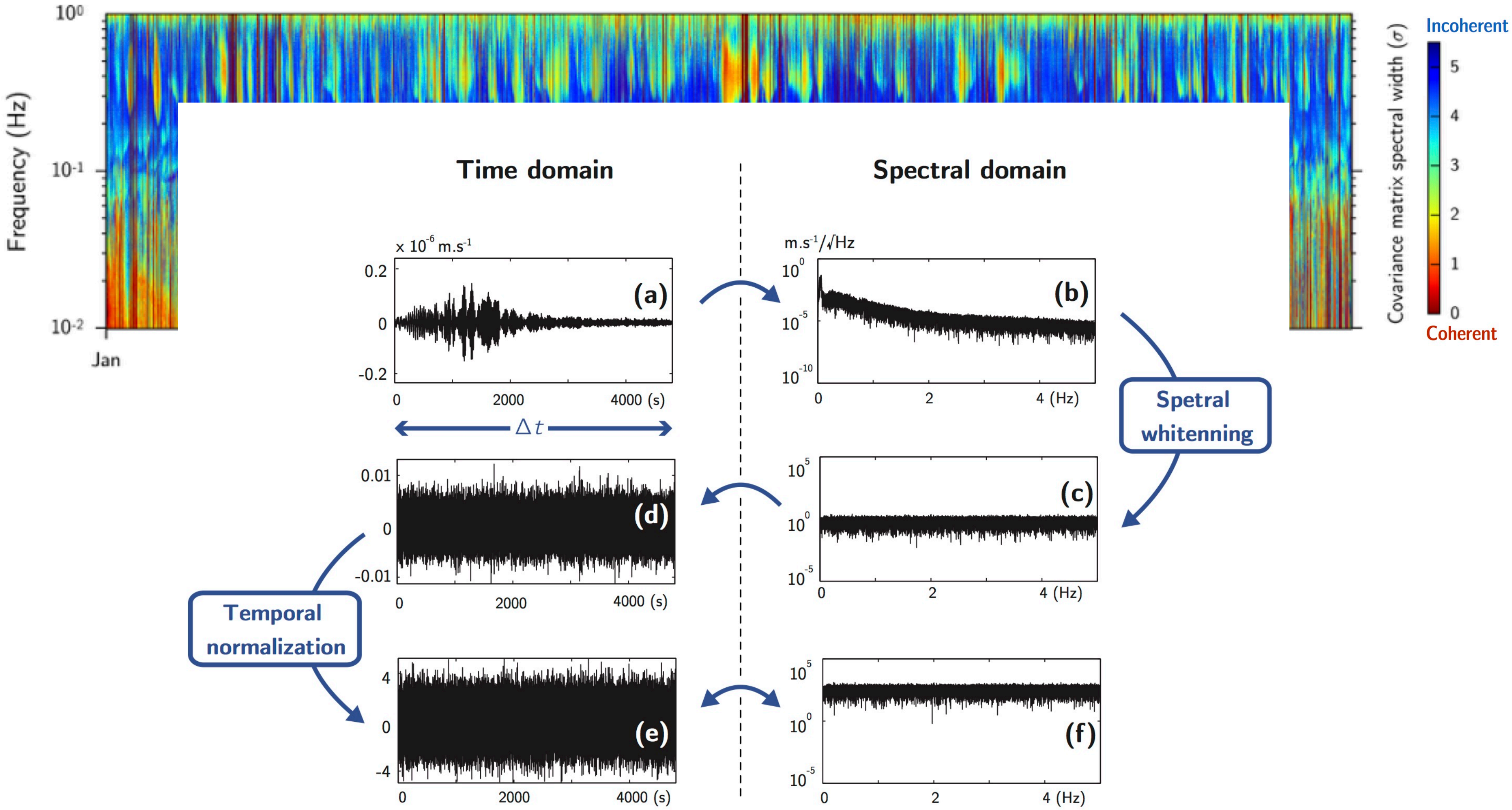
From **raw data** (no equalization applied)



Effect of the energy equalization on the covariance matrix spectral width

Seydoux et al., subm. to GRL

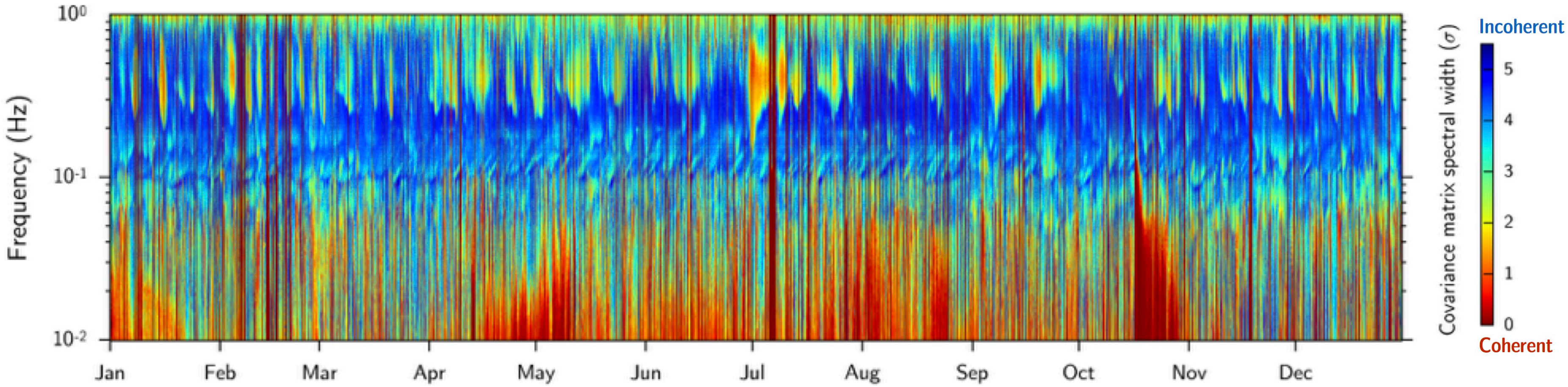
From **raw data** (no equalization applied)



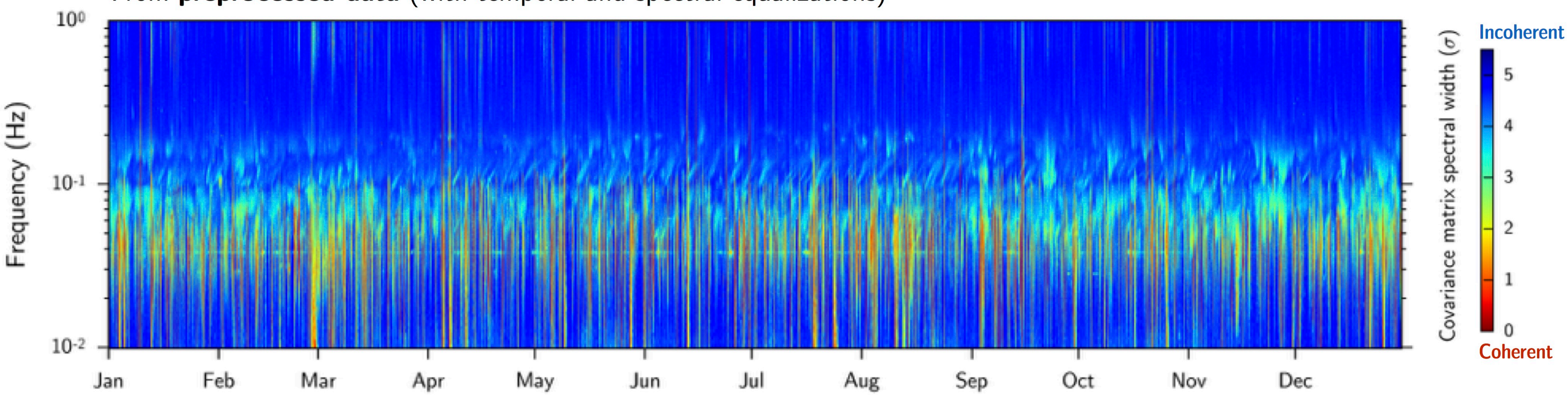
Effect of the energy equalization on the covariance matrix spectral width

Seydoux et al., subm. to GRL

From **raw data** (no equalization applied)



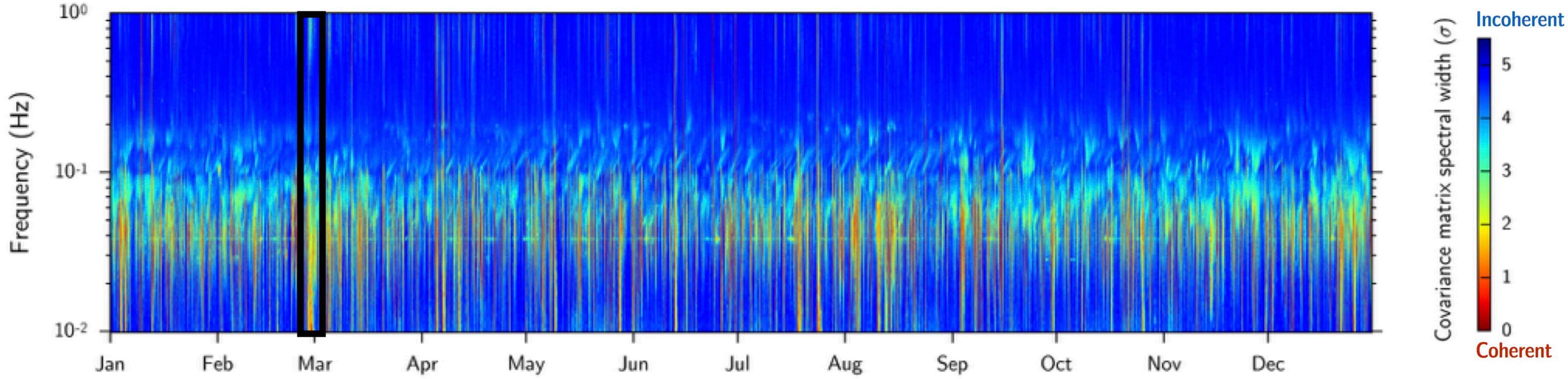
From **preprocessed data** (with temporal and spectral equalizations)



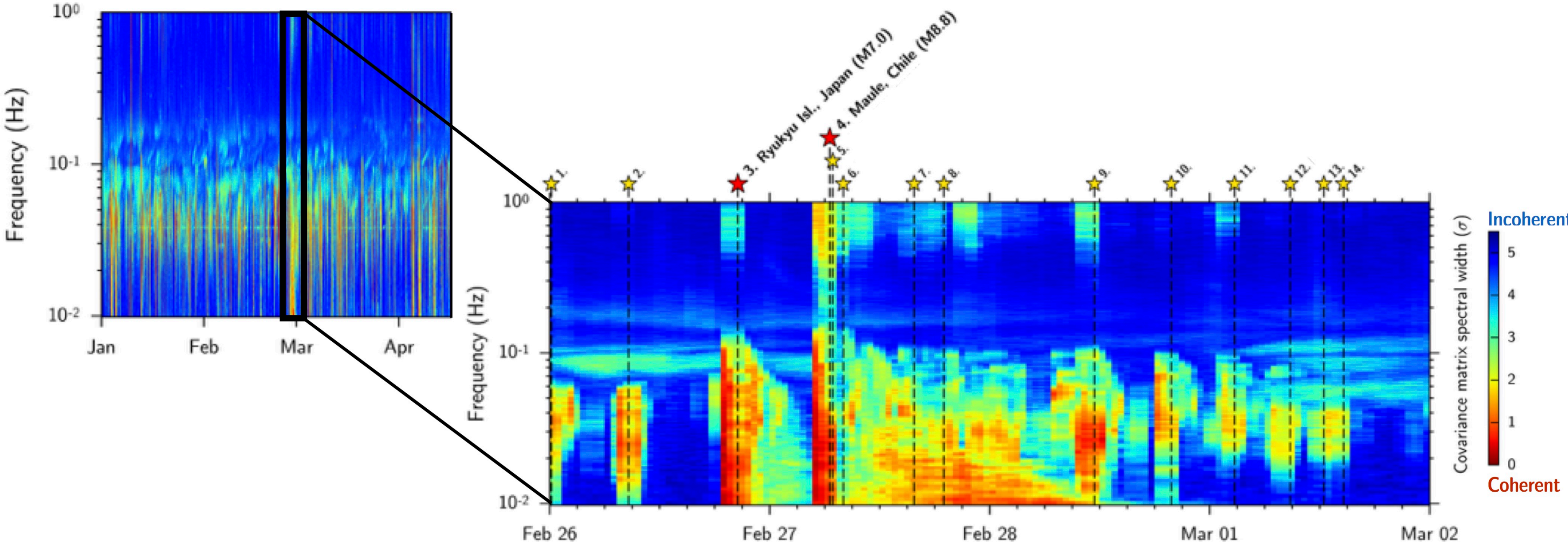
Equalization process clearly improves the seismogram stationarity

However, coherent signals are still present after the equalization

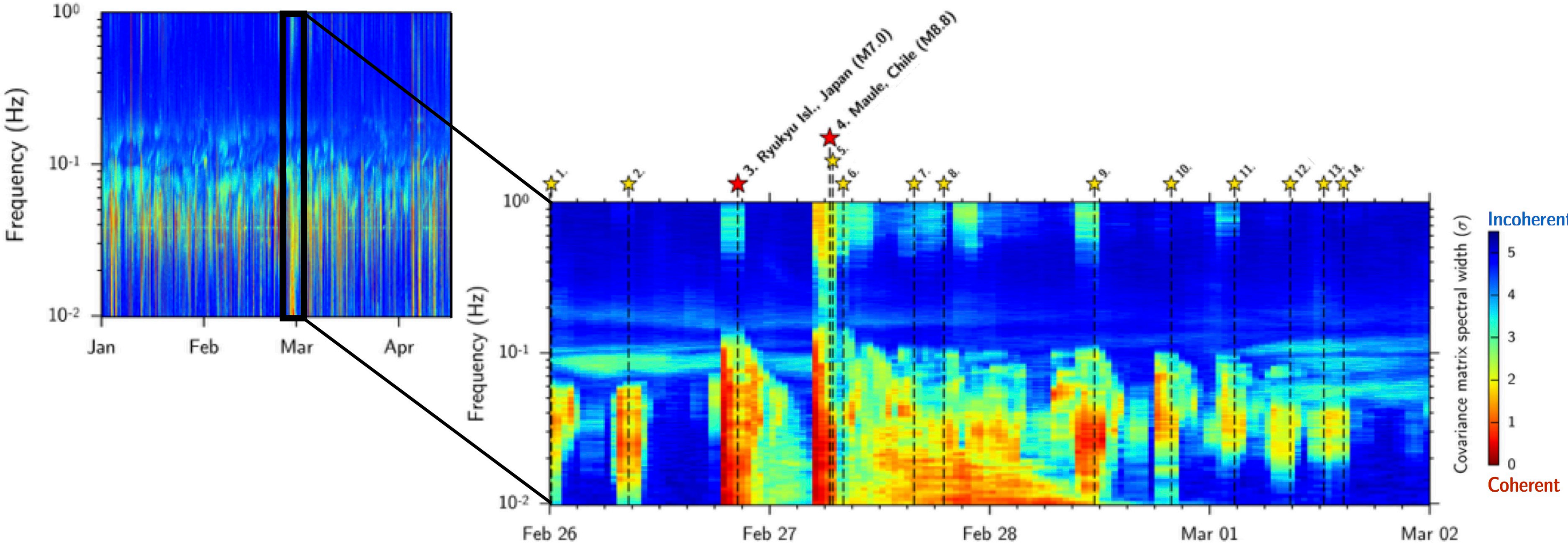
Earthquakes still induce drops of the covariance matrix spectral width

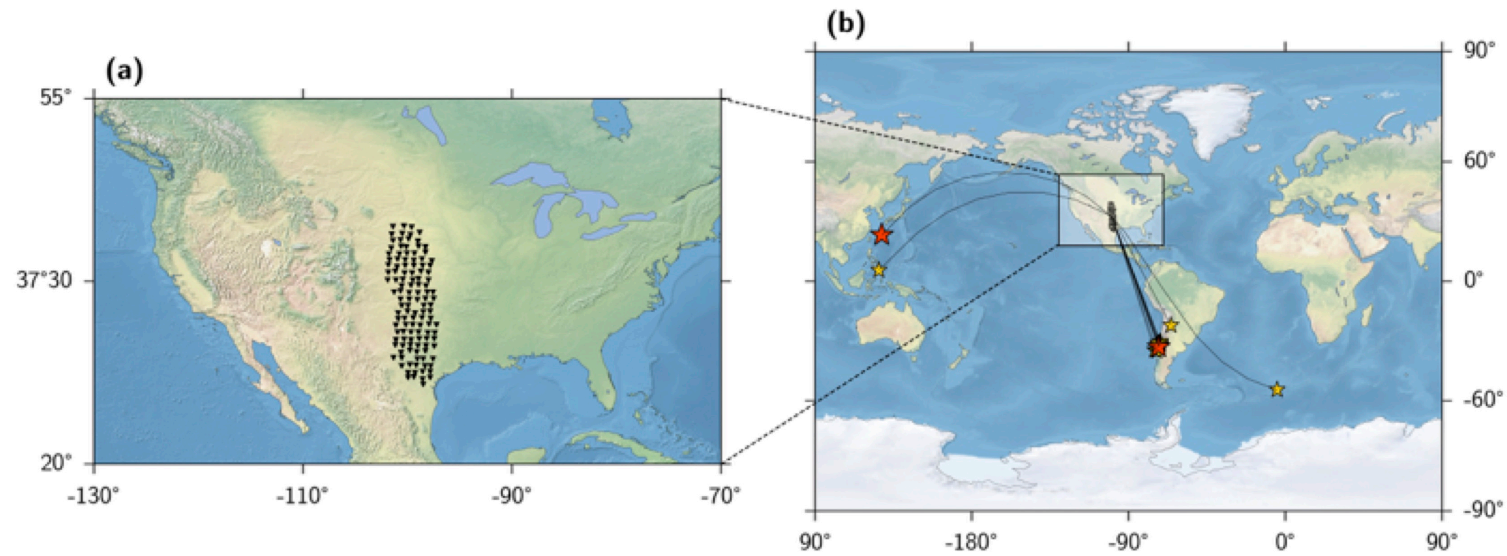


Earthquakes still induce drops of the covariance matrix spectral width

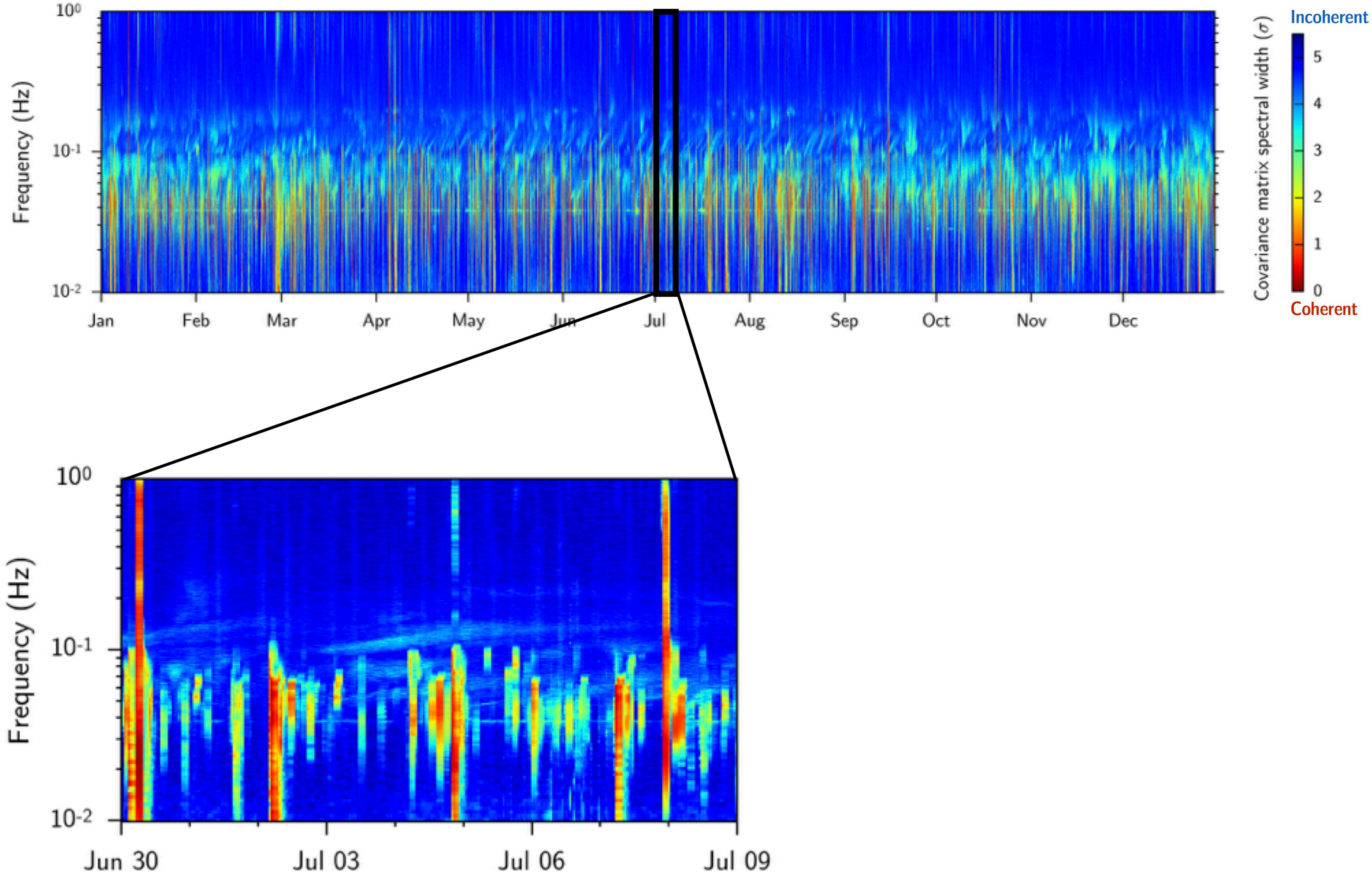


Earthquakes still induce drops of the covariance matrix spectral width

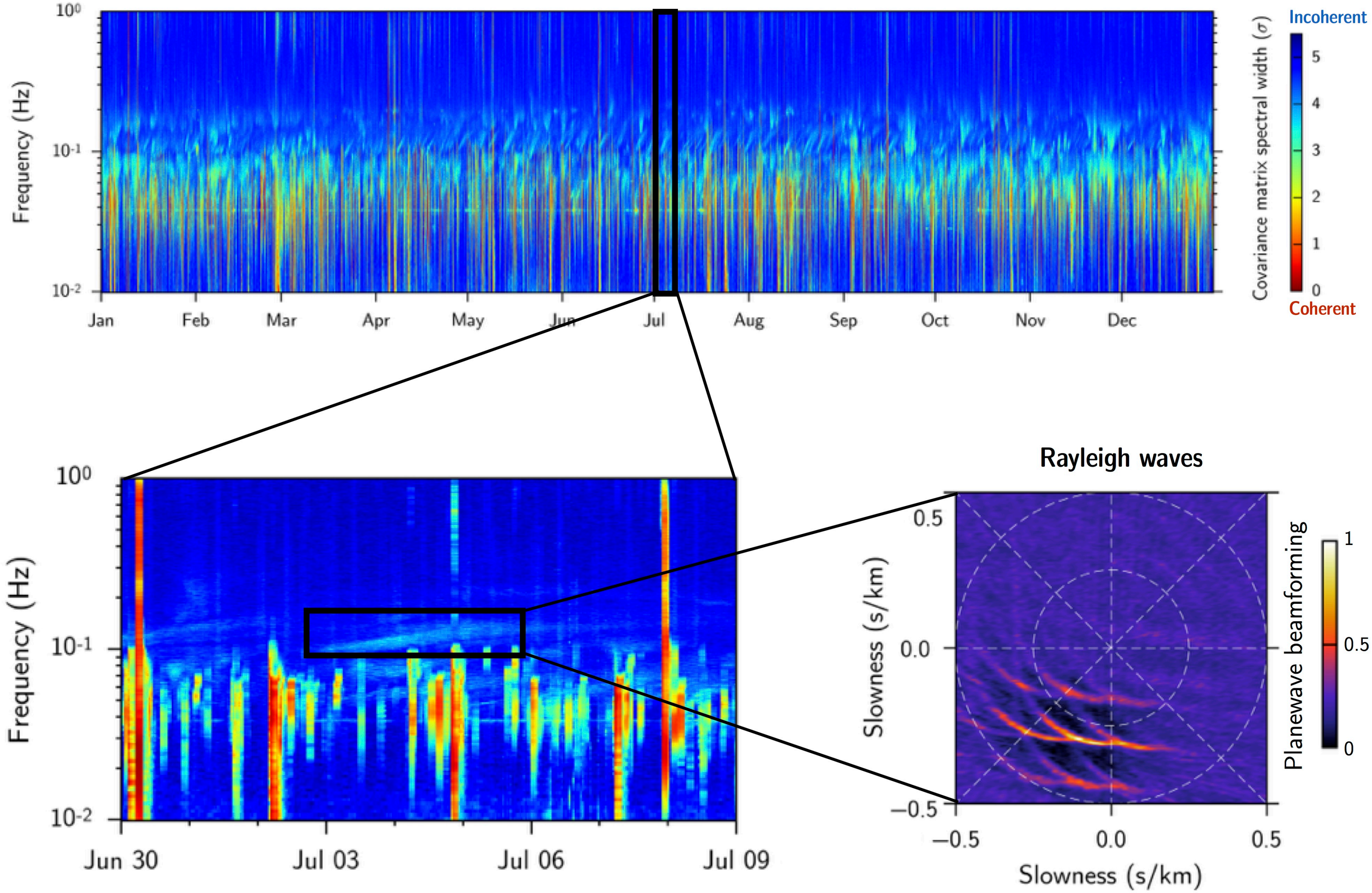




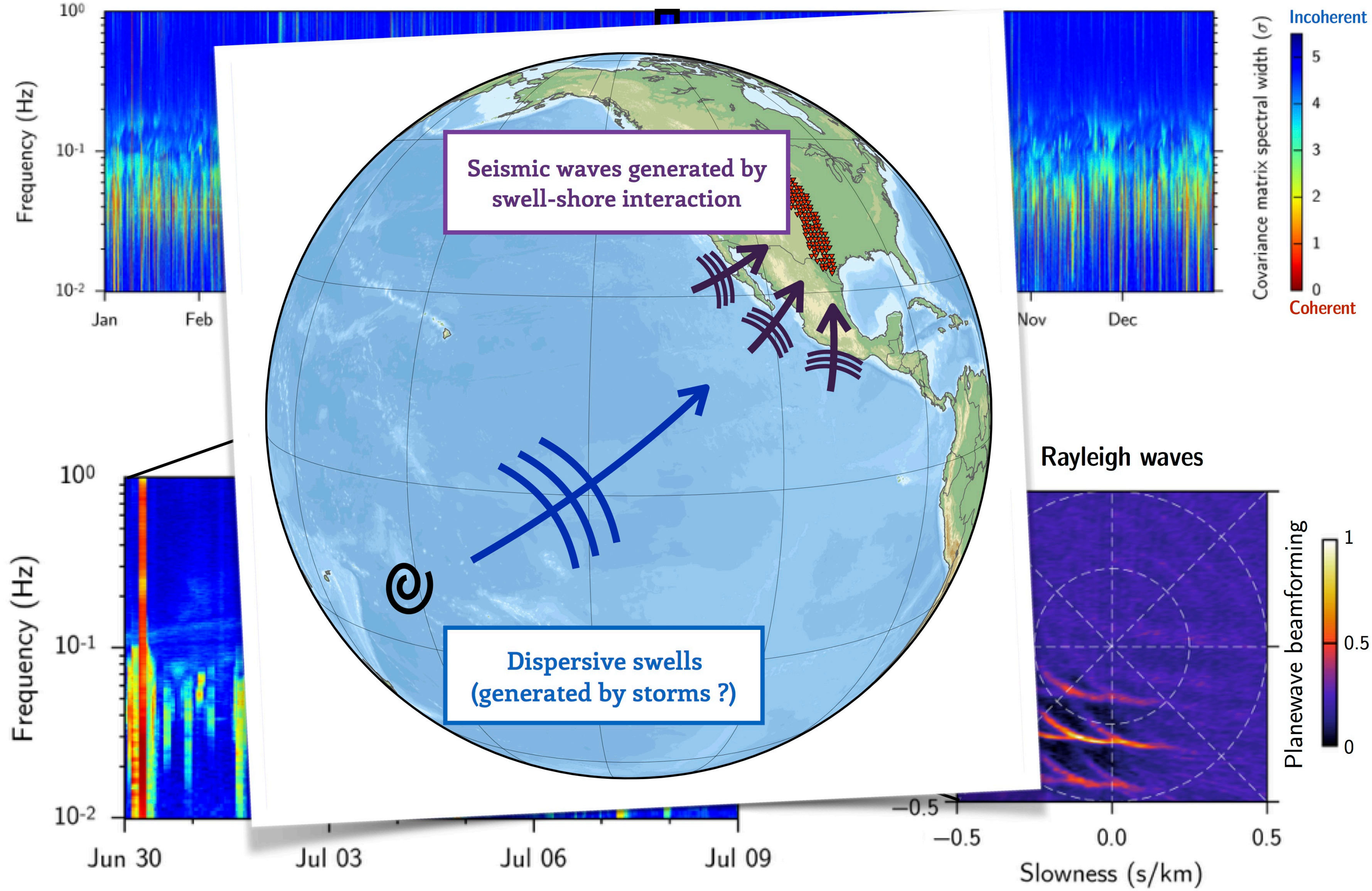
Dispersive signals are still visible around 0.1 Hz



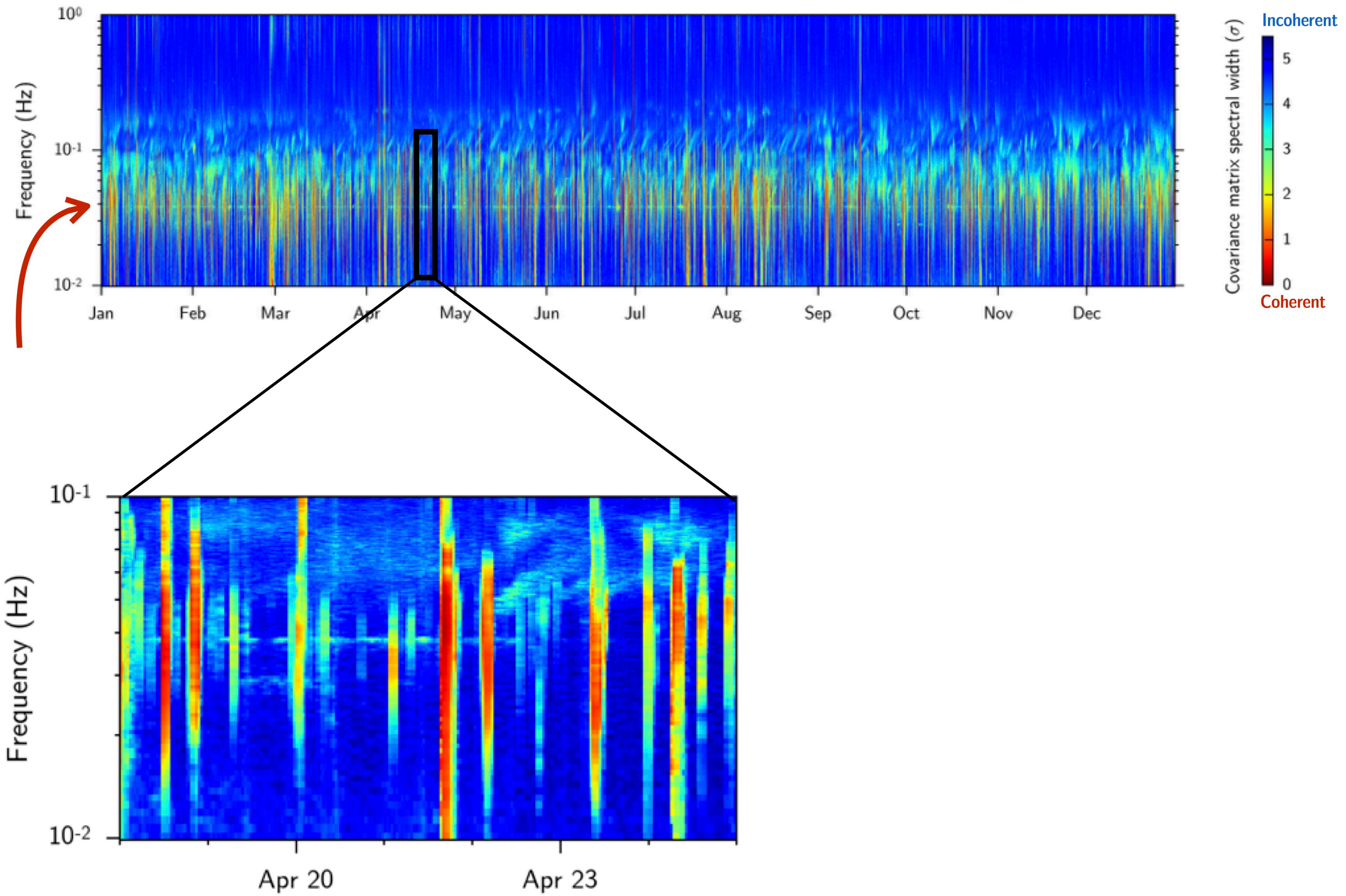
Dispersive signals are still visible around 0.1 Hz



Dispersive signals are still visible around 0.1 Hz

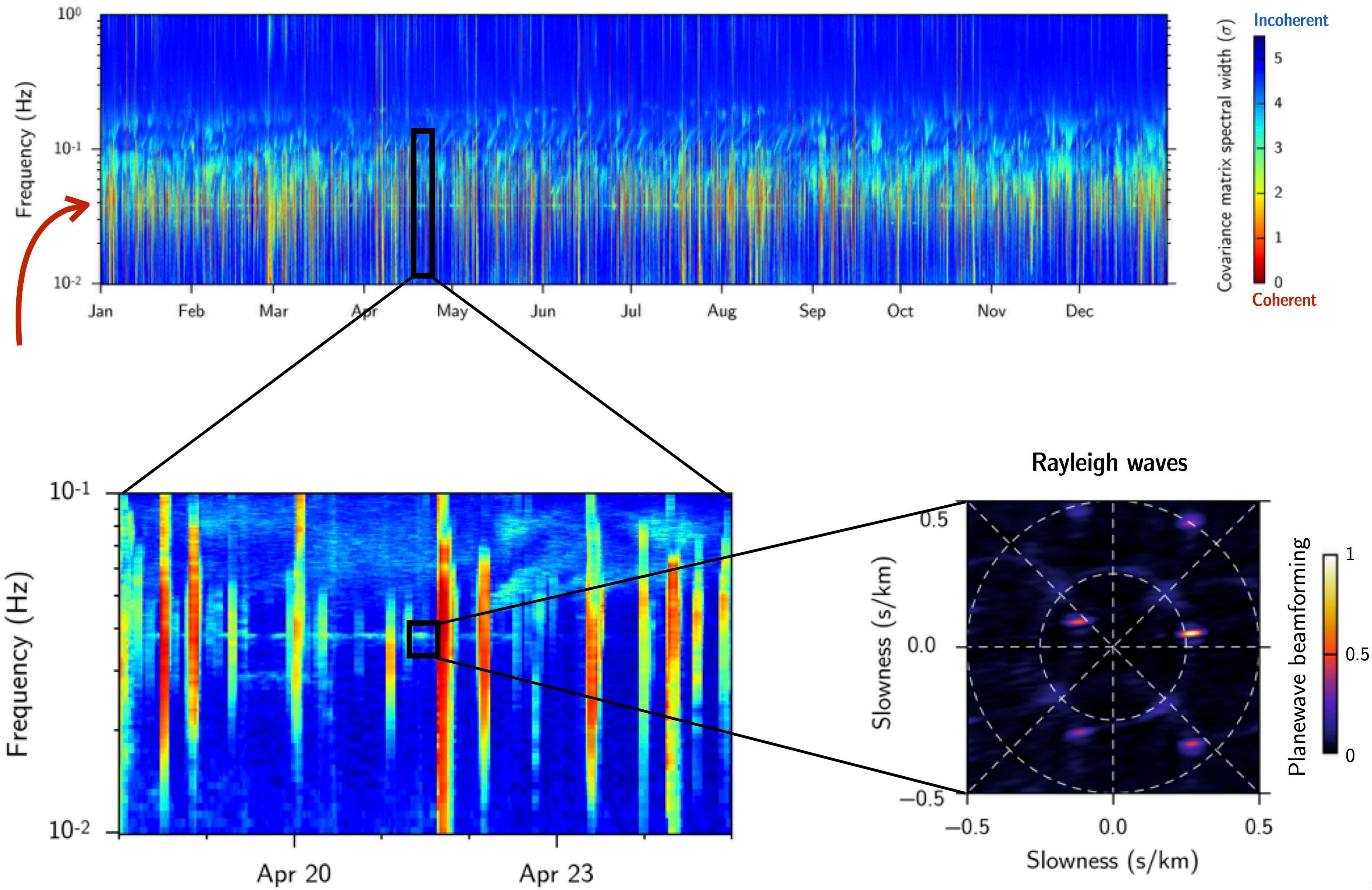


A nearly-continuous and quasi-monochromatic signal is still visible around 26-sec of period



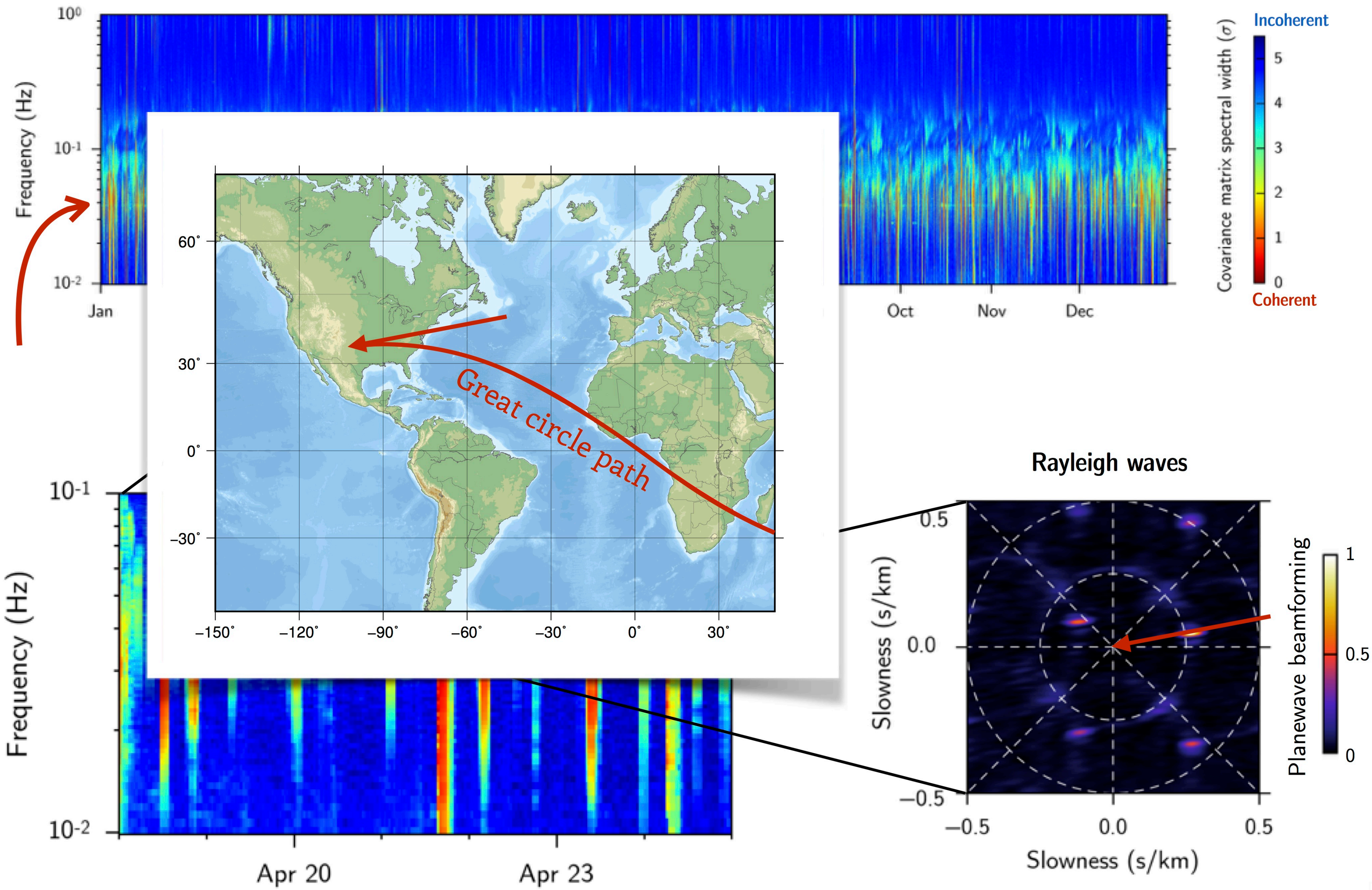
Detection of a monochromatic signal

A nearly-continuous and quasi-monochromatic signal is still visible around 26-sec of period

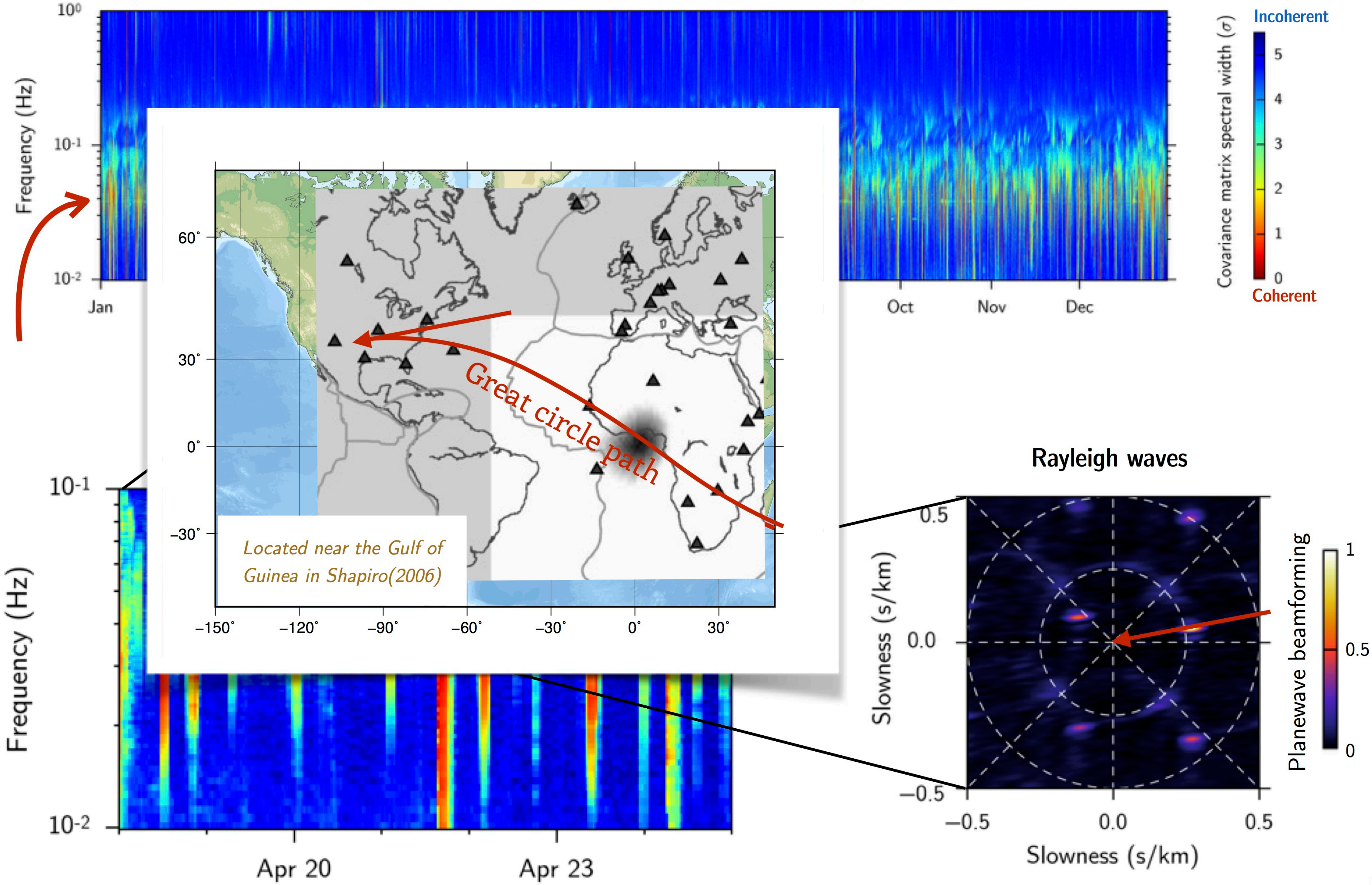


Detection of a monochromatic signal

A nearly-continuous and quasi-monochromatic signal is still visible around 26-sec of period



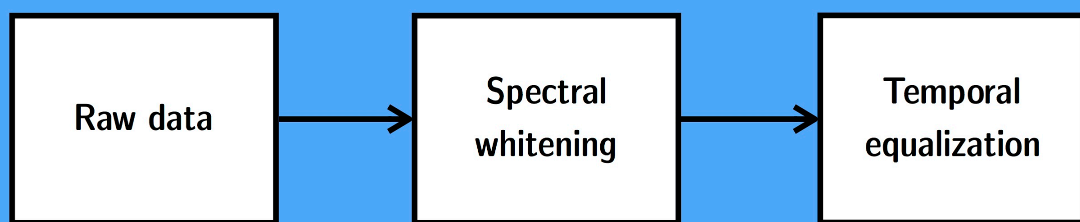
A nearly-continuous and quasi-monochromatic signal is still visible around 26-sec of period



- The covariance matrix spectral width is a robust estimator of the wavefield coherence

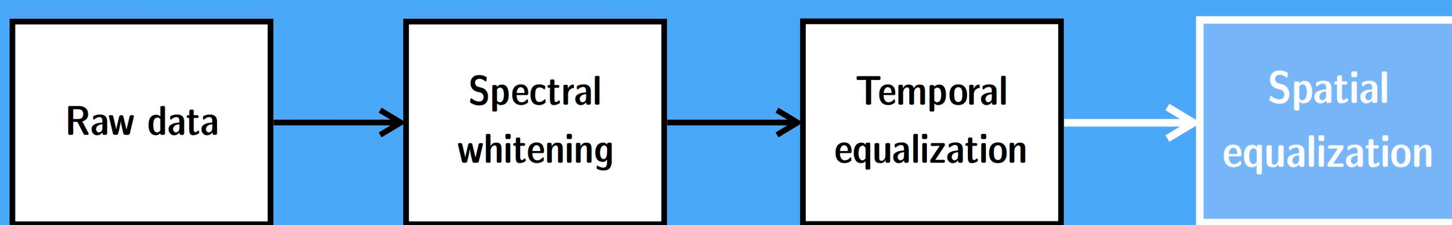
- It can detect weak and emergent signals with low signal-to-noise ratio

- Temporal and spectral equalizations only partially correct for the inhomogeneous source distribution



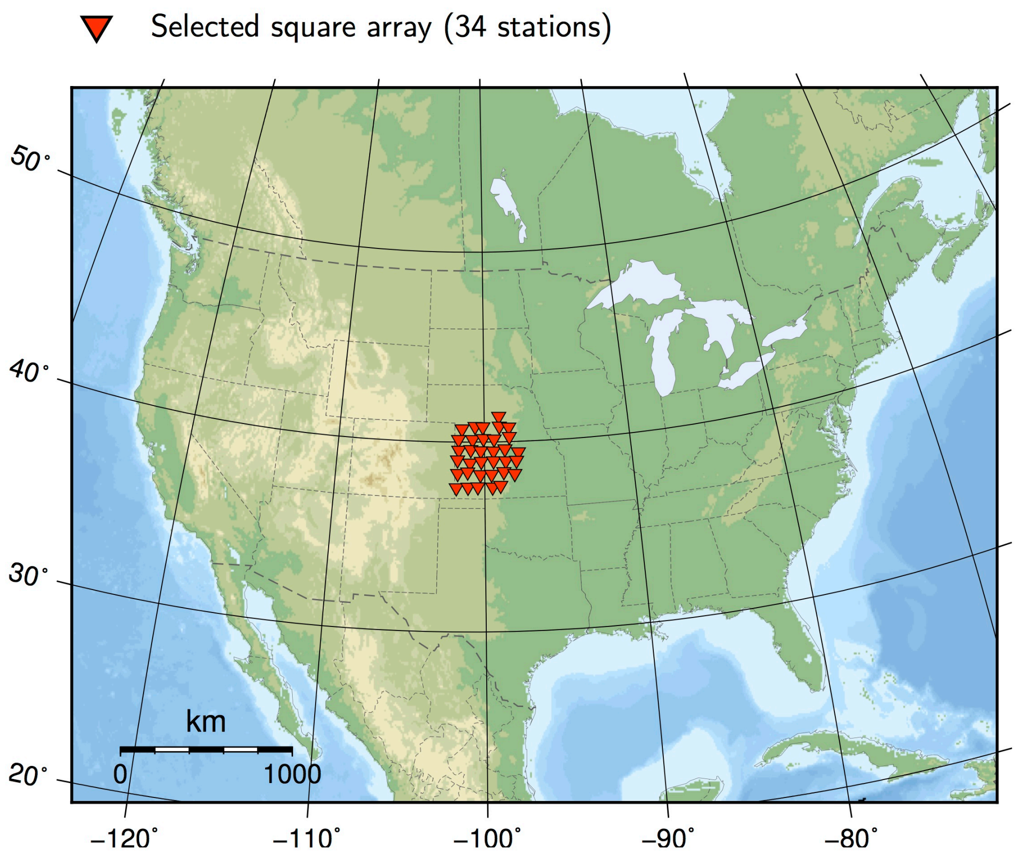
How to deal with the signals that resist this equalization?

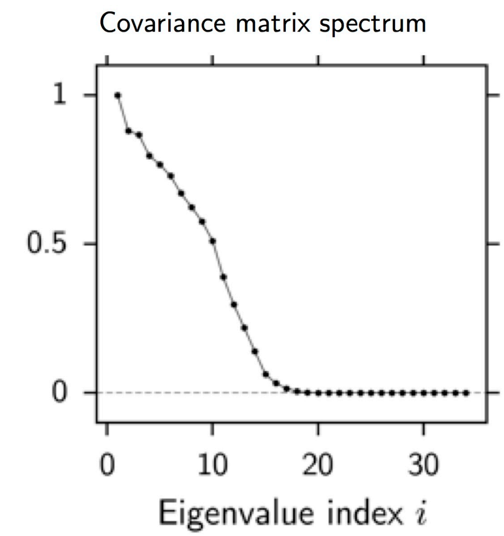
- **The covariance matrix spectral width is a robust estimator of the wavefield coherence**
- **It can detect weak and emergent signals with low signal-to-noise ratio**
- **Temporal and spectral equalizations only partially correct for the inhomogeneous source distribution**

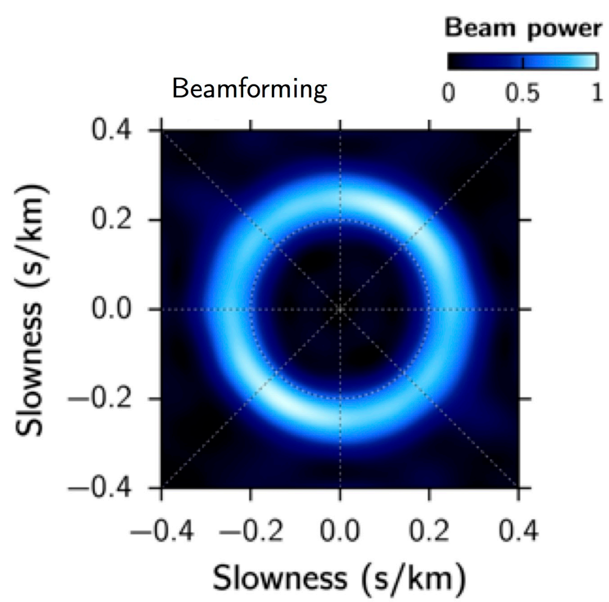


We could also equalize the covariance matrix spectrum

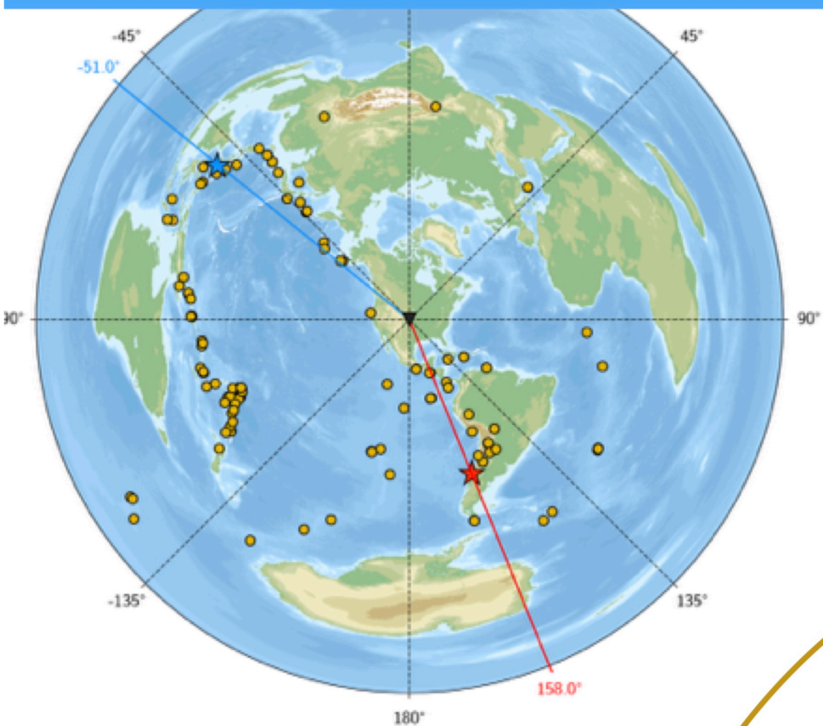
Synthetic isotropic noise seen by a square array

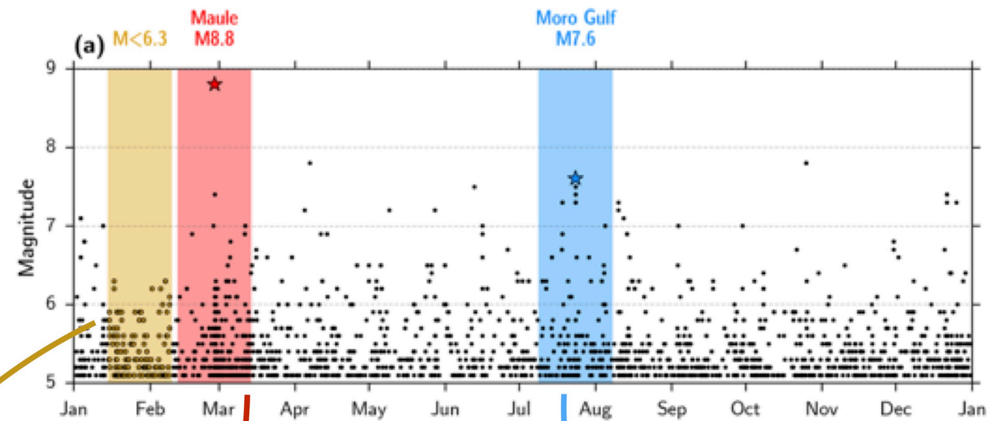




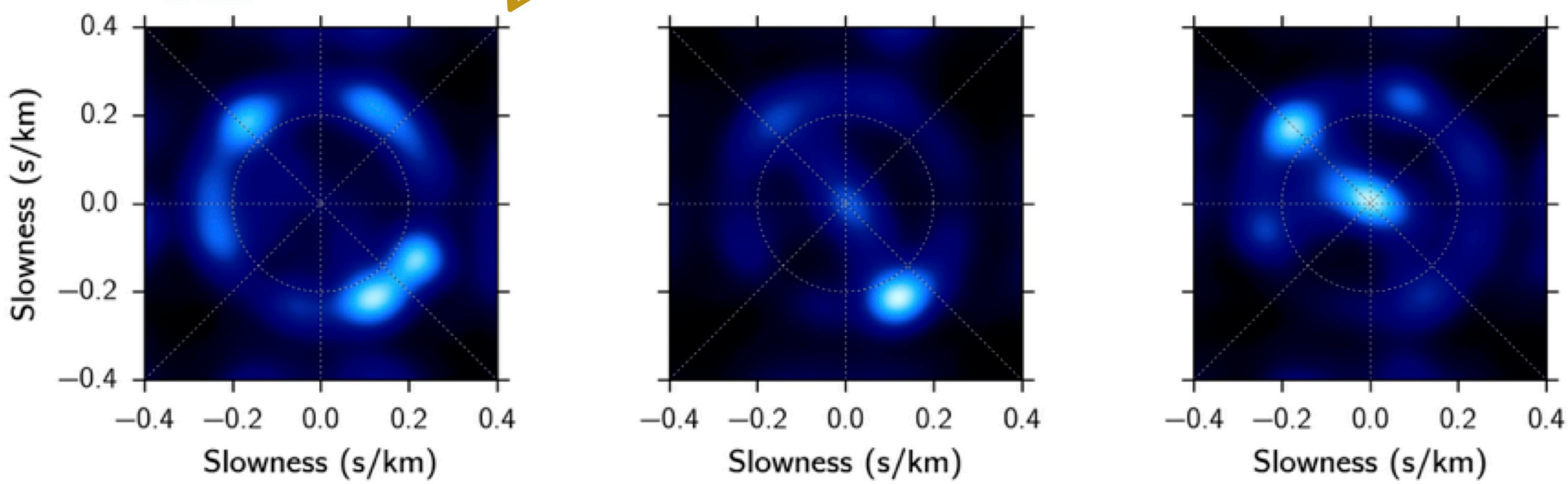


Beamforming on real data with spectral and temporal equalization



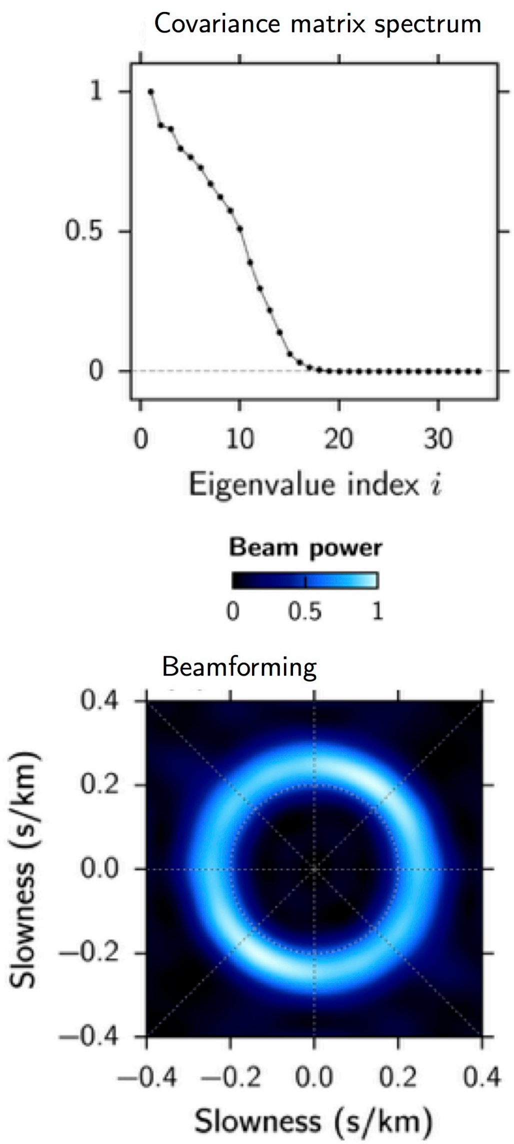


Beamforming nearly circular

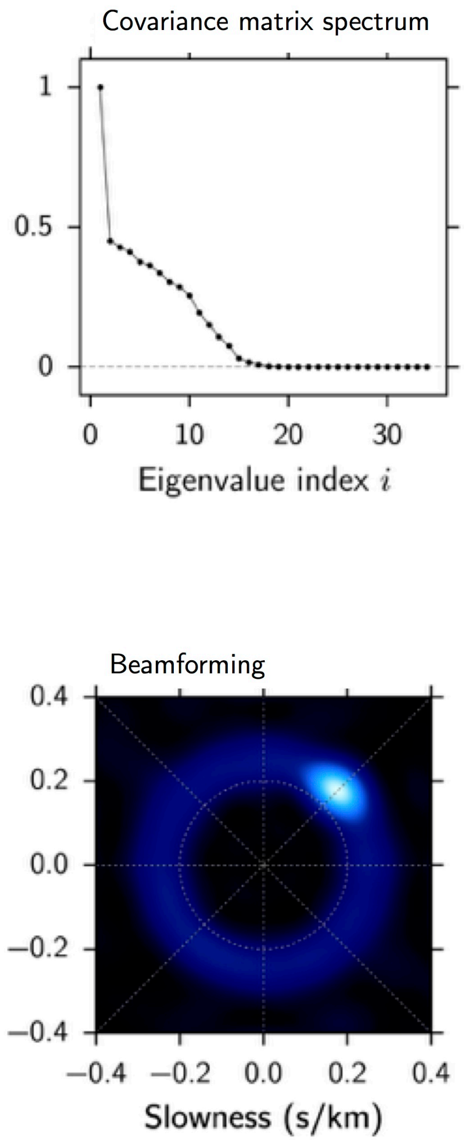


Spatial equalization (synthetic case)

Isotropic noise
(what we want)

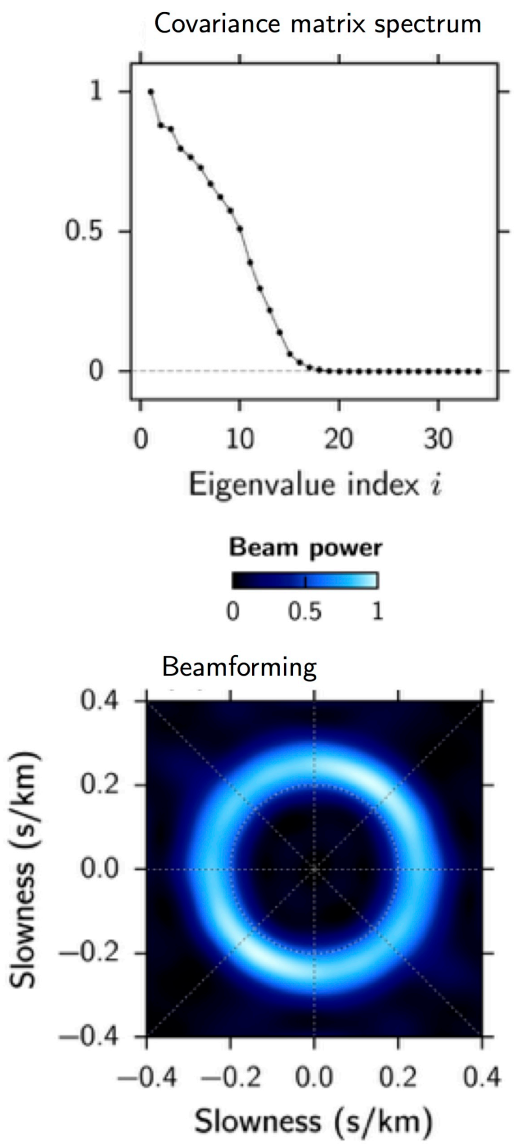


Isotropic noise + source
(what we actually have)

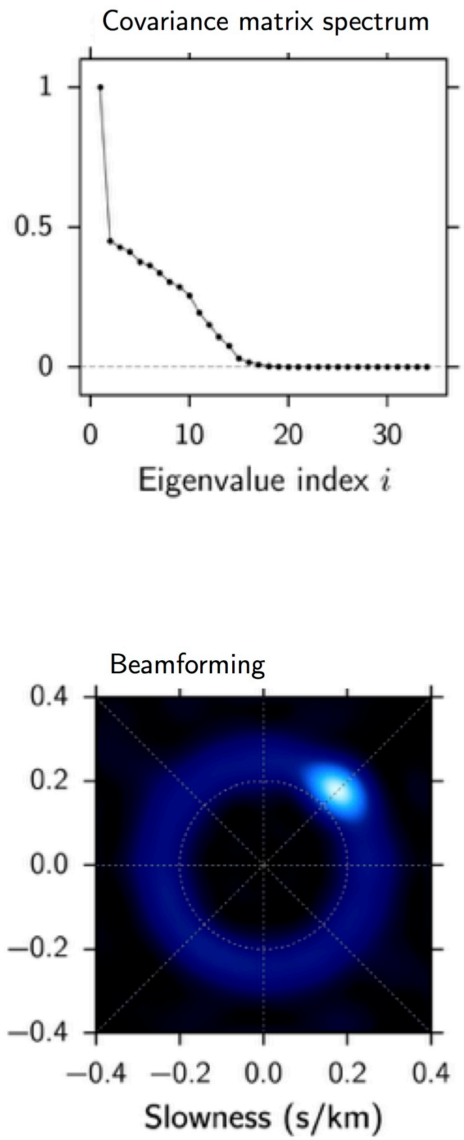


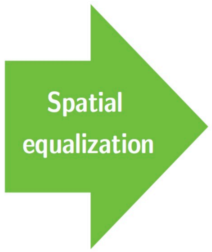
Spatial equalization (synthetic case)

Isotropic noise
(what we want)

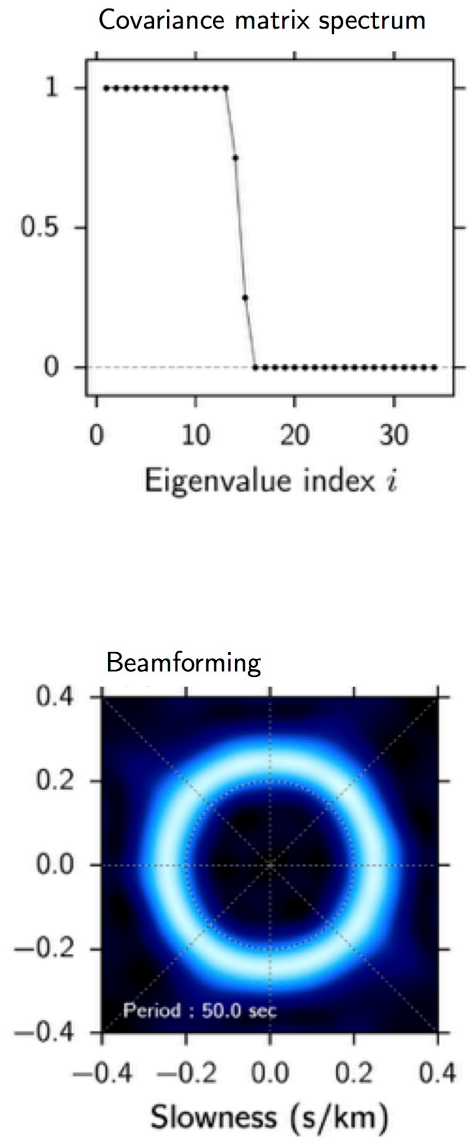


Isotropic noise + source
(what we actually have)

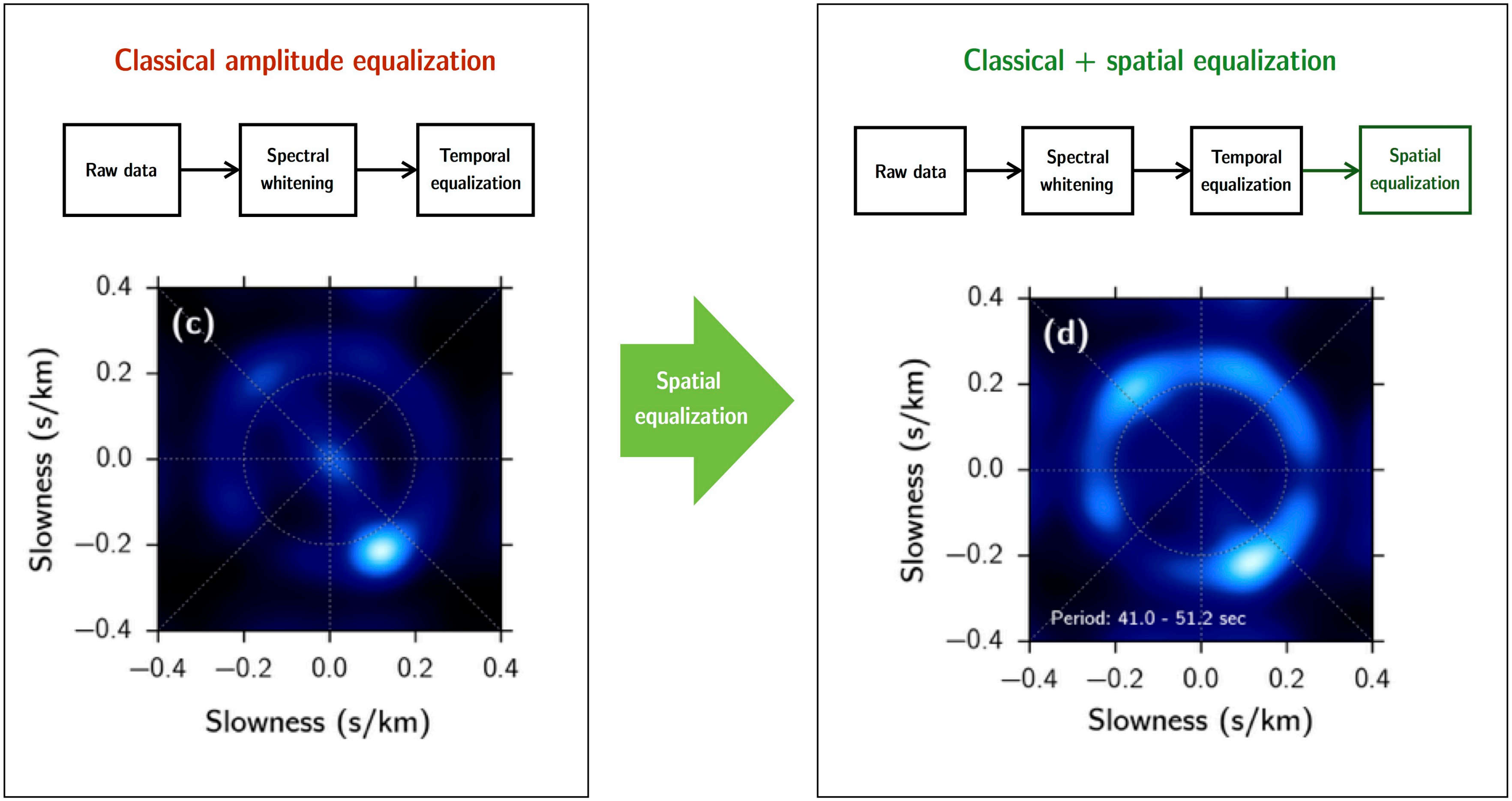




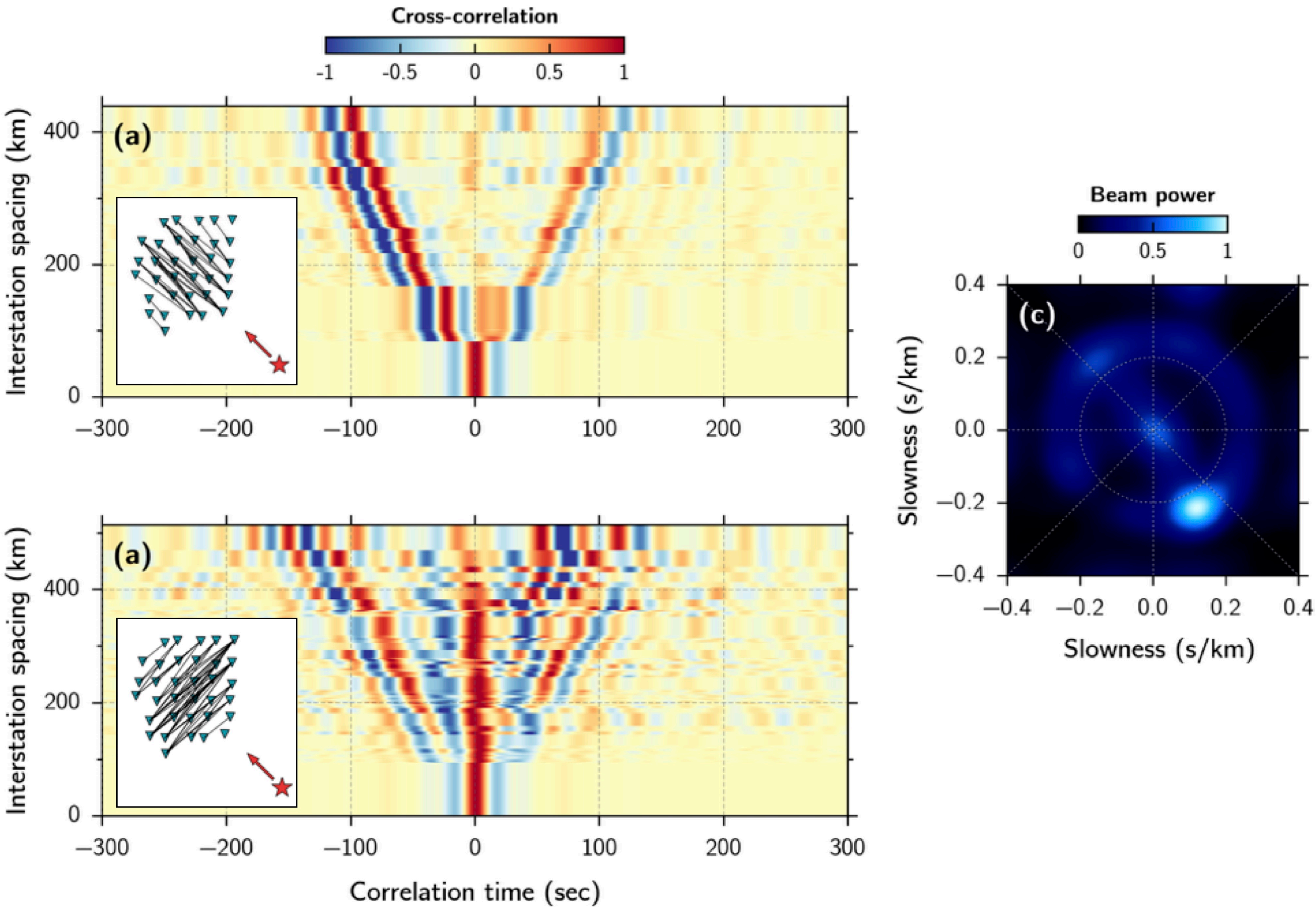
Equalized source



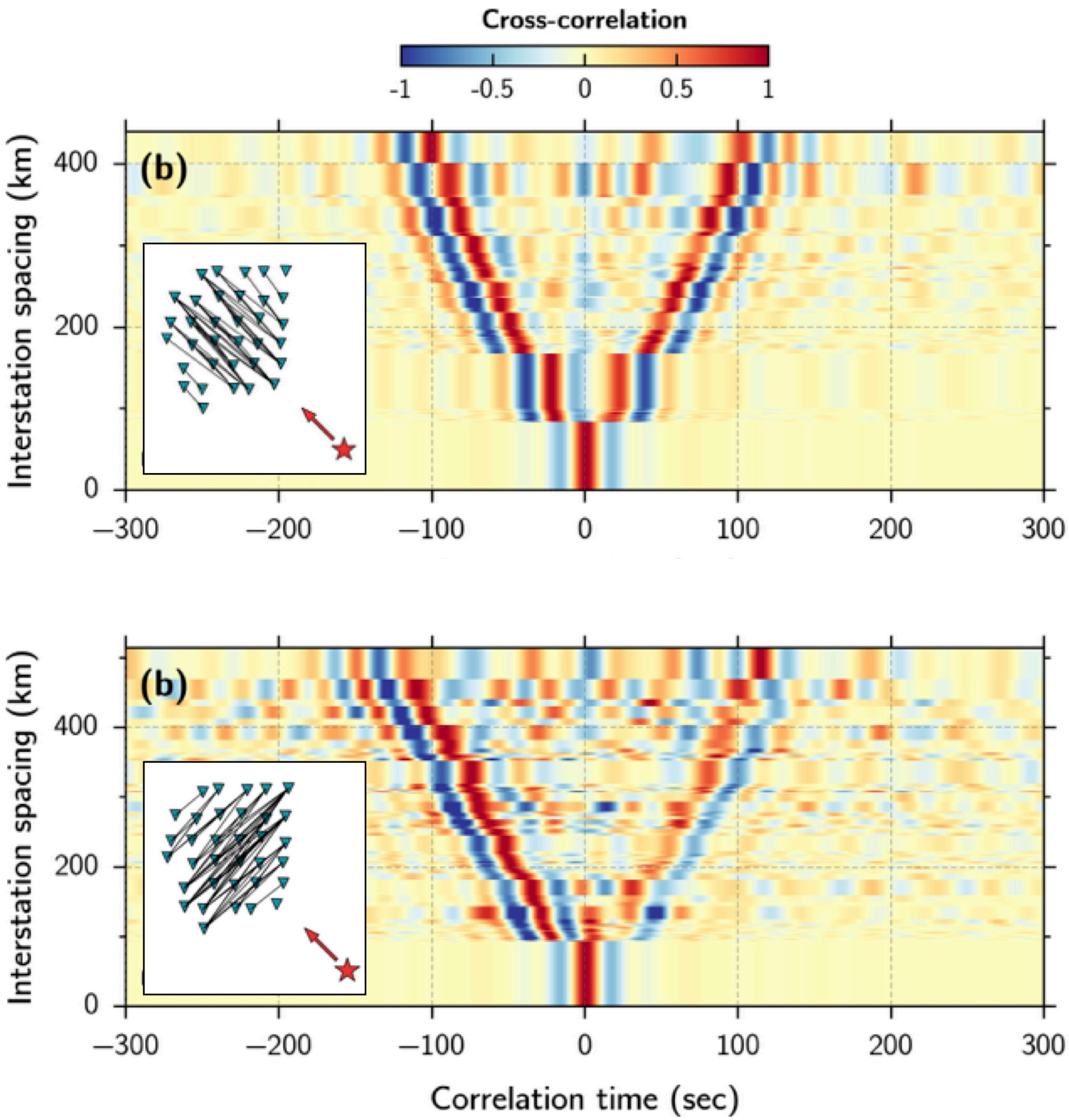
Spatial equalization of the M.8.8 Maule earthquake

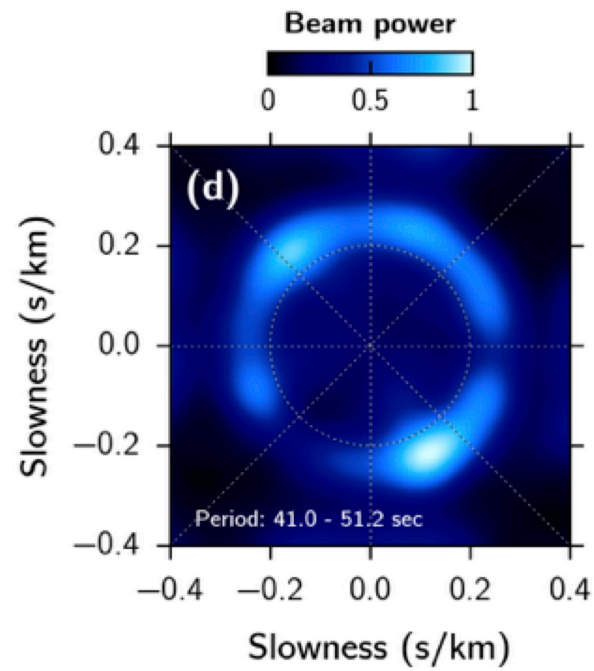


Cross-correlation around the Maule earthquake with temporal and spectral equalizations



Cross-correlation around the Maule earthquake with temporal, spectral and spatial equalizations





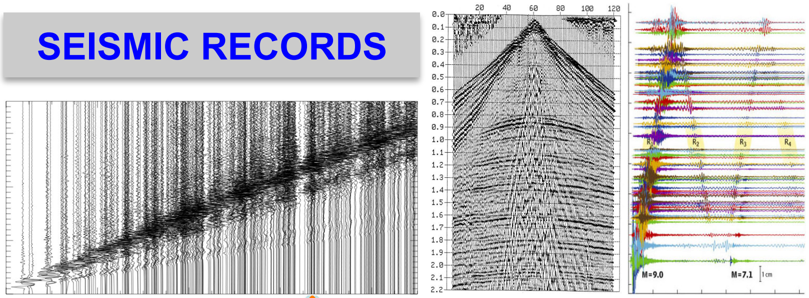
Conclusions

- Seismic wavefield in the Earth is not fully random and diffuse
- Seismic records must be pre-processed before cross-correlation to obtain a reasonable approximation of Green functions
- Time and spectral normalization of records at individual stations homogenizes the wavefield only partially
- Array-based methods can be used to further improve the pre-processing

seismic networks : large scale antennas

Data intensive seismology

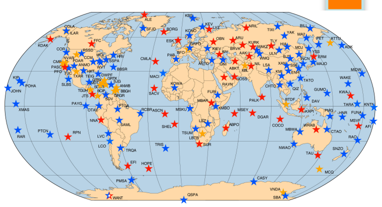
SEISMIC RECORDS

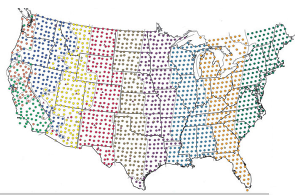


DATA ARCHIVING

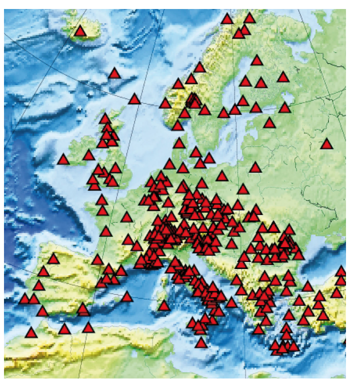
DATA CENTERS

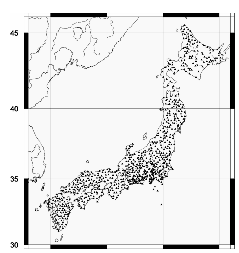
DATA COLLECTION





SEISMIC NETWORKS





DATA INTEGRATION

HPC resources

infrastructure for efficient exploration of rapidly growing volumes of seismological observations

DATA ANALYSIS, MODELING, AND INVERSION

Earth's interior and dynamic

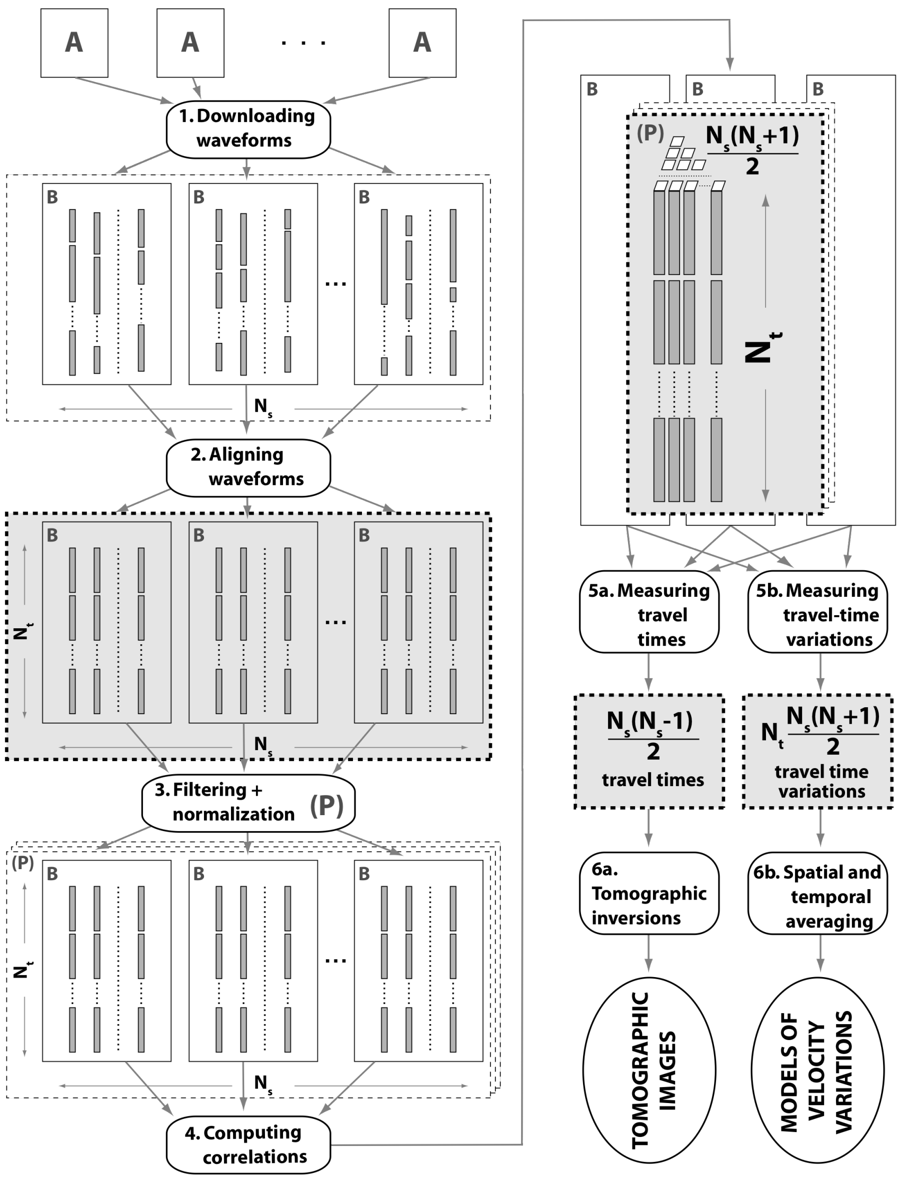
Earthquakes, volcanoes, natural hazards

Interaction of solid Earth with Ocean and Atmosphere

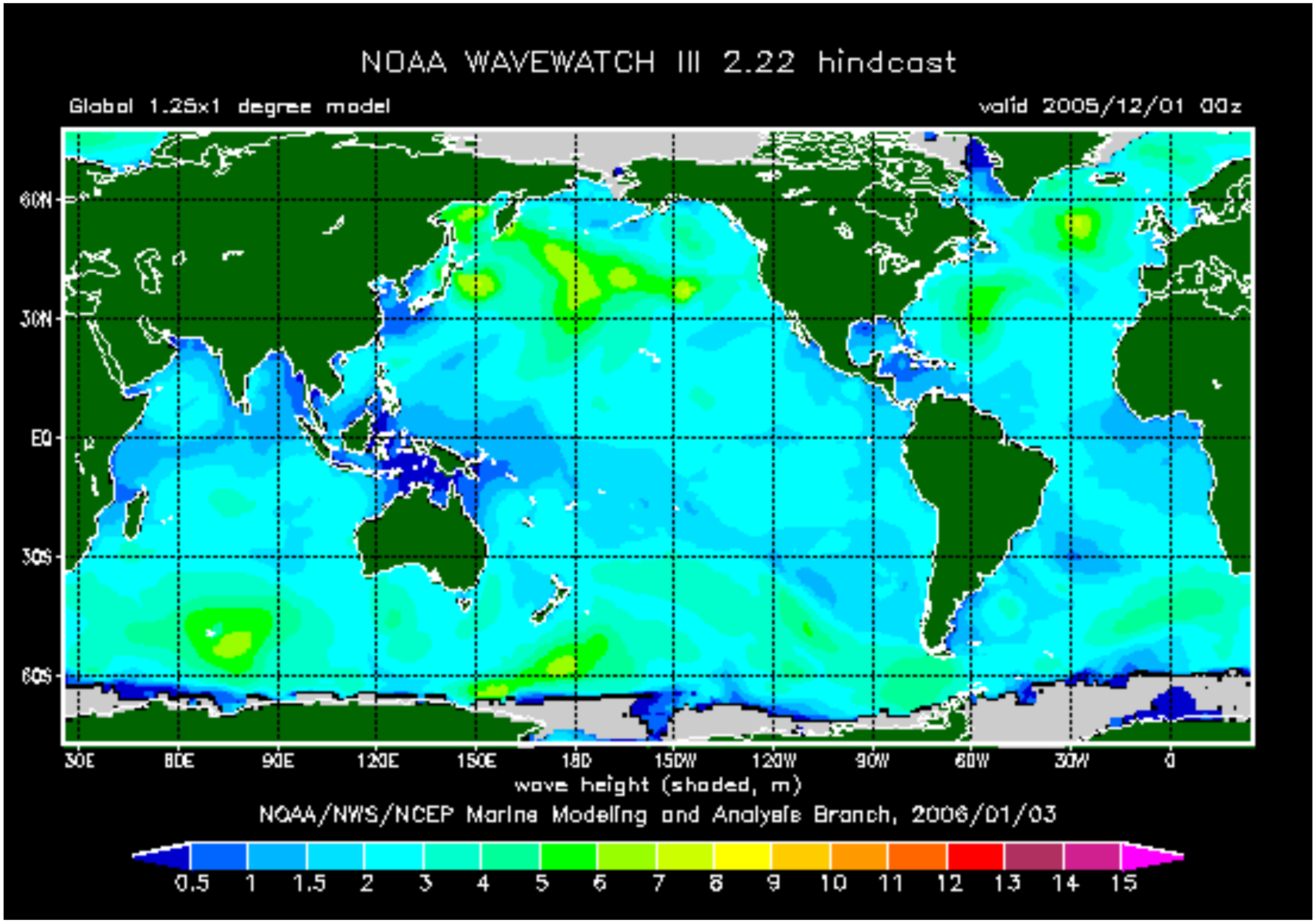
END

Analysis of continuous seismic data

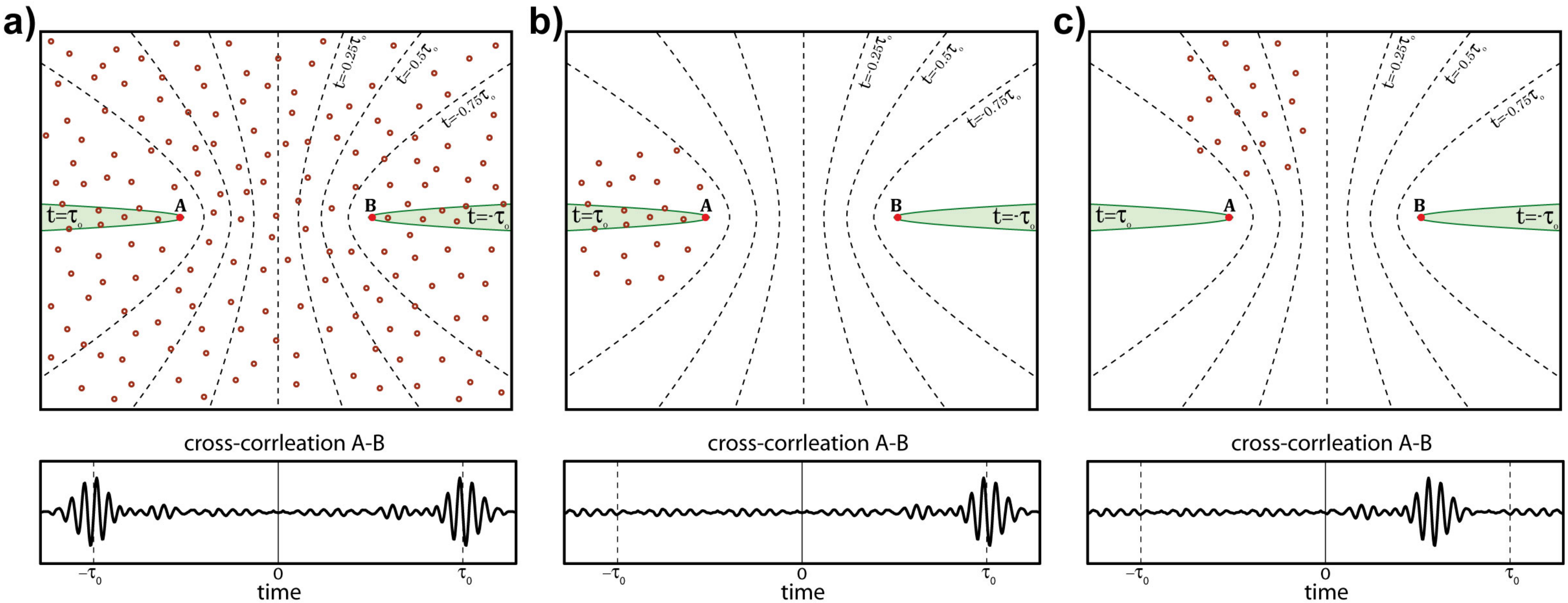
(A) seismological datacenters (B) data processing platforms



- **Large volumes of input raw data**
- **Analysis: sequences of very large number of simple operations: (digital filters, Fourier transforms, dot-products, …)**
- **Huge number of output files (N input-> N(N-1)/2 output)**
- **Repeating analyses**

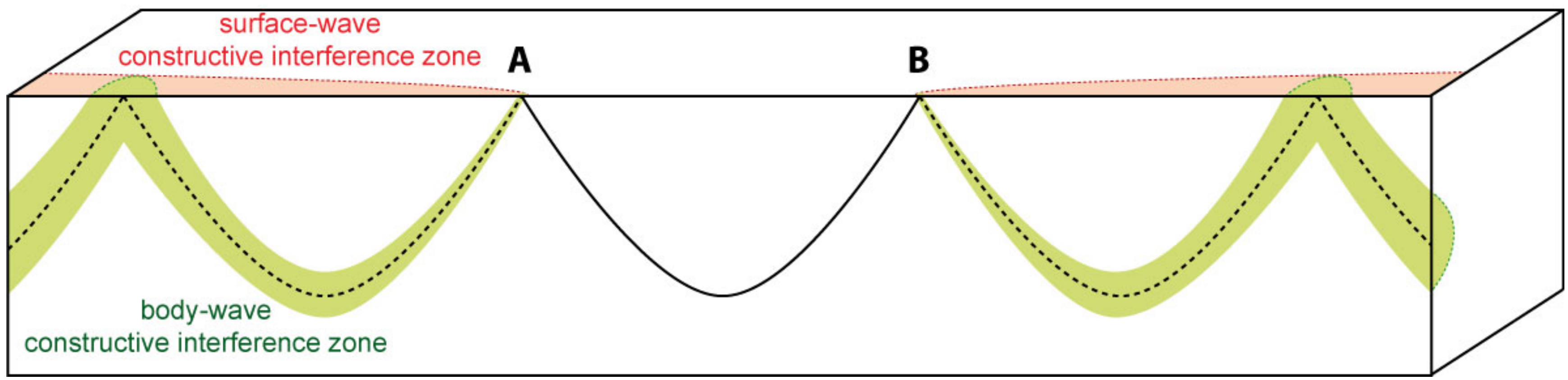


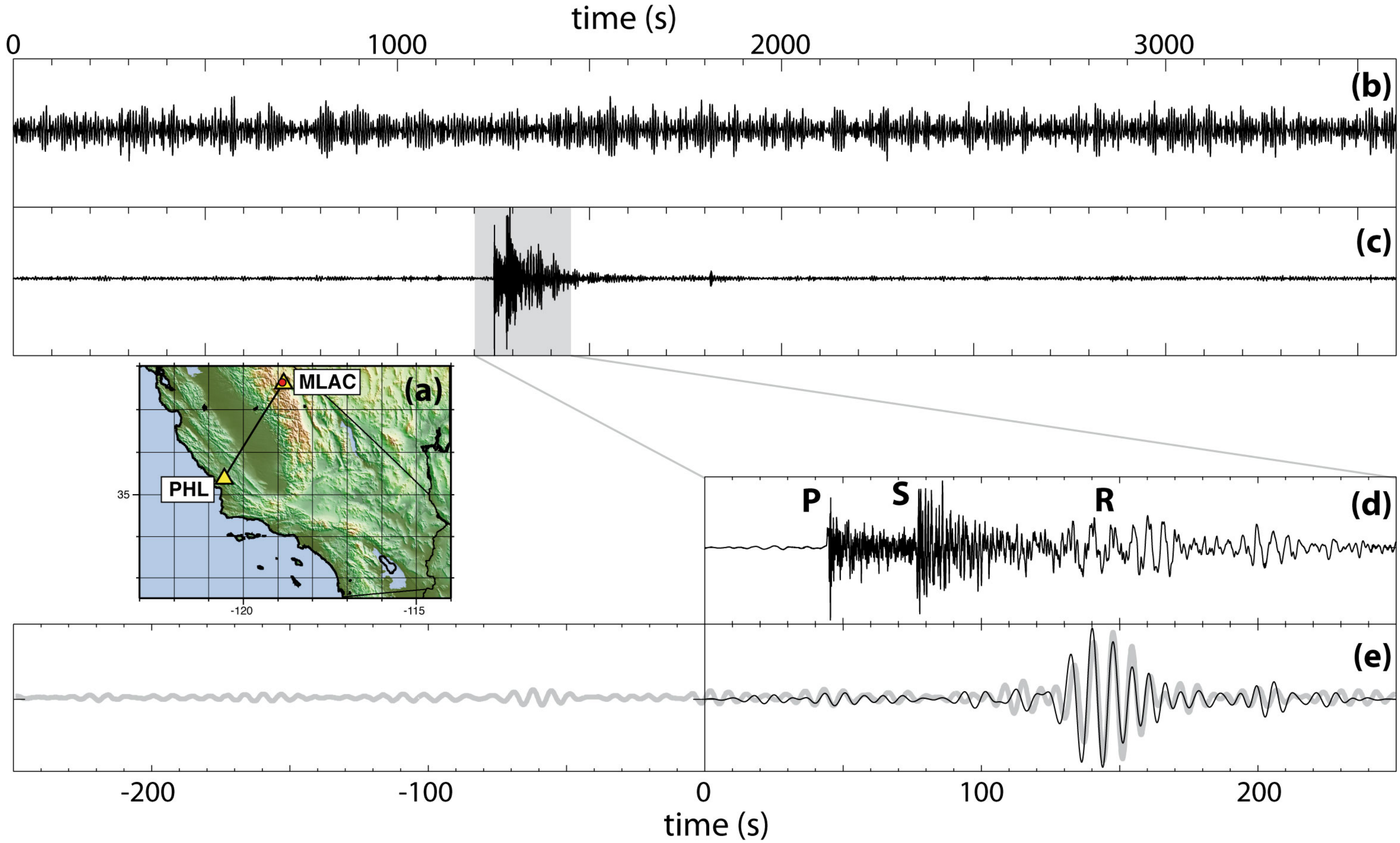
NOAA WAVEWATCH III 2.22 hindcast
Global 1.25x1 degree model
valid 2005/12/01 00z
60N
30N
EQ
30S
60S
30E
60E
90E
120E
150E
180
150W
120W
90W
60W
30W
0
wave height (shaded, m)
NOAA/NWS/NCEP Marine Modeling and Analysis Branch, 2006/01/03
0.5
1
1.5
2
3
4
5
6
7
8
9
10
11
12
13
14
15



$$(f \star g)(\tau) \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} f^*(t)\ g(t+\tau)\,dt,$$

Constant phase: $t_A - t_B = \text{const}$

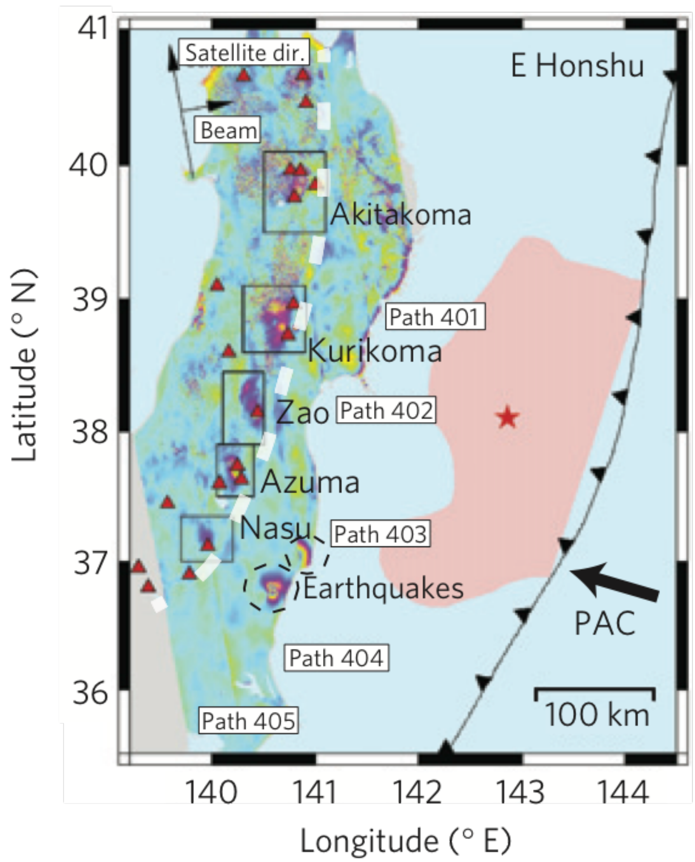

surface-wave
constructive interference zone
A
B
body-wave
constructive interference zone

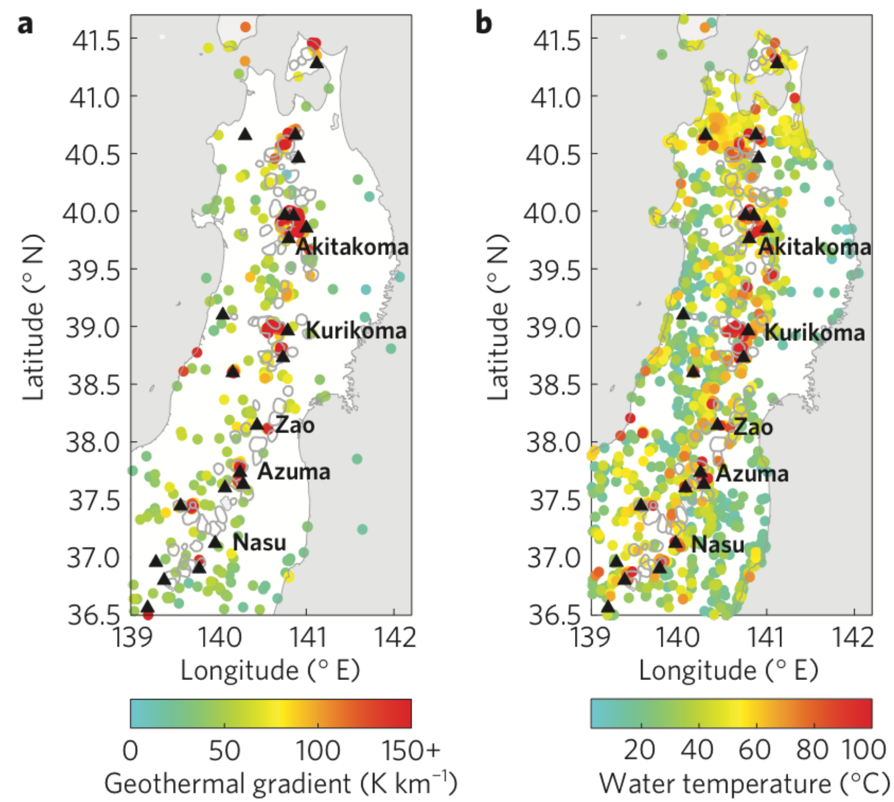


time (s)
0
1000
2000
3000
(b)
(c)
MLAC
(a)
PHL
35
-120
-115
P
S
R
(d)
(e)
-200
-100
0
100
200
time (s)

Volcanic subsidence triggered by the 2011 Tohoku earthquake in Japan

Subsidence observed with the satellite interferometry

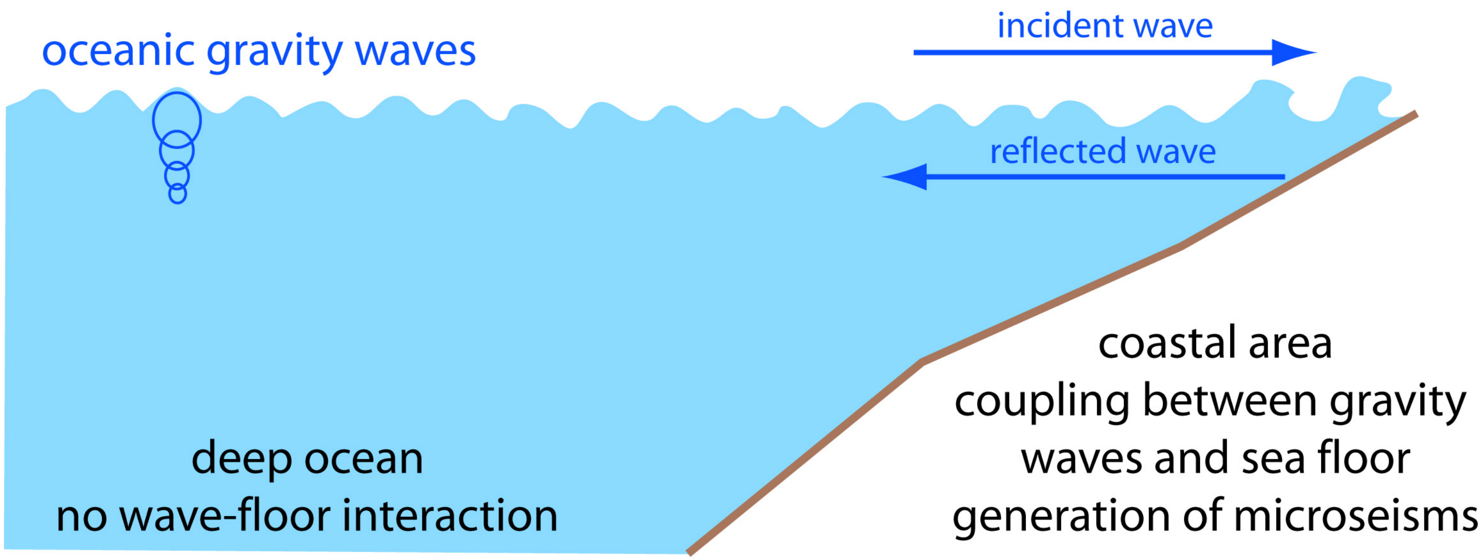




Takada and Fukushima, 2013

Generation of microseisms

theory from Longuet-Higgins (1950)

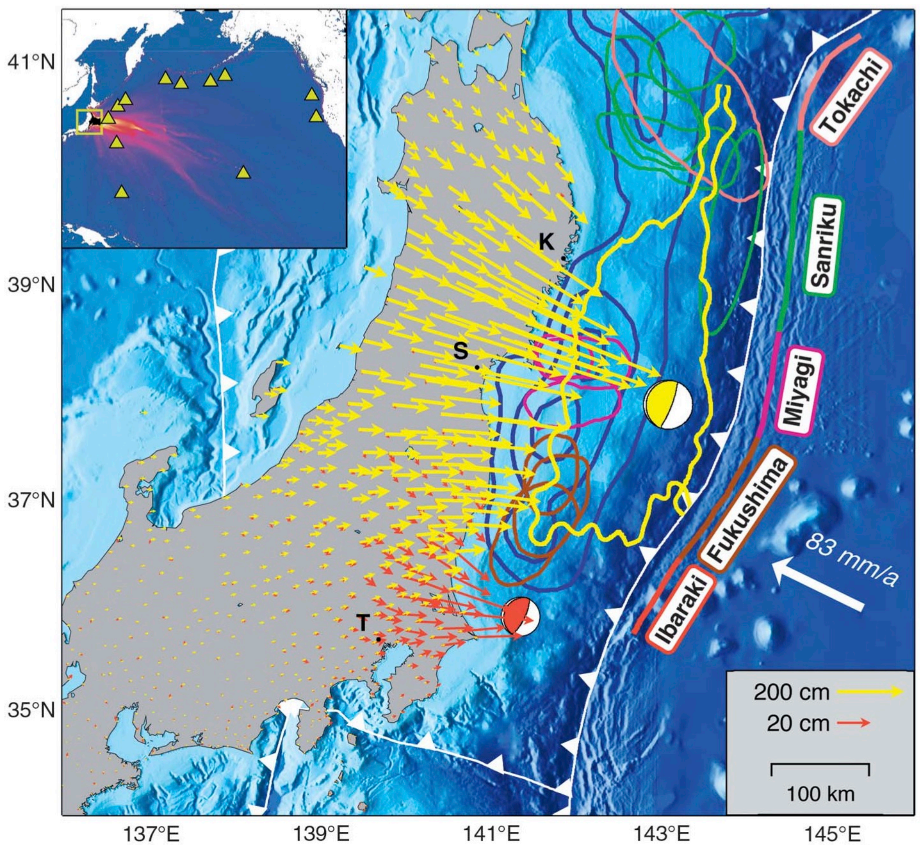


primary microseism is excited at frequencies corresponding to the spectrum of incoming oceanic gravity waves (periods of **10-20 s**)

secondary microseism is exited at doubled frequencies due to the nonlinear interaction between incident and reflected waves (periods of **5-10 s**)

Crustal velocity changes during the 2011 Tohoku earthquake in Japan

Tohoku earthquake



Hi-net seismic network