Analyse intensive de données sismiques

Nikolai Shapiro (IPGP)

with contributions from:

Florent Brenguier (IsTerre, Grenoble)

Michel Campillo (IsTerre, Grenoble)

William Frank (IPGP)

Matthieu Landès (IPGP)

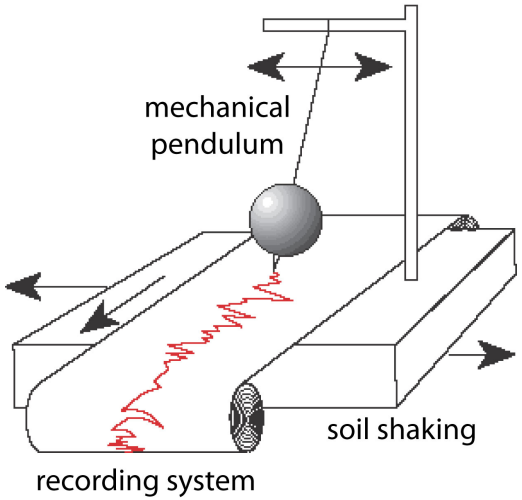
Aurelien Mordret (IPGP)

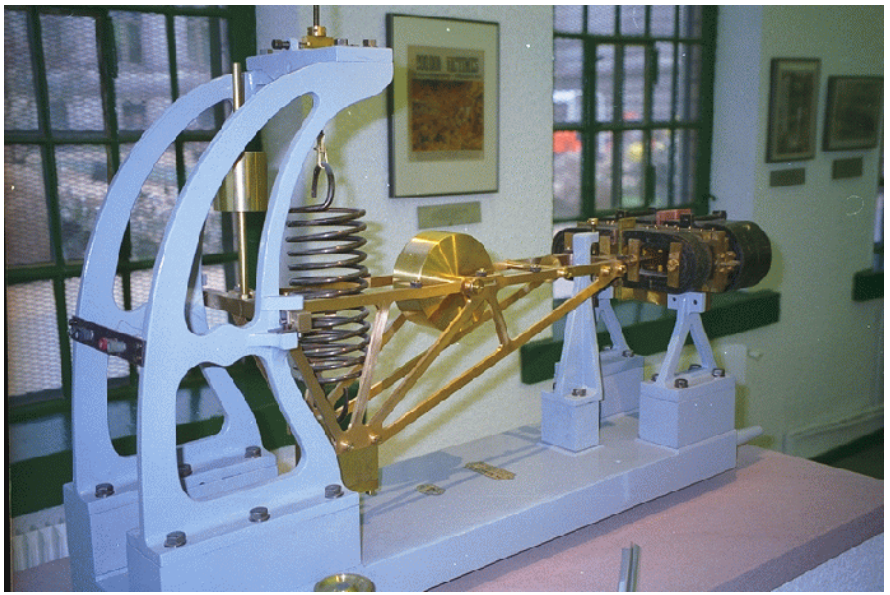
Diane Rivet (Geoazur, Nice)

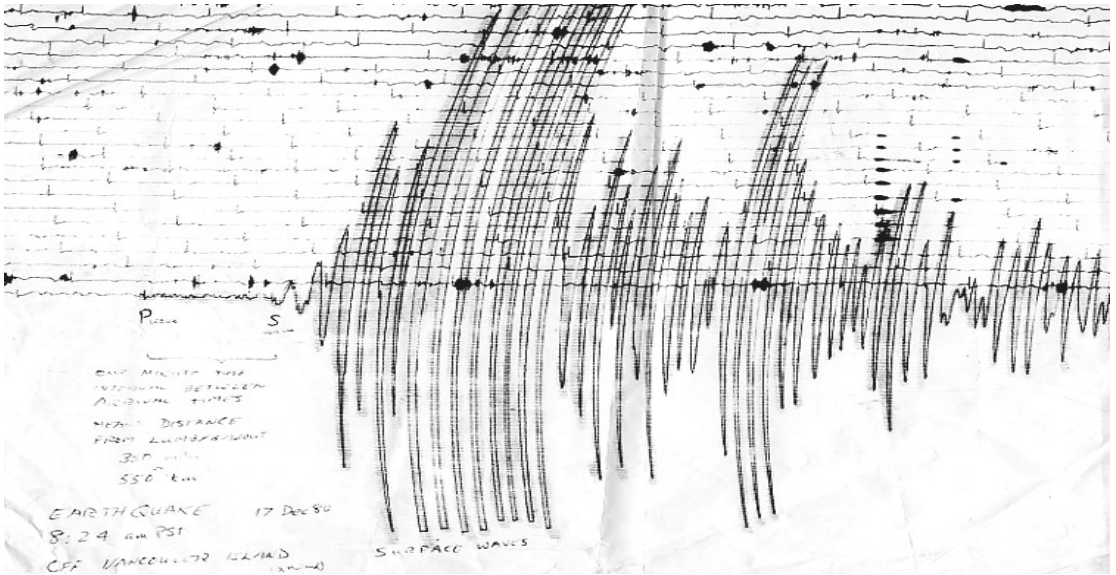
Alexey Romanenko (Novosibirsk State University)

Geneviève Moguilny (IPGP)

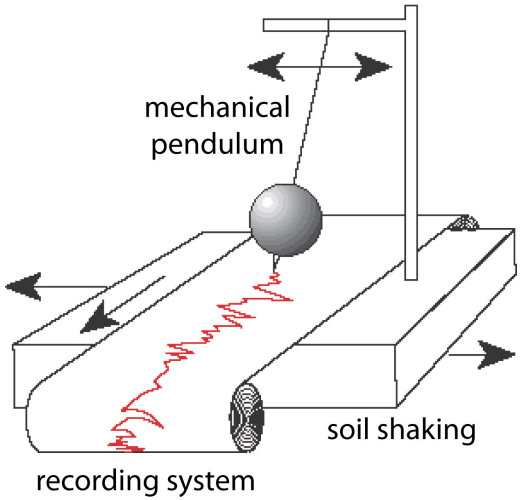
Seismic data: records of the motion of the Earth's surface

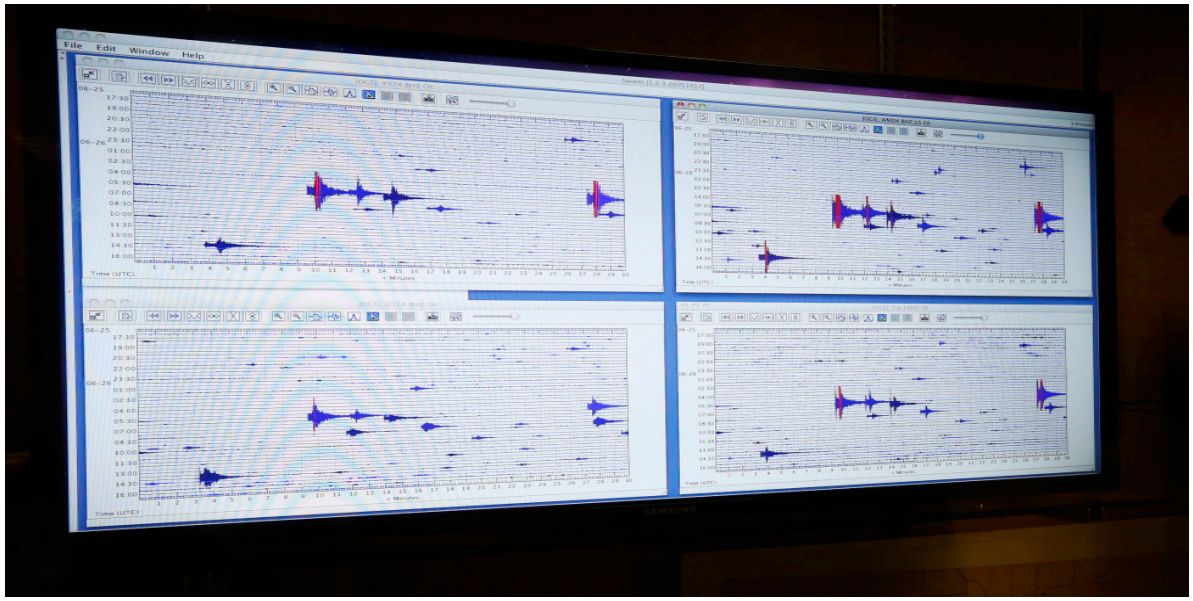






Seismic data: records of the motion of the Earth's surface





Seismology

- Monitoring and mitigation of natural hazards (earthquake, tsunamis, and volcanoes)
- Studies of the Earth's interior and dynamics

Universities and government funded research agencies

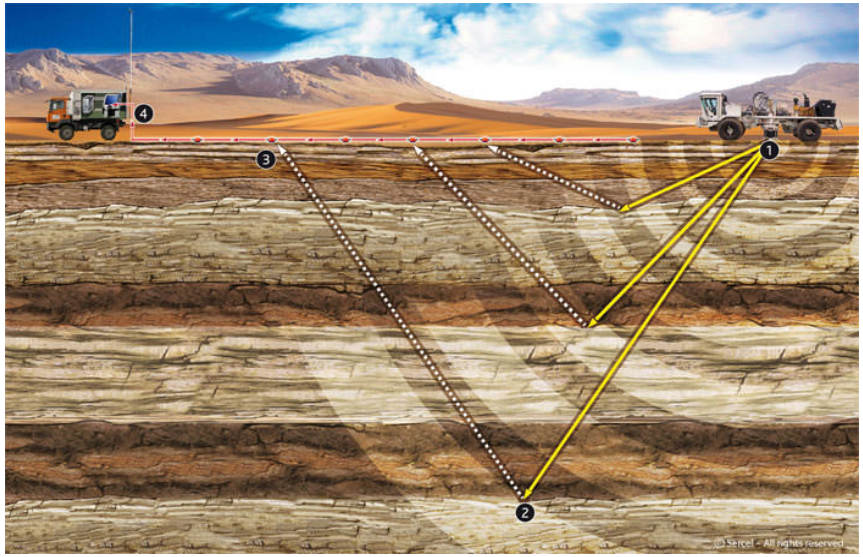
Exploration geophysics

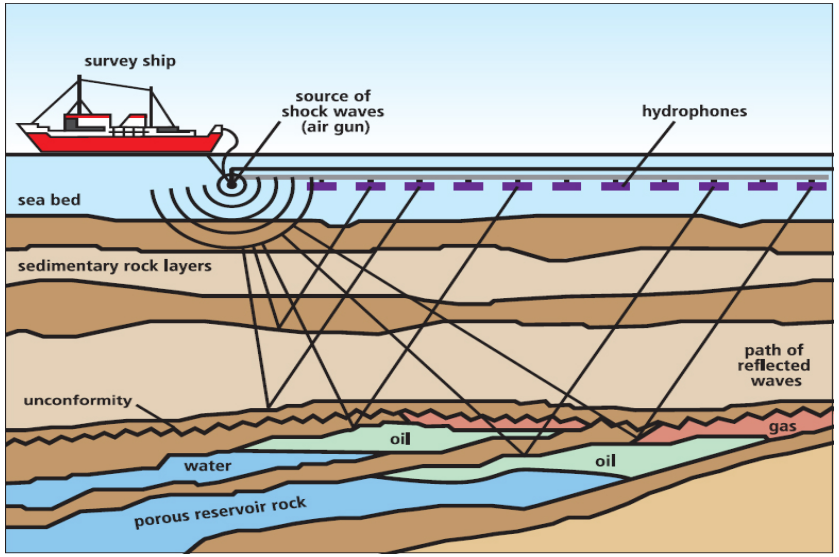
- Exploration of natural resources (mainly oil and gas)

Oil and gas and geophysics service companies

Exploration geophysics

Methods based on active sources (explosions, vibroseis, air guns)

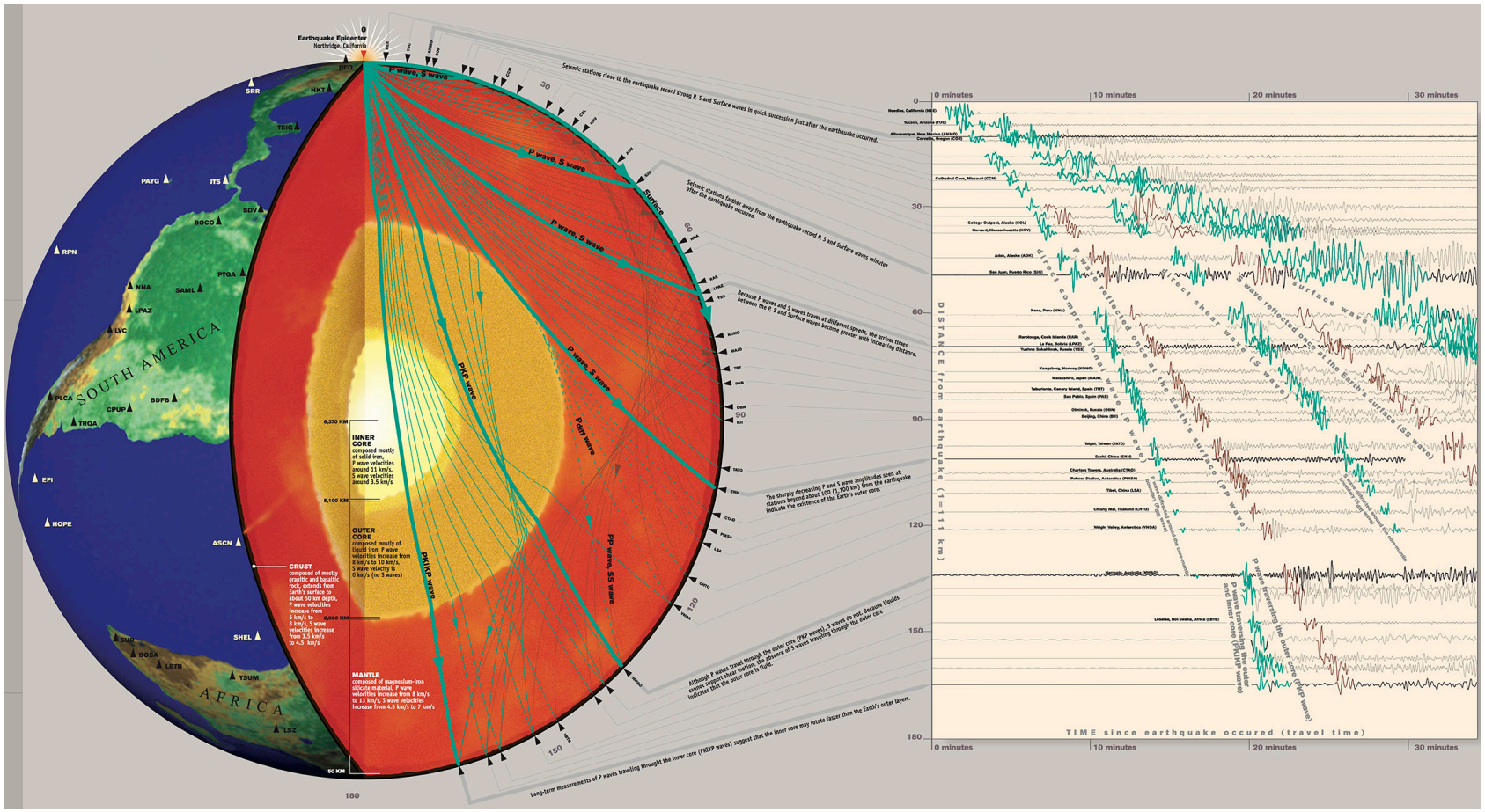




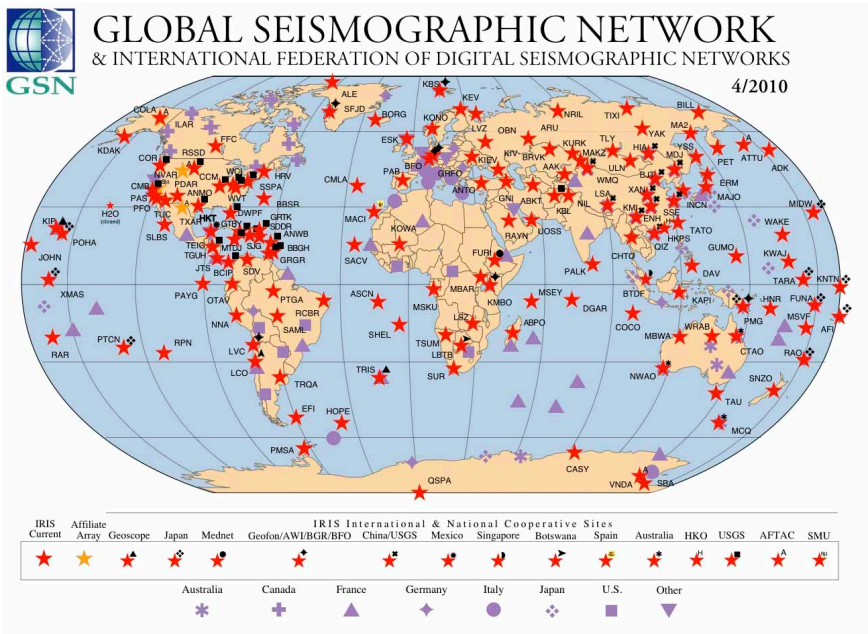
- Experiments with many thousands of receivers
- 4-component records
- ~250 samples per seconds
- Raw datasets up to many Pb

Seismology

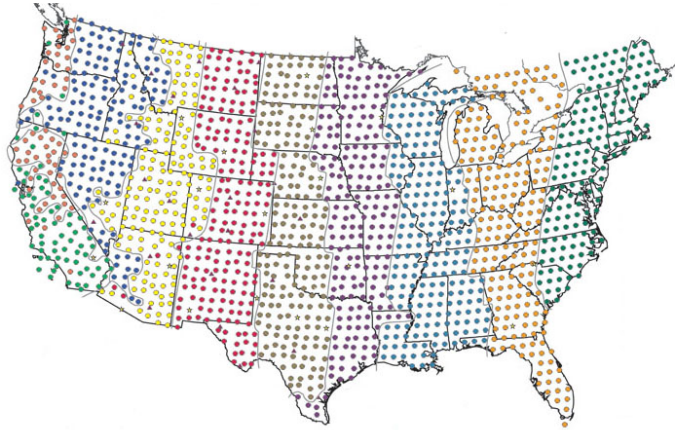
“Passive” methods based on natural seismic sources (earthquakes)

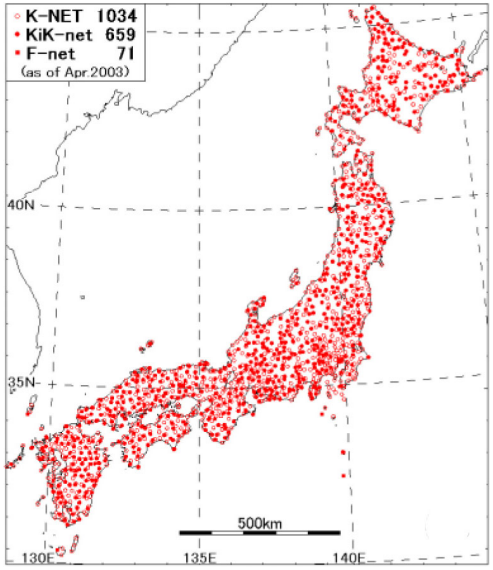


Seismology



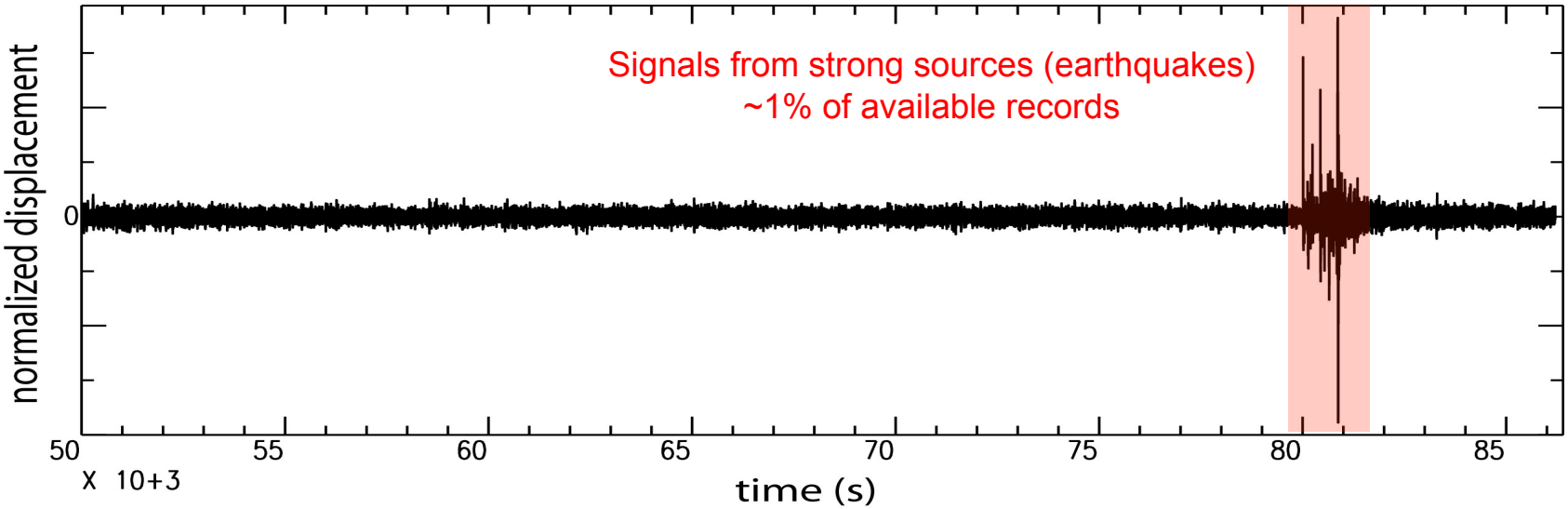
dense regional networks



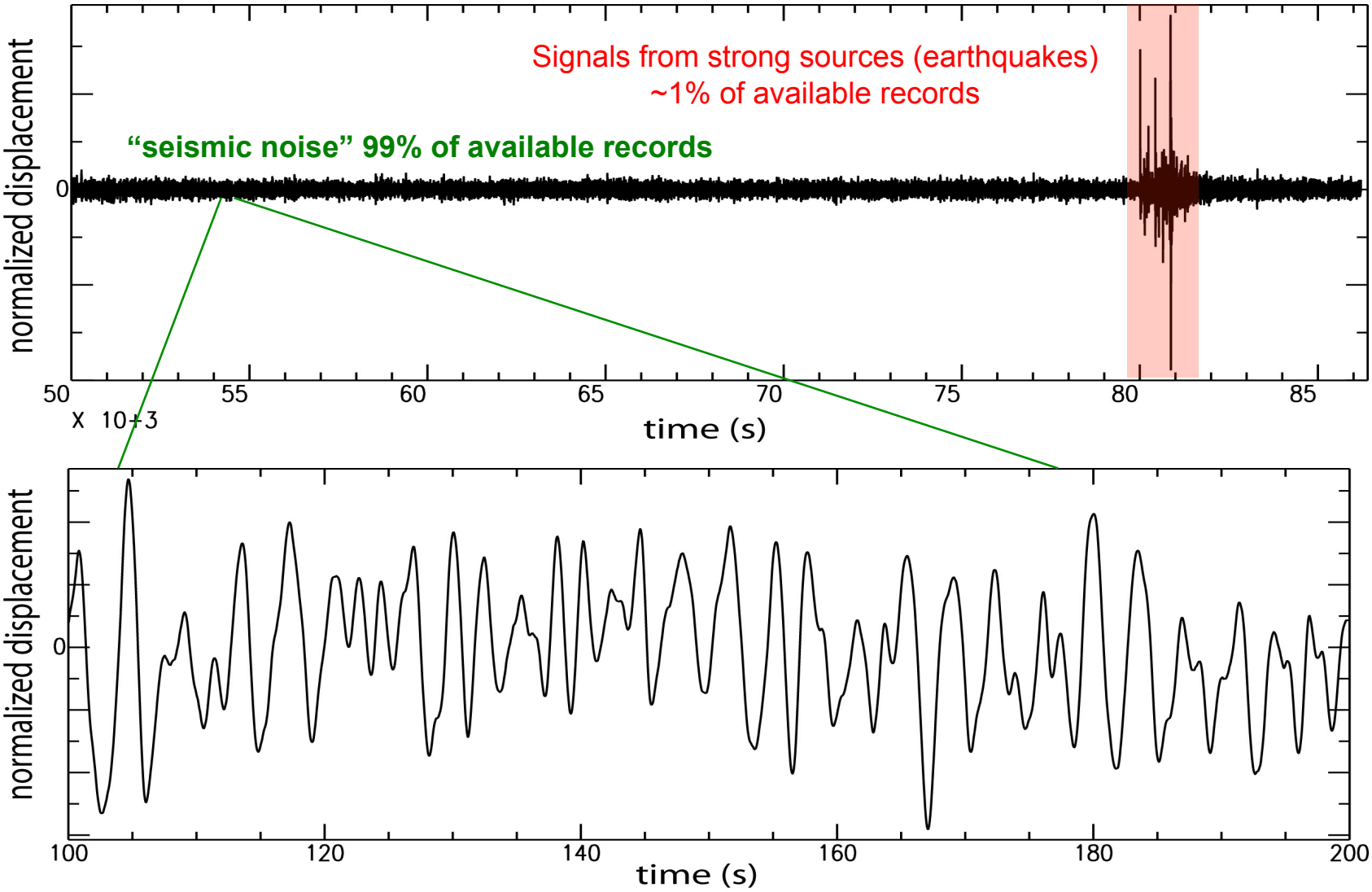


- **thousands of digital seismographs recording continuously**
- **3-component record**
- **~50 samples per second**
- **raw datasets of ~10 or ~100 Tb**

New paradigm in seismology: using full continuous records

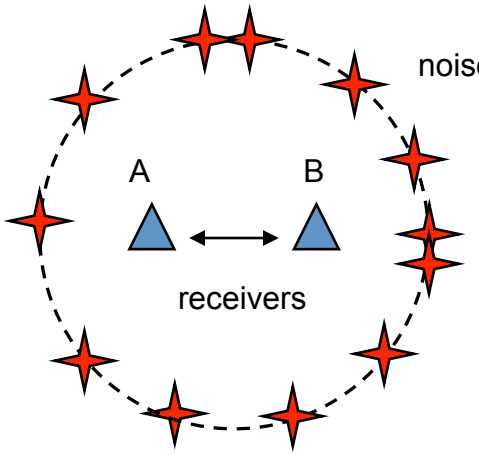


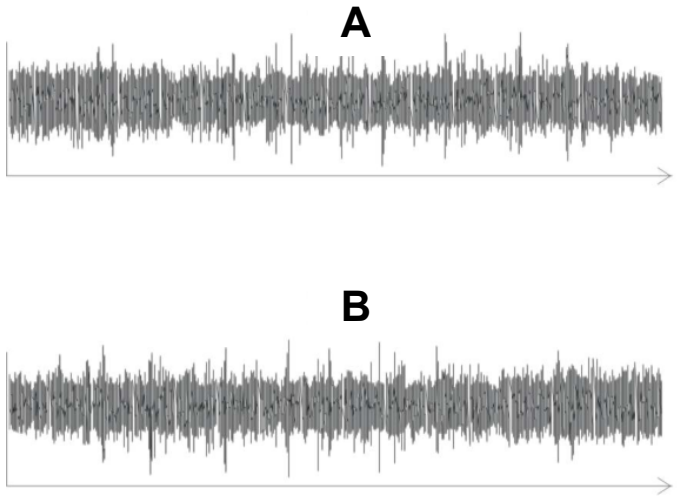
New paradigm in seismology: using full continuous records



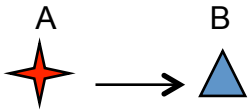
Using continuous seismic records for imaging and monitoring

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



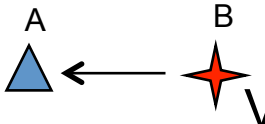


Virtual source



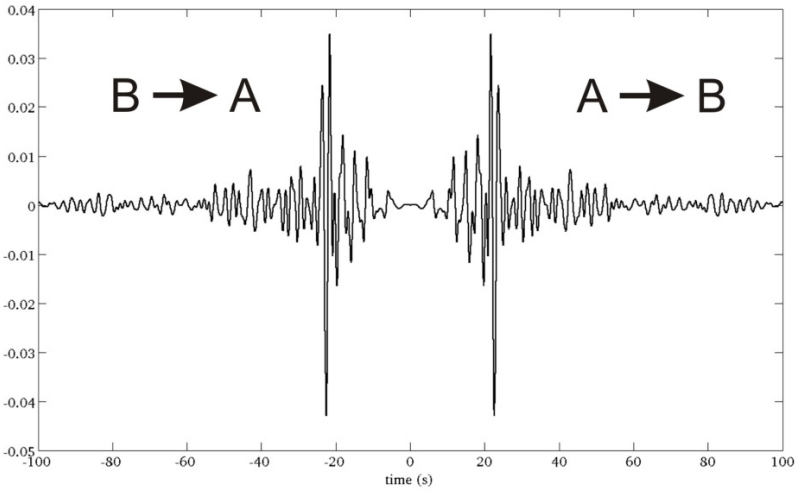
receiver

receiver

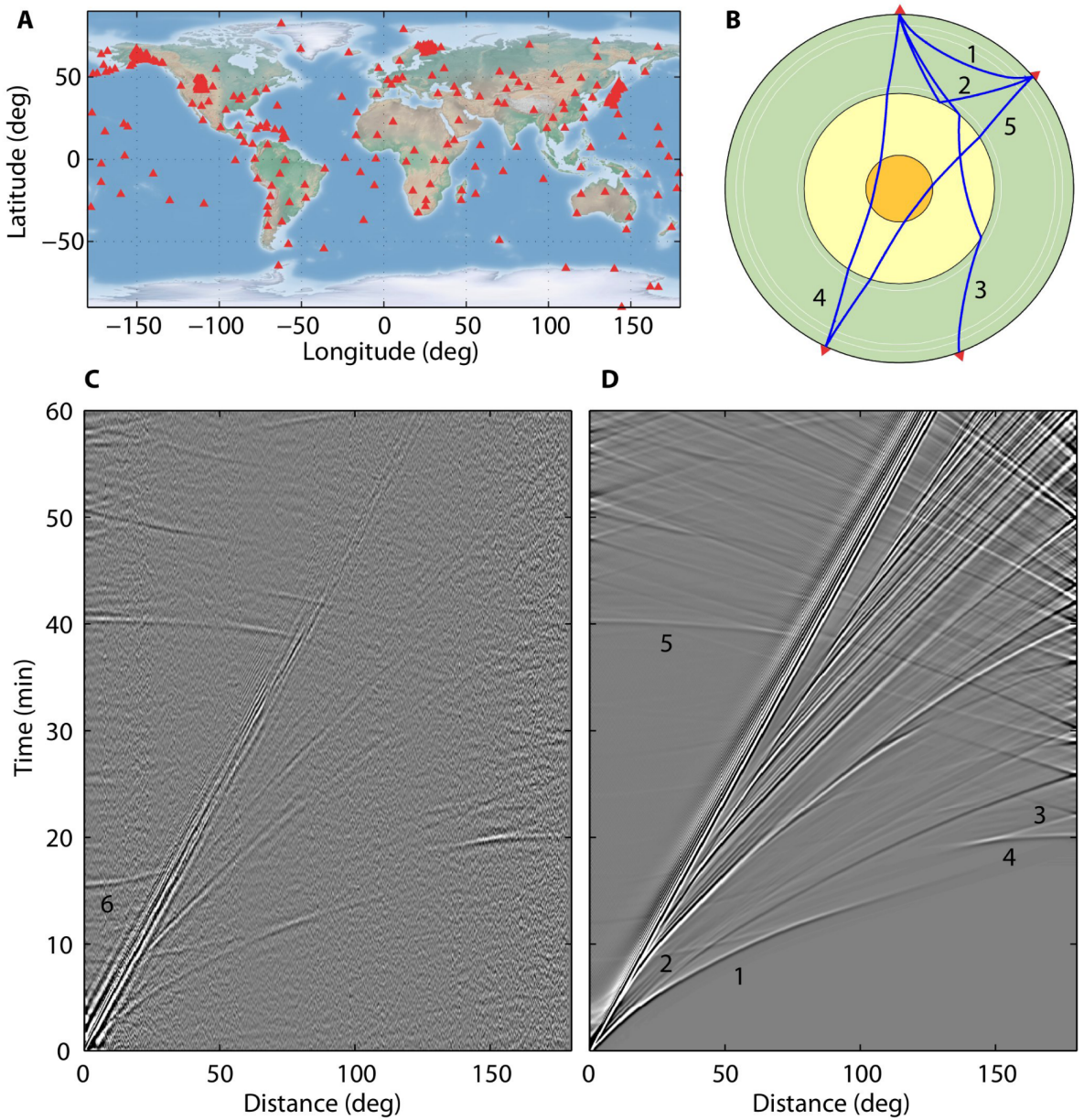


Virtual source

Correlation function = Green's function



Noise-based “virtual seismograms” from the global seismic network

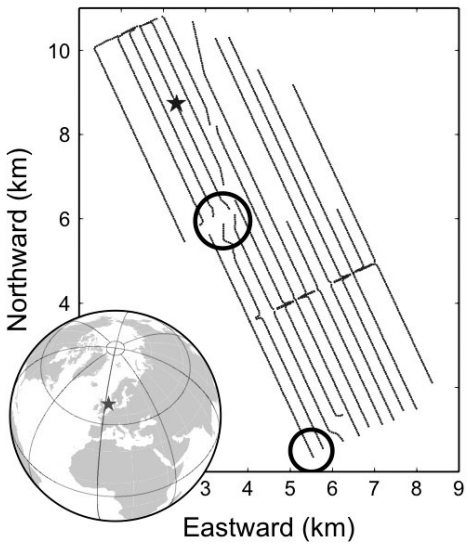


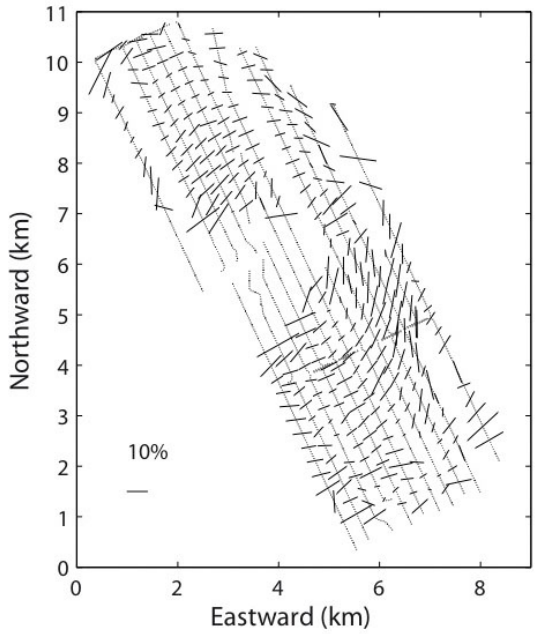
Boué et al., 2013

Noise based imaging with the industrial data

Valhall
Network
2400 receivers

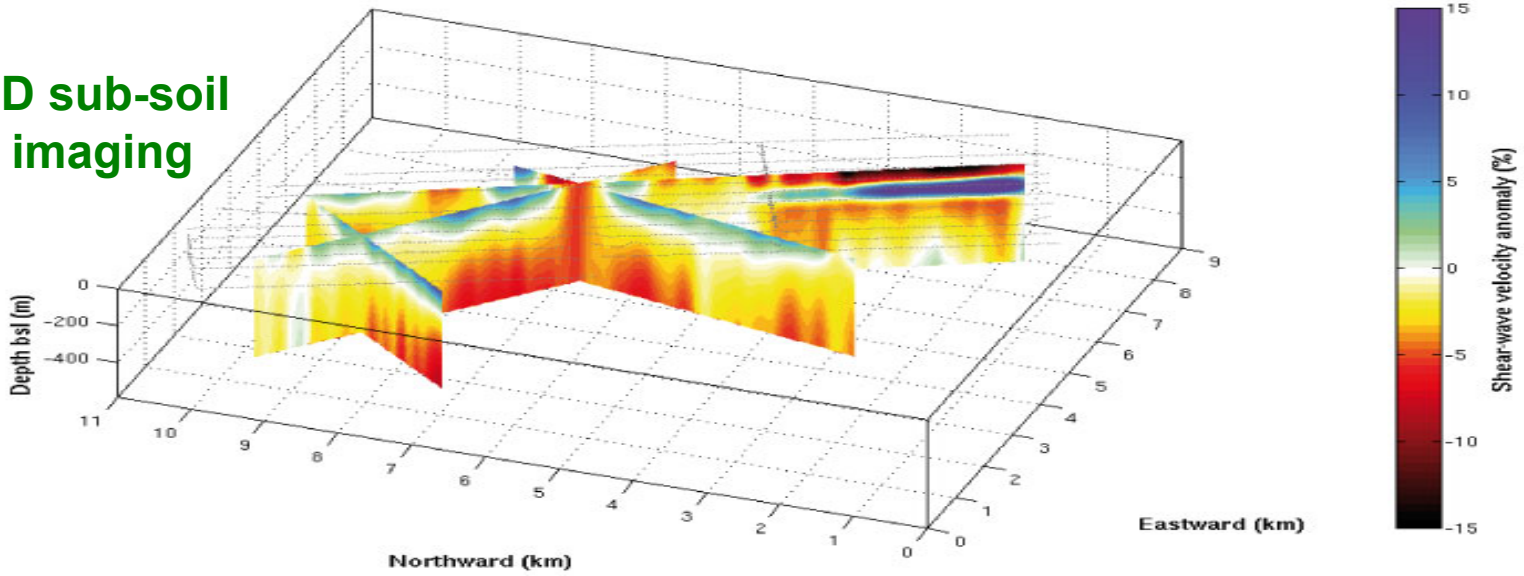
6 hours of
passive records





Seismic anisotropy

3D sub-soil imaging



Mordret et al., 2013, 2014, 2015

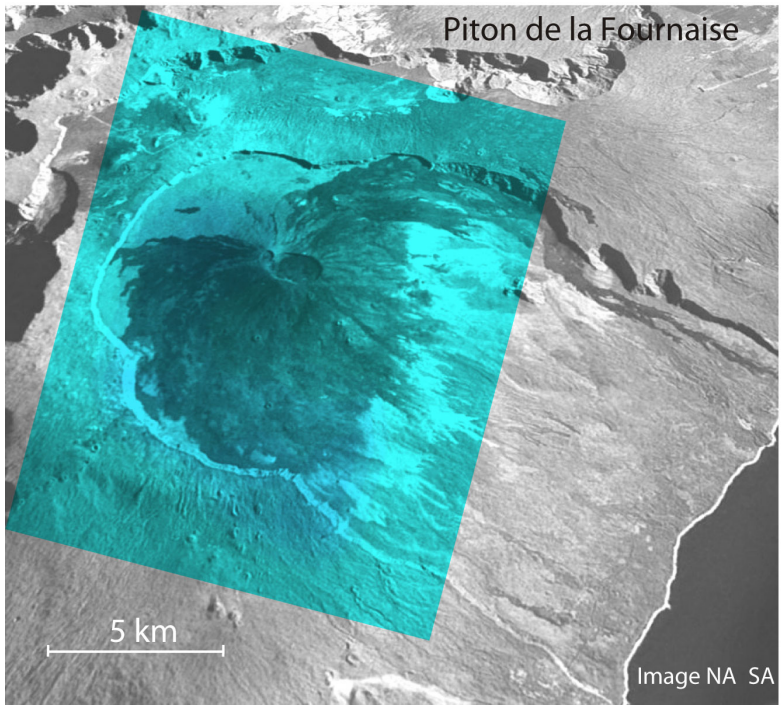
Noise-based seismic monitoring

- computing cross-correlation of seismic noise between two stations from long enough records is equivalent to an experiment when a source is acting at location of one of stations and recorded at another

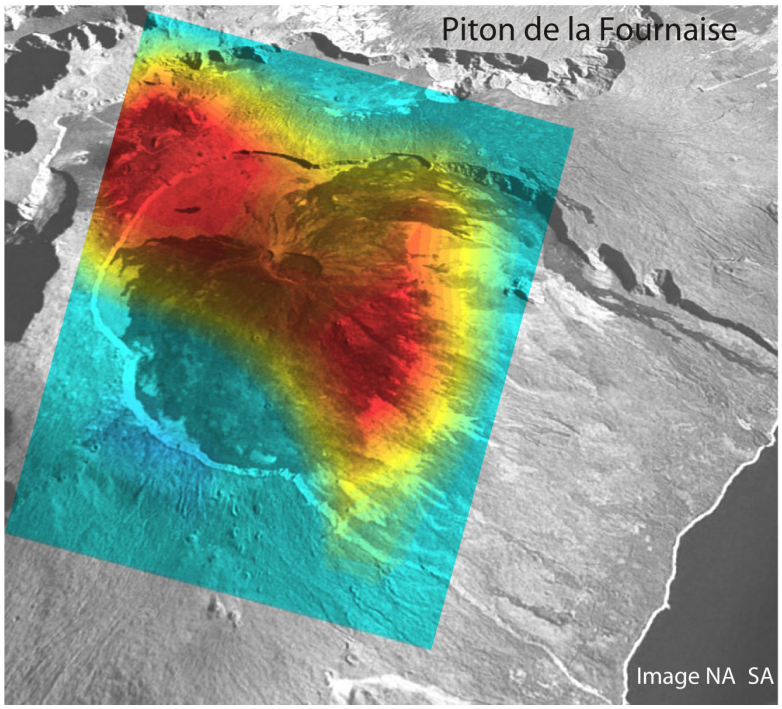
- repetitive computations of noise cross-correlations are equivalent to using repetitive seismic sources and can be used to detect changes in the medium

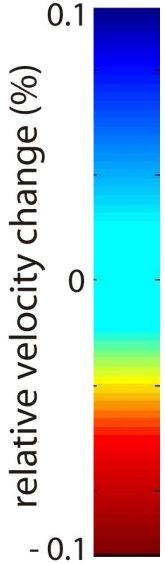
Monitoring of the Piton de la Fournaise volcano at La Réunion Island

9 days before eruption of June 2000



4 days before eruption of June 2000

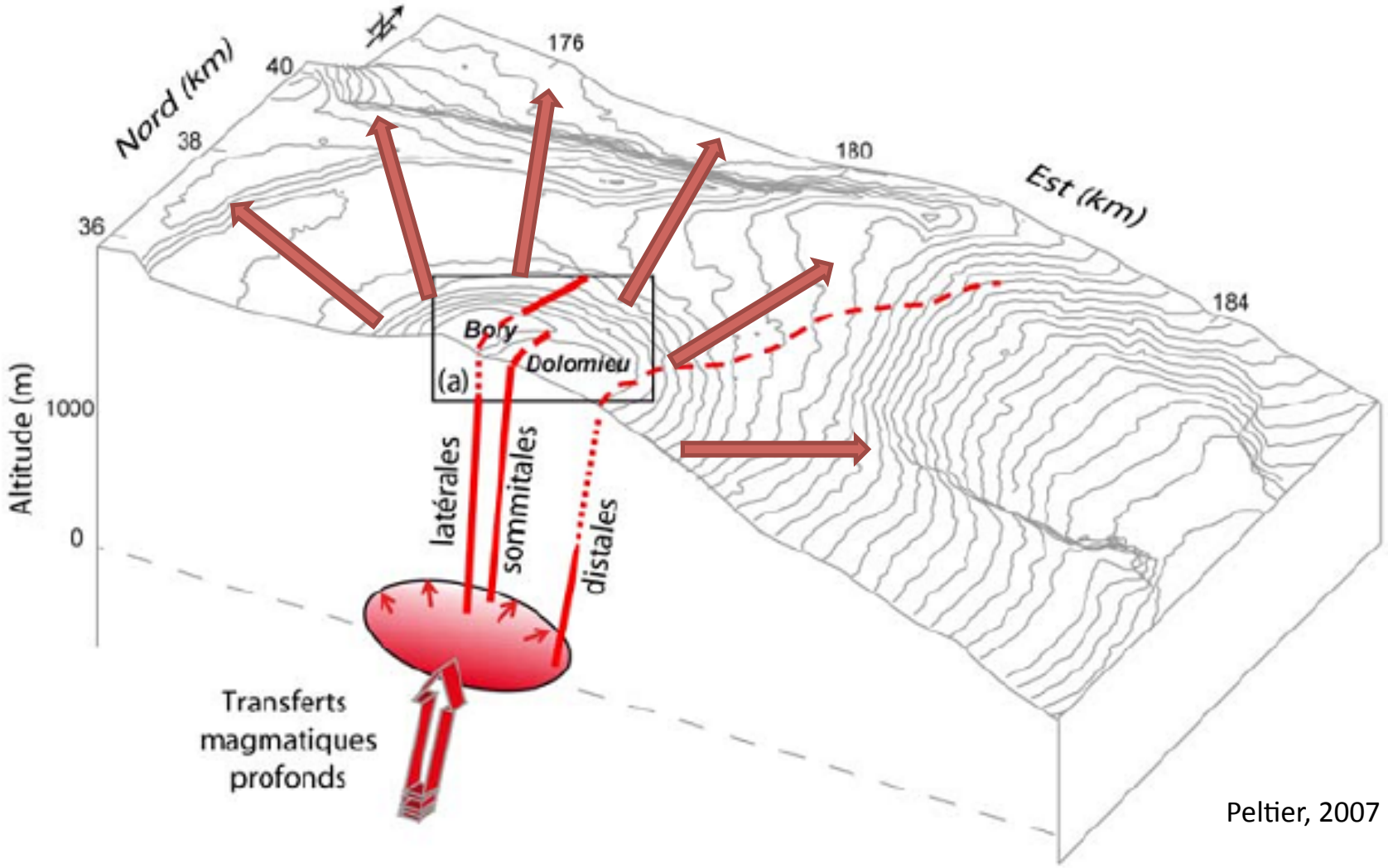




Detected velocity variations are localized in the vicinity of the main crater:
consistent with a shallow source of deformation

Brenguier et al., Nature Geoscience, 2008

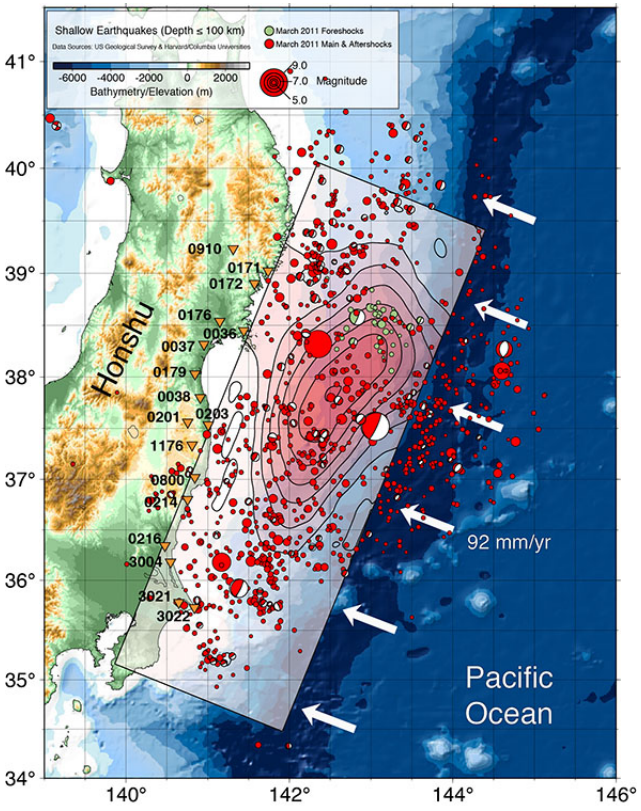
Stress build-up within the reservoir "dilates" the edifice and opens cracks



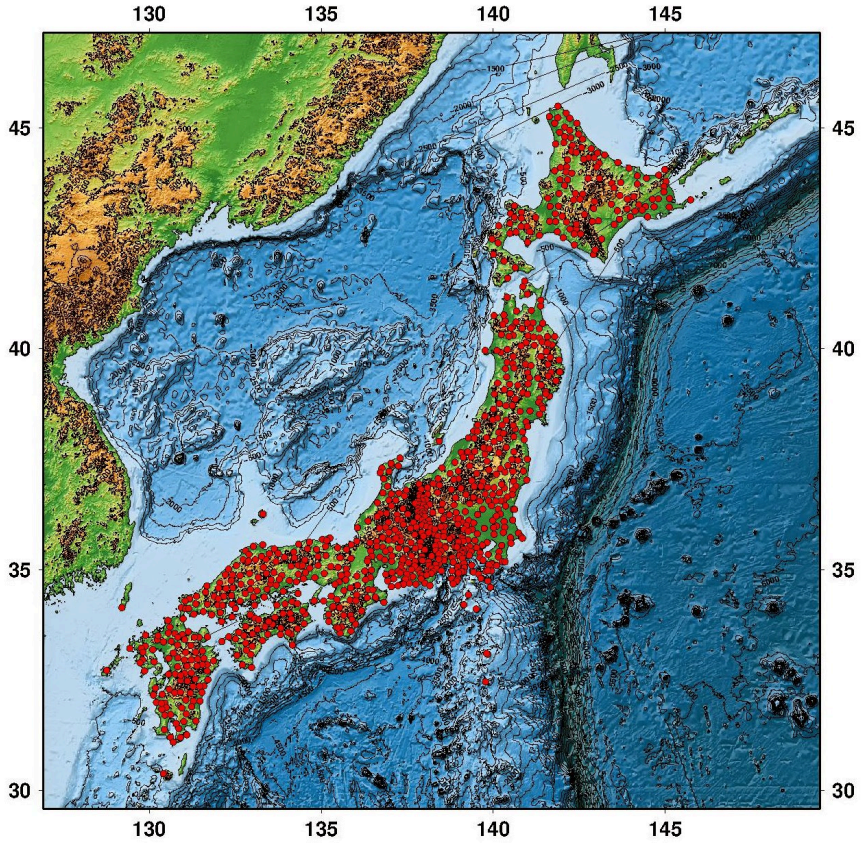
Peltier, 2007

Crustal velocity changes during the 2011 Tohoku earthquake in Japan

Tohoku earthquake



Hi-net seismic network

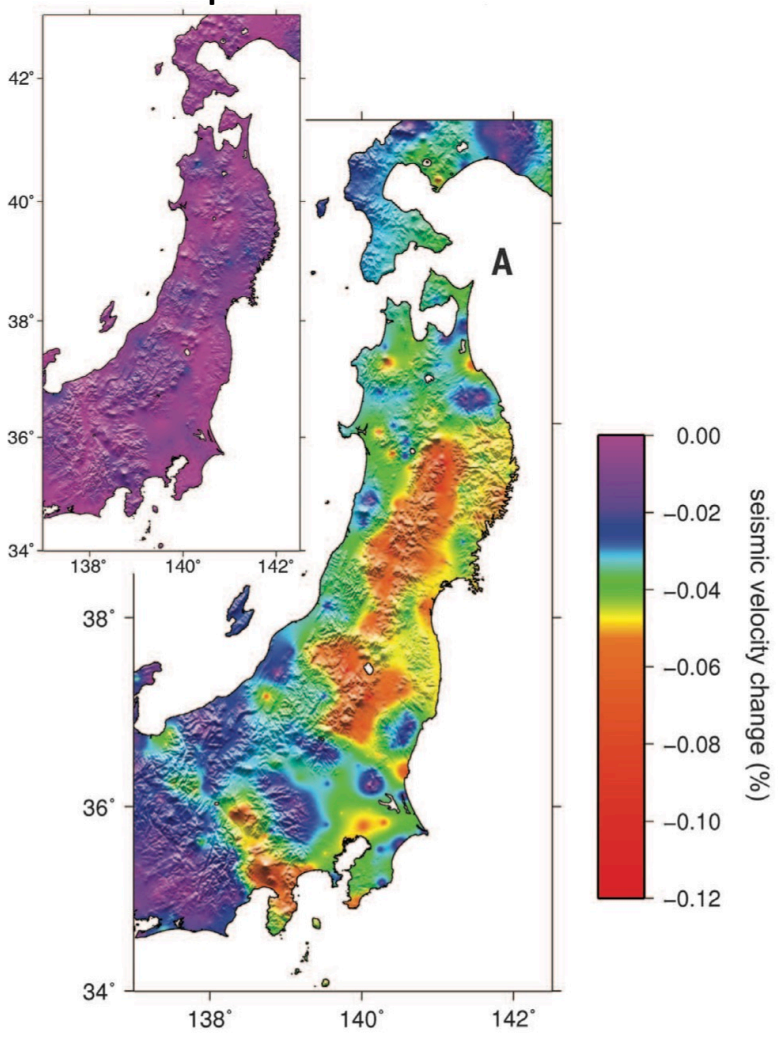


~100 Tb of raw data

Analysis made with the CIMENT MesoCentre at Grenoble

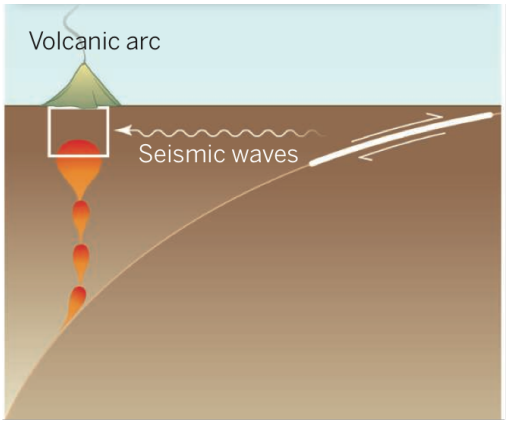
Regionalized crustal velocity changes during the 2011 Tohoku earthquake

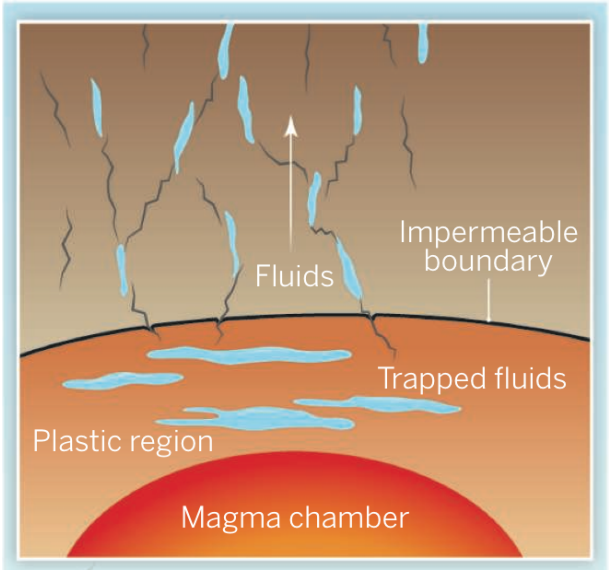
5 days before the earthquake



5 days after the earthquake

Perturbation of widespread hydrothermal fluids in volcanic areas

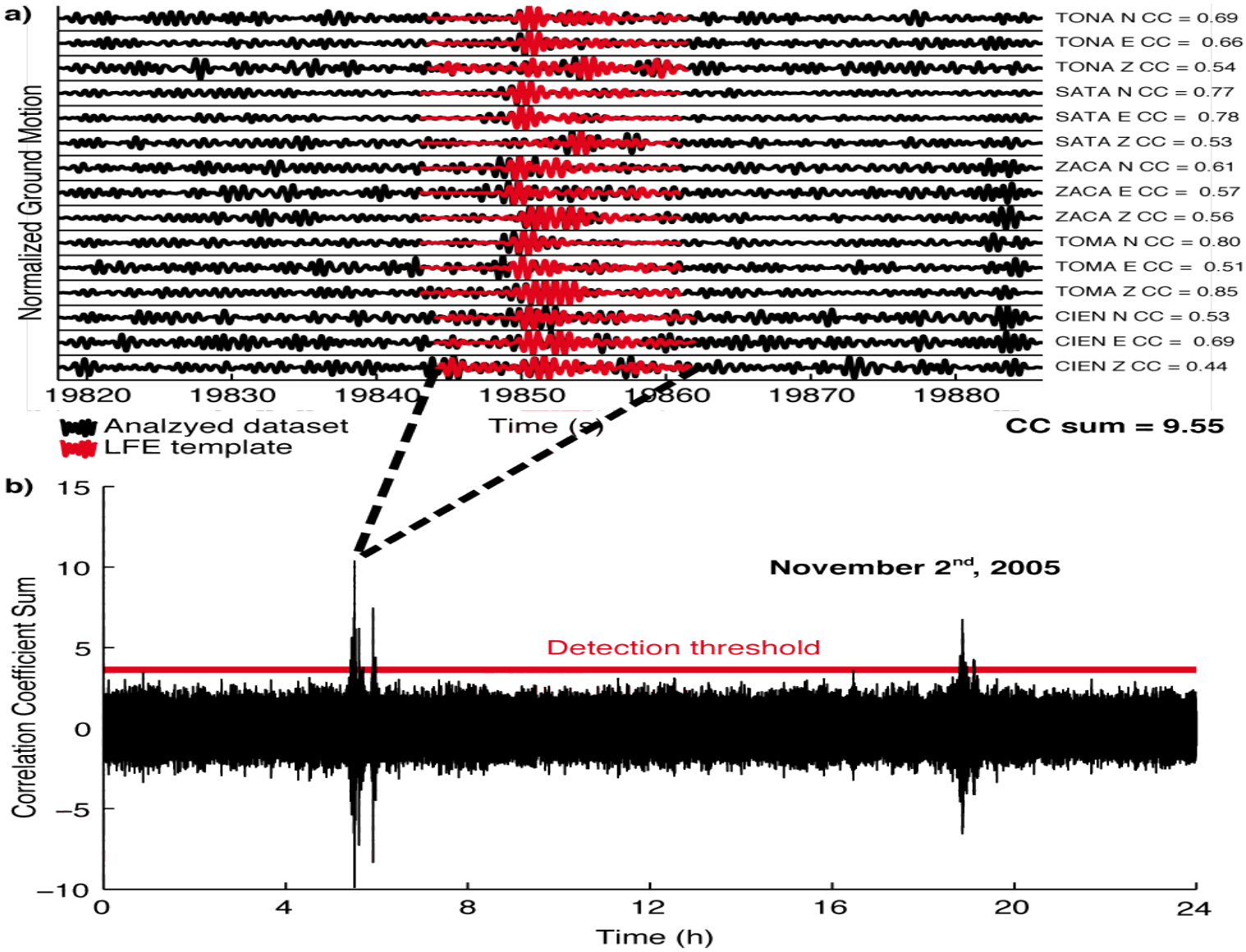




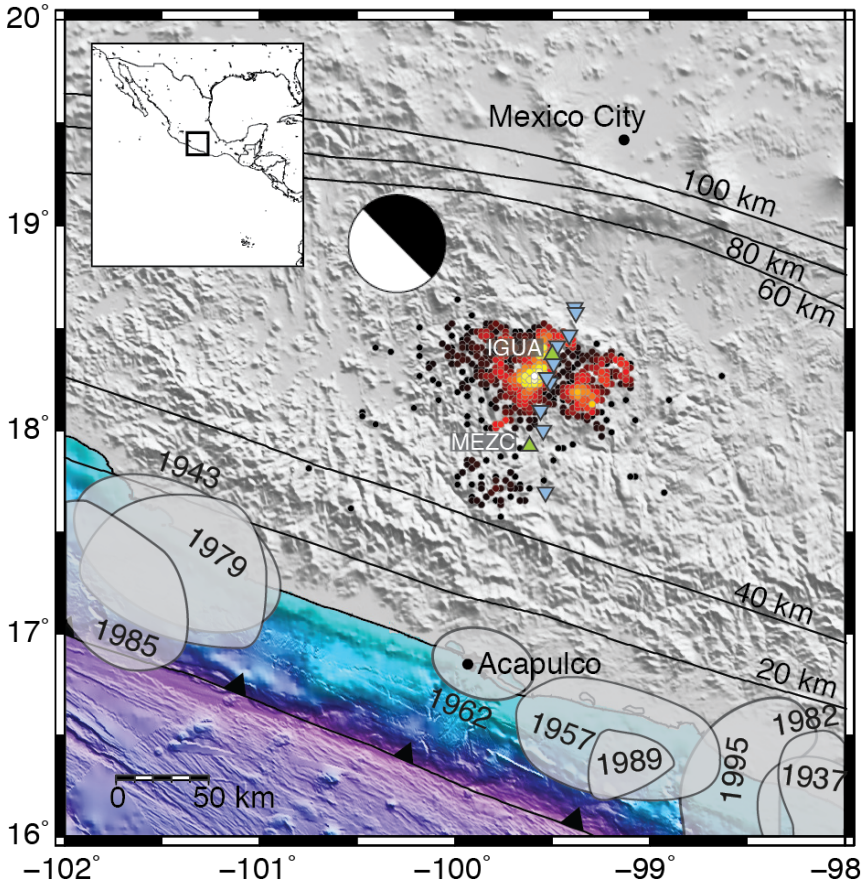
Brenguier et al., Science 2014

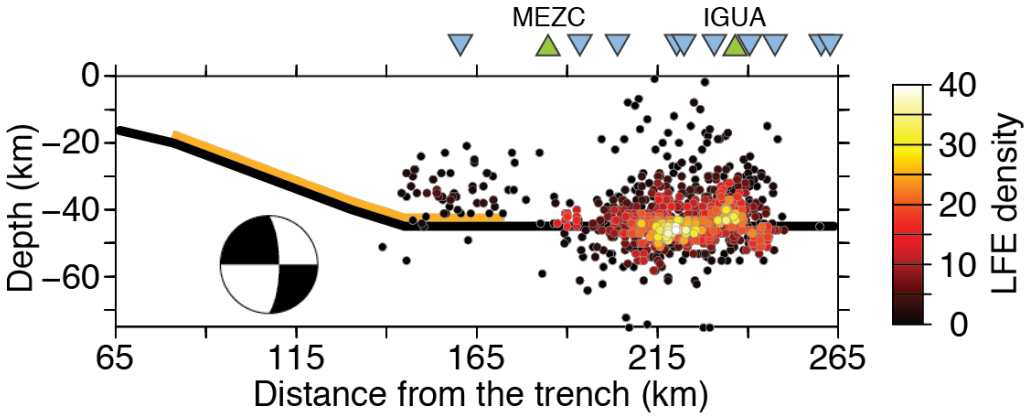
Discovery of weak seismic signals associated with slow deformation

Detecting template multiplets: Multi-station multi-component correlation



Application in Mexico: largest catalog of seismic events (low-frequency earthquakes, LFE)



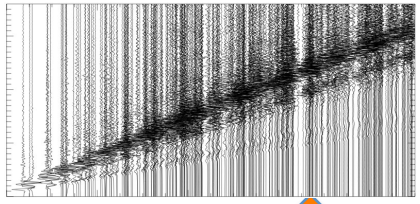


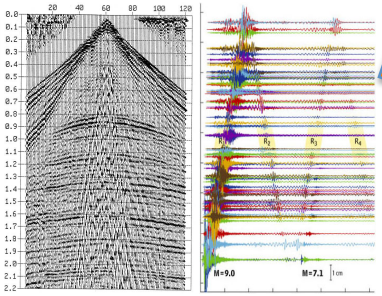
Frank et al., JGR, 2014 and Frank et al., GRL, 2013

1,849,486 LFEs detected over 2.5 years

- 1120 unique sources (800 more than next biggest catalog)
- >2000 LFEs per day on average

SEISMIC RECORDS



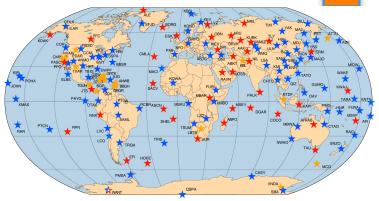


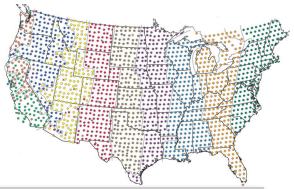
DATA ARCHIVING

DATA CENTERS

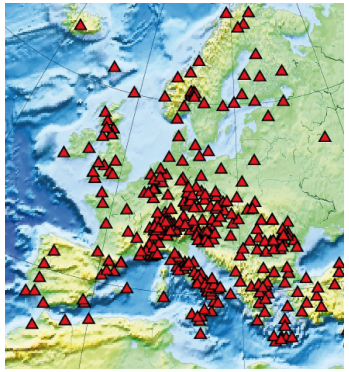
DATA INTEGRATION

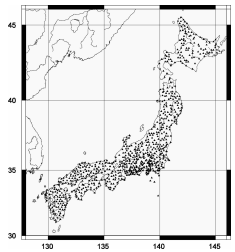
DATA COLLECTION





SEISMIC NETWORKS





HPC resources

infrastructure for efficient exploration of rapidly growing volumes of seismological observations

DATA ANALYSIS, MODELING, AND INVERSION

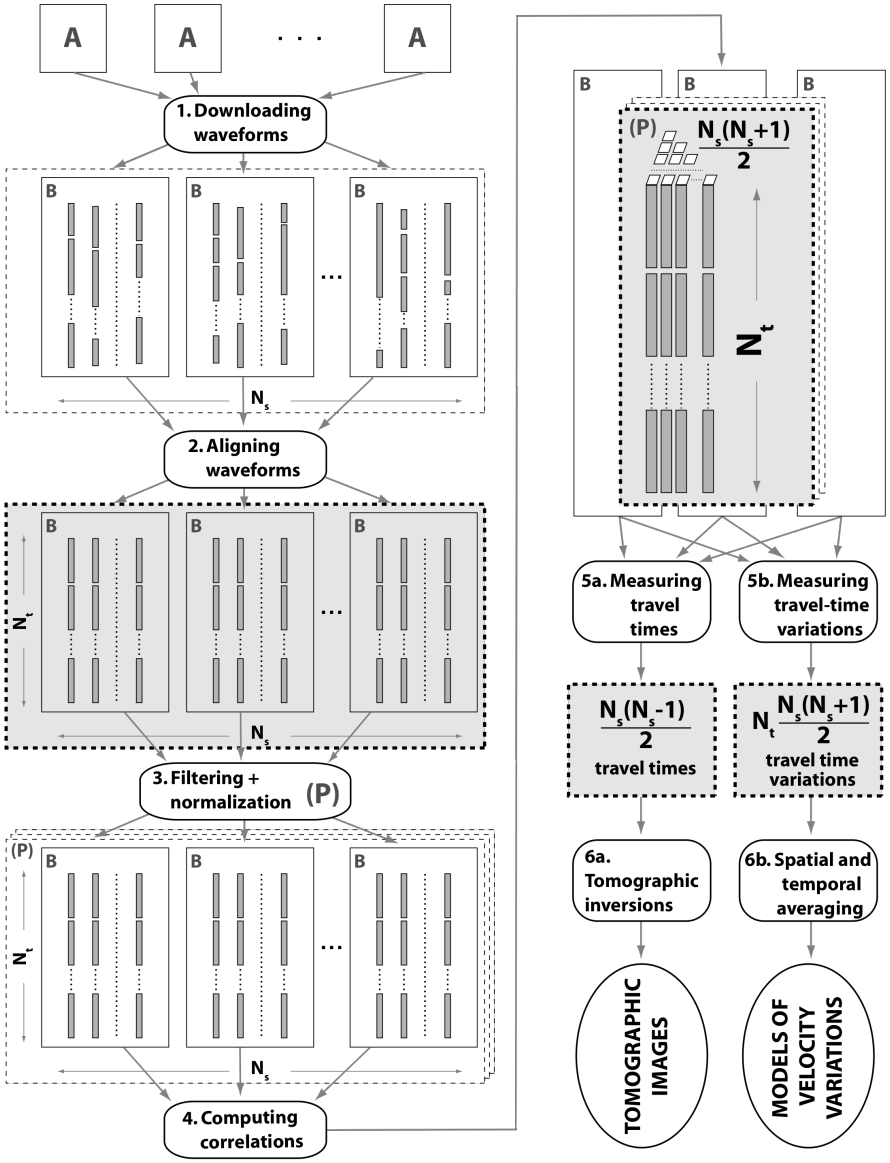
Earth's interior and dynamic

Earthquakes, volcanoes, natural hazards

Interaction of solid Earth with Ocean and Atmosphere

Analysis of continuous seismic data

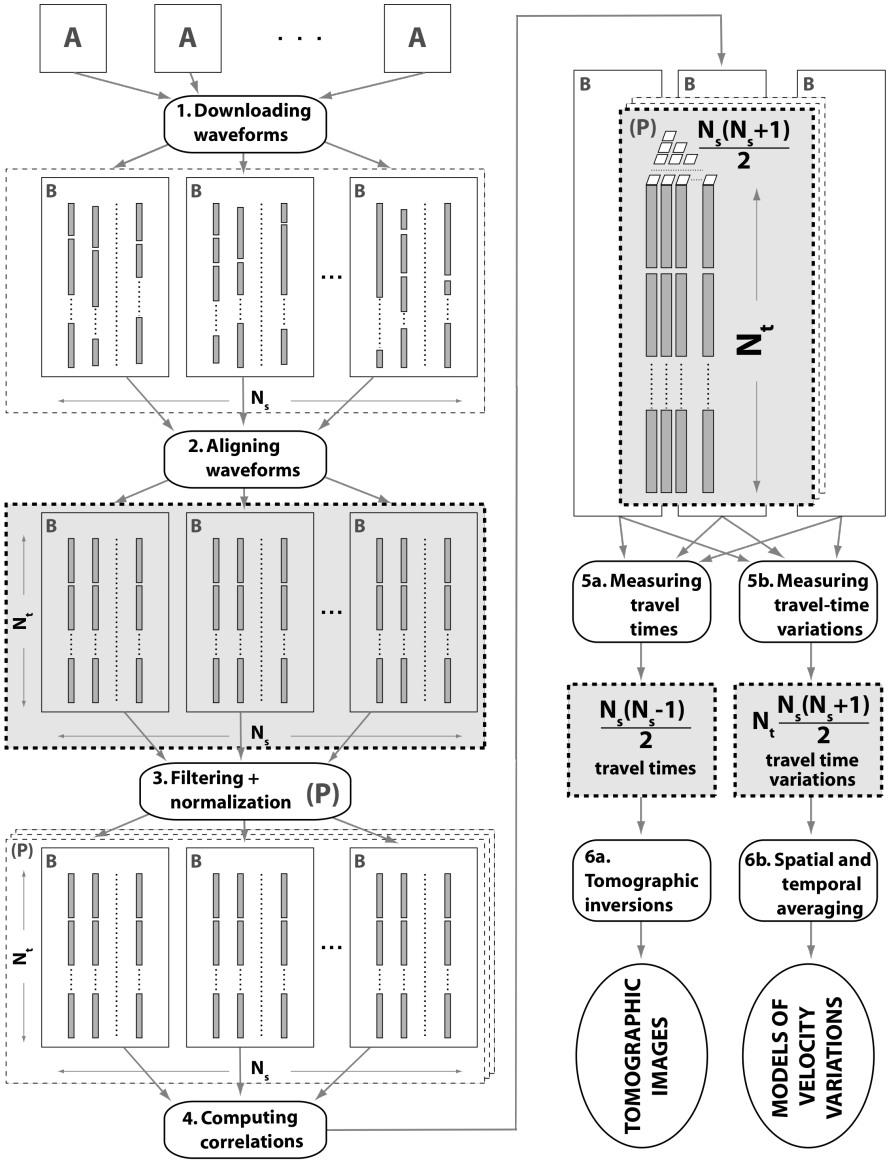
(A) seismological datacenters (B) data processing platforms



- **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**

Analysis of continuous seismic data

(A) seismological datacenters (B) data processing platforms

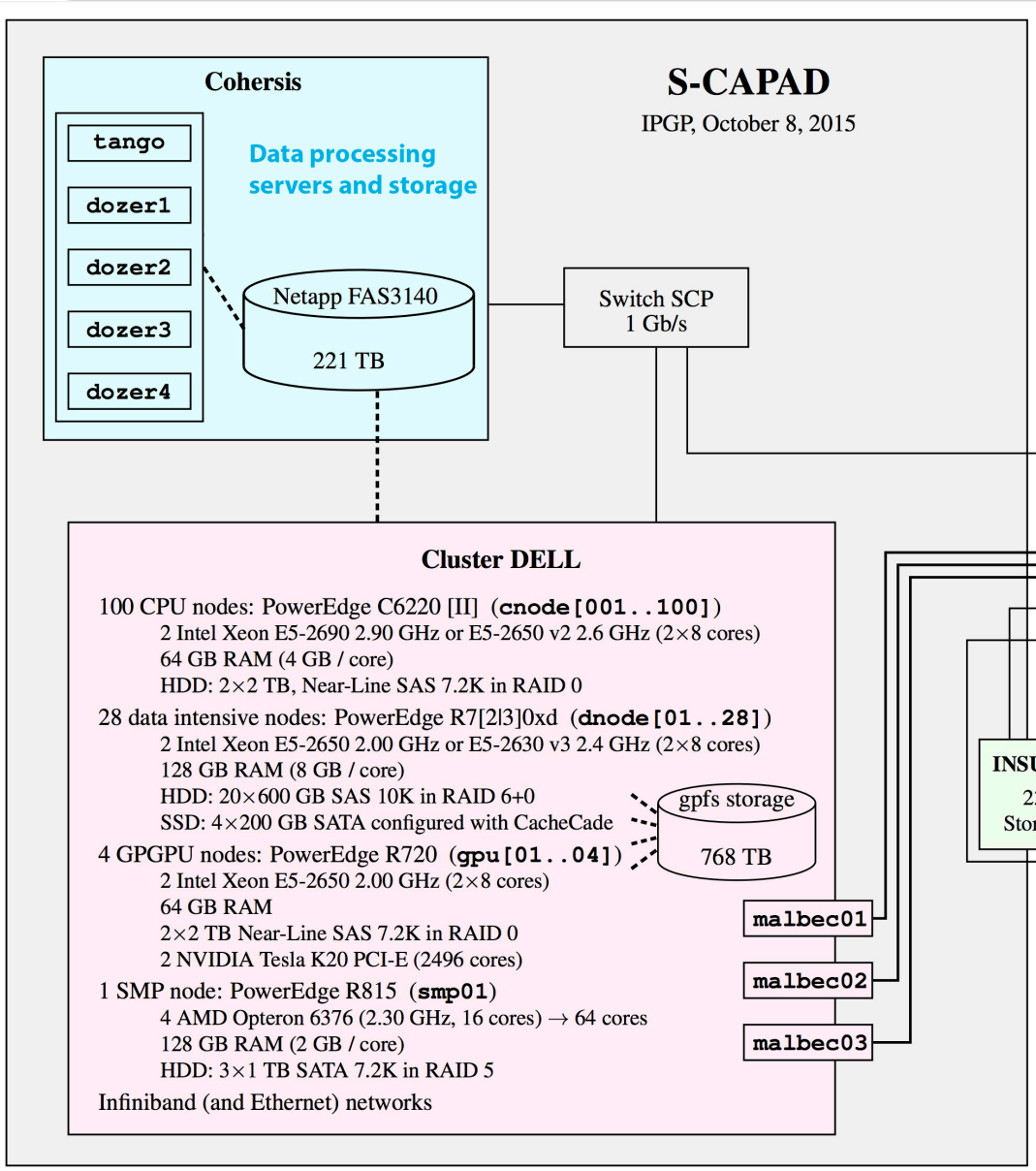


- **Large volumes of input raw data**
- **Analysis: sequences of very large number of simple operations: (digital filters, Fourier transforms, dot-products, …)**
- **Huge number of output files (N input-> N(N-1)/2 output)**
- **Repeating analyses**

Constraints on the computing resources

- **Long-term storage of large volumes of input and output data**
- **Parallelization**
- **Optimization of the data transfer and access**

MesoCentre S-CAPAD (IPGP): architecture dedicated to intensive data analysis



clustervision DELL

IPGP
INSTITUT DE PHYSIQUE DU GLOBE DE PARIS

Responsable technique: G. Moguilny

MesoCentre S-CAPAD (IPGP): architecture dedicated to intensive data analysis

- ~ 1 Pb of parallel file storage: efficient solution for processing large volumes of data with simultaneous computing at many nodes
- Computing power: CPU + GPU
- Parallel programming
- Stimulating environment to learn advanced computing approaches

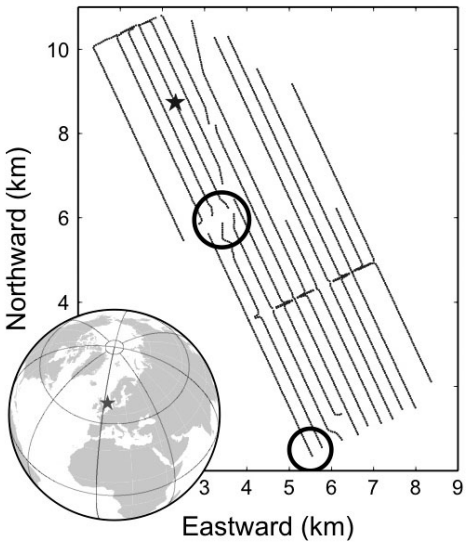
clustervision DELL

IPGP
INSTITUT DE PHYSIQUE DU GLOBE DE PARIS

Responsable technique: G. Moguilny

Example 1: noise based tomography at Valhall

(A. Mordret, D. Weissenbach, M. Landes, A. Romanenko)



6 hours of continuous records

4 components

250 samples per second

~110 Gb

> 26 millions cross-correlations

~2.2 Tb

GRID

3 month of computation on the European Grid Infrastructure

300 parallel jobs

MesoCentre S-CAPAD

2 days of computation on 8 NVIDIA Tesla GPU

GPU optimized code with CUDA

Example 2: Massive event detection in Mexico

(W. Frank and A. Romanenko)

Main operation: dot product repeated billions of times

~ 5 Tb of input data

Sequential algorithm

29 templates with ~10000 total detections

months of computations

Parallel, GPU-accelerated at S-CAPAD

~ 5000 templates with many millions of detections

~20 days of computation on 8 NVIDIA Tesla GPU

GPU optimized code with CUDA

end