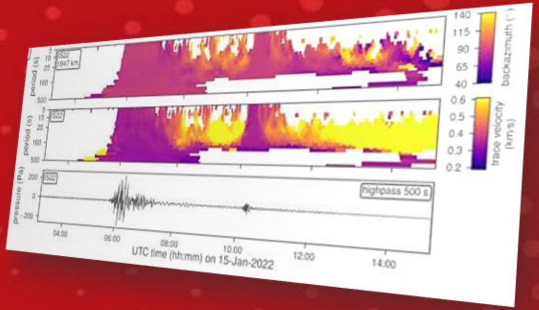
High-resolution array processing method:

MLE vs. TDOA – Evaluation, implementation and applications

A. Le Pichon[1], M. Charbit[2], B. Poste[1], J. Vergoz[1]

[1] CEA-DAM-DIF, F-91297 Arpajon, France

[2] Telecom ParisTech, Institut Polytechnique de Paris



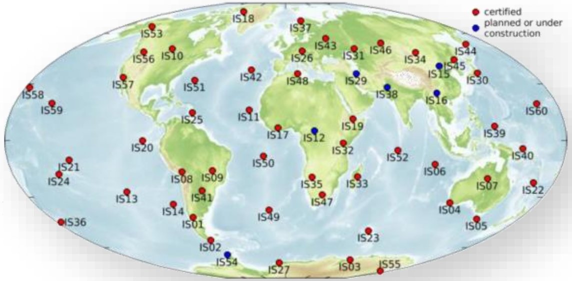
Outline

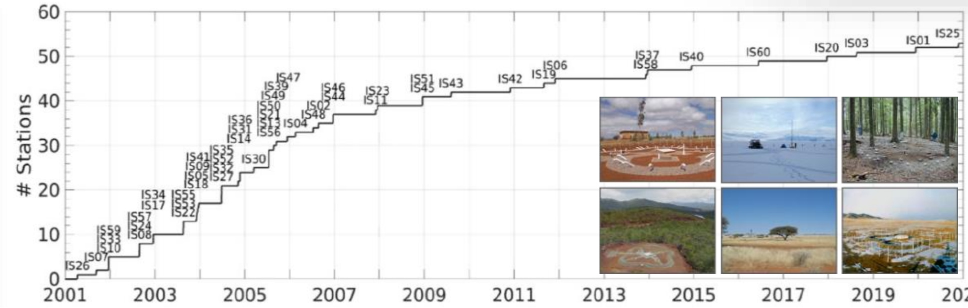
- Overview of operational detection method used at the French NDC
- Examples of applications of infrasound products
- Recent advances in the implementation of MLE approach
- Evaluation through synthetic dataset and real cases

IMS infrasound network today

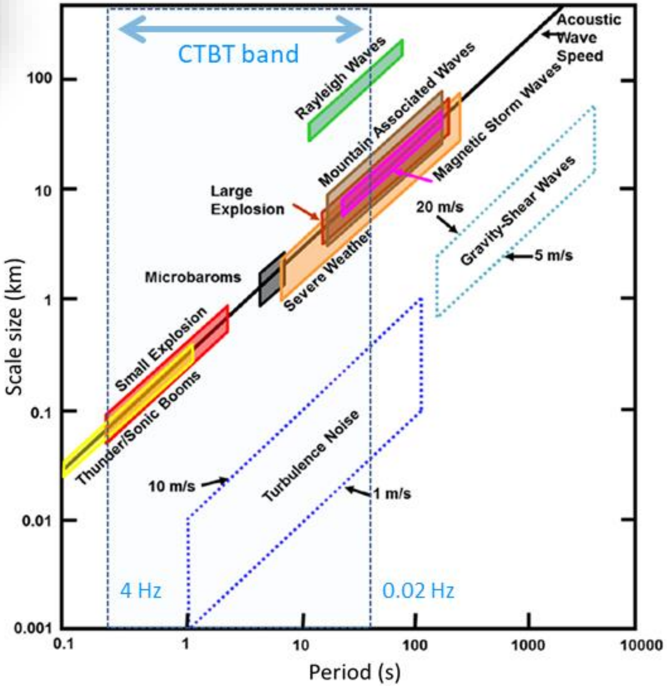
International Monitoring System (IMS)

- Operational global network of 60 arrays
- ~90% operating stations
- High-quality continuous measurements since 2000





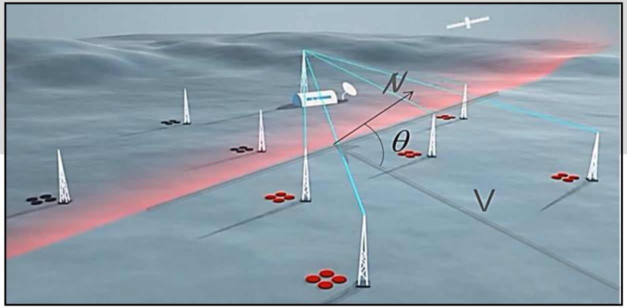
Adapted from Silber, 2014



A "zoo" of infrasound sources (0.02-4 Hz)

- Ocean waves, bolides, earthquakes, volcanoes, hurricanes…
- Need robust detection, location and source characterization methods
- Opportunity to calibrate the network and promote civil and scientific applications

Broadband detection algorithm



Time Difference Of Arrival (TDOA)

- PMCC = Progressive Multi-Channel Correlation (Cansi, 1995) ➔ Julien's Lab. Session
- Challenge: highly variable noise conditions
- Unique source: raw data are filtered in narrow bands
- Time delay estimates of planar wavefront

$$\tau_{ij} = \arg\max_{\tau}\left(C_{ij}(\tau)\right)$$

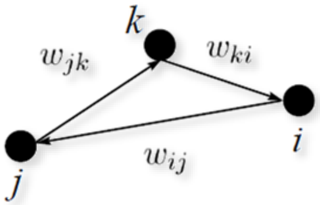
- Function of test: consistency

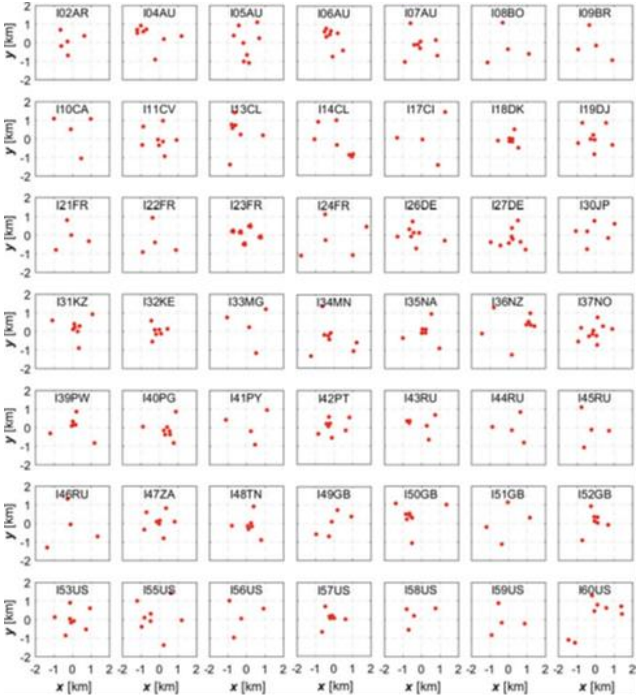
$$C = \frac{6}{M(M-1)(M-2)} \sum_{i>j>k} r_{ijk}^2 < c_{Threshold}$$

$$r_{ijk} = w_{ij} + w_{jk} + w_{ki}$$
$$w_{ij} = \tau_i - \tau_j . \; r_{ijk}$$

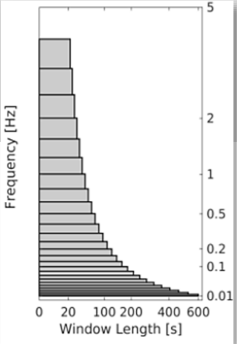
- 2D least-squared solution

$$S = \left(A^T A\right)^{-1} A^T \Delta T, \quad V = \frac{1}{\sqrt{S_x^{\,2} + S_y^{\,2}}}, \theta = \operatorname{atan}\left(\frac{S_y}{S_x}\right)$$

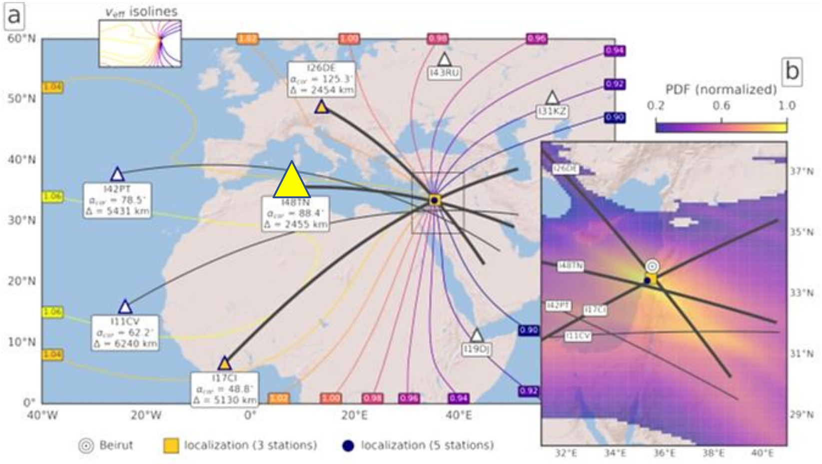


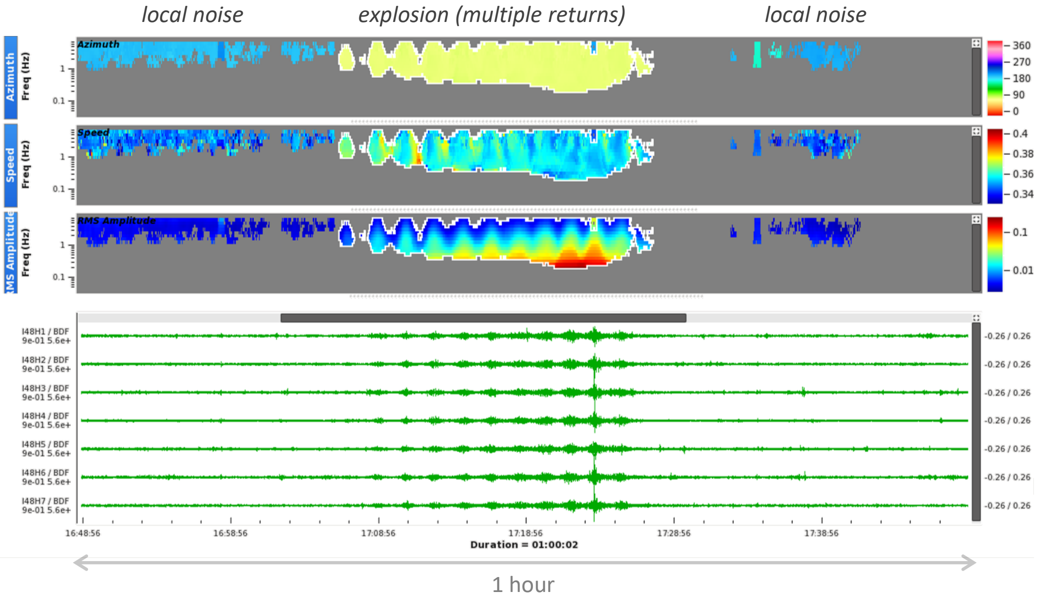


Example of PMCC analysis
Beirut explosion (2020)



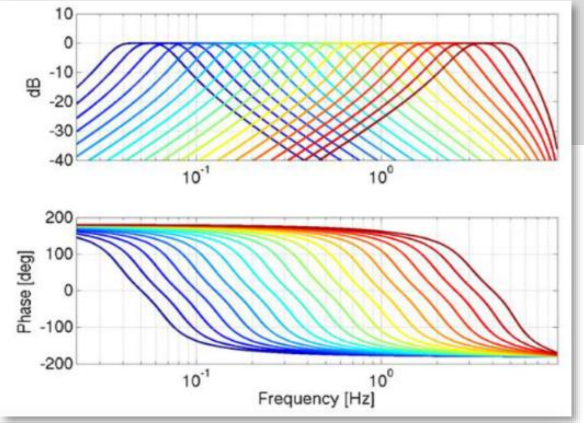
- Explosion of ammonium nitrate stored at the port of Beirut
- 5 IMS stations: from 2450 to 6200 km
- Localization: Δloc = 38 km, Δt = 20 s
- 0.01-4 Hz, 1/3-octave band scheme (Garcés, 2013)
- Window length scaled to the period
- 11 stratospheric returns at IS48-Tunisia





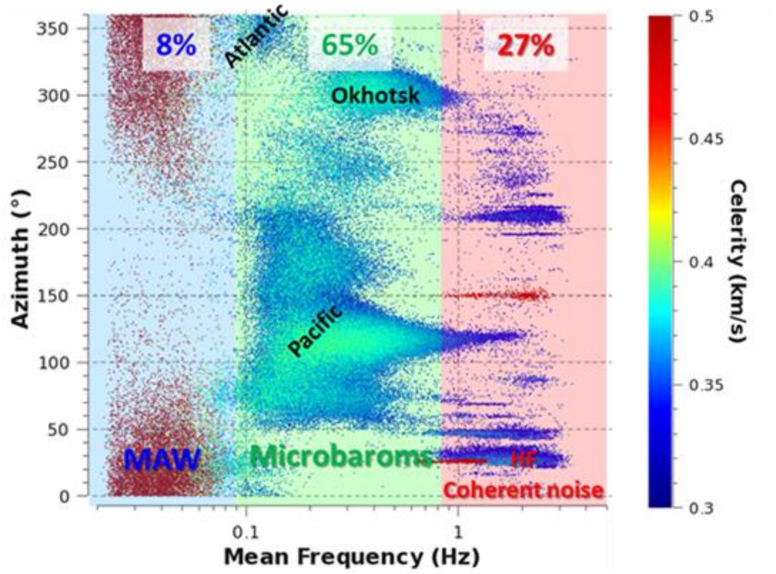
Pilger, C. et al. (2021). Sci. Rep. https://doi.org/10.1038/s41598-021-93690-y

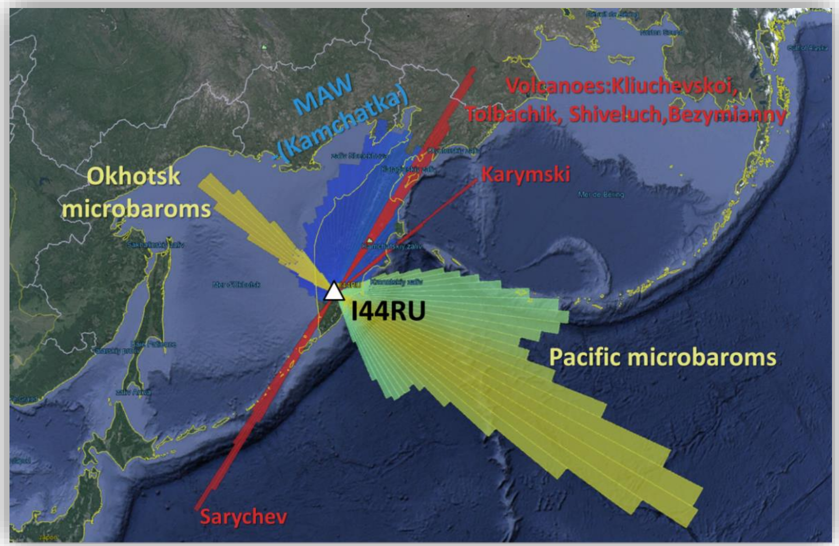
Characterizing the global coherent infrasound wavefield



Standardized processing scheme

- Real time process, 26 log-scaled frequency bands (0.01-4 Hz)
- Stable responses in phase and amplitude
- Statistically and systematically characterizing coherent infrasound





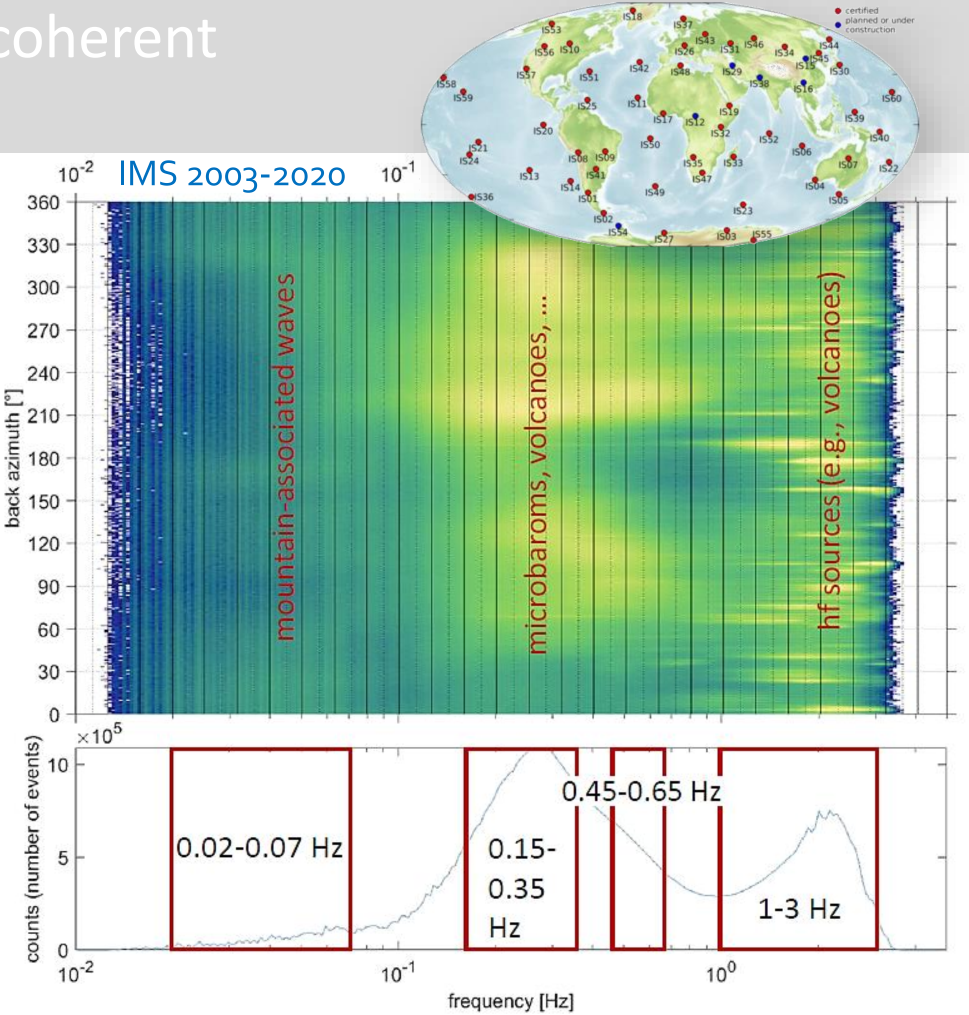
Characterizing the global coherent infrasound wavefield

New IMS broadband bulletin products for atmospheric studies

- Full IMS infrasound data set reprocessed (2003-2020)
- One-third octave frequency band between 0.01 and 4 Hz
- Detection lists: >81 million entries with center frequencies between 0.02 and 3.5 Hz (all stations)
- Infrasound products (Hupe et al., 2022, ESSD) https://doi.org/10.25928/bgrseis_mbhf-ifsd

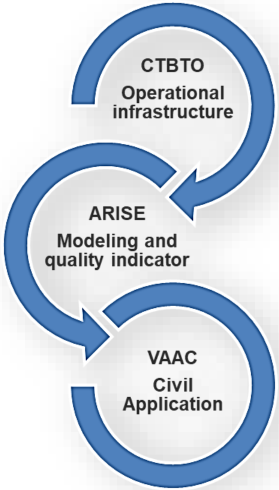
Examples of recent applications

- Global mountain wave characterization (Hupe et al., 2019)
- Benchmarking microbarom radiation and propagation model (De Carlo et al., 2021)
- Identifying signatures from rocket launches (Pilger et al., 2021)
- Lightning activity (Farges et al., 2021)
- Volcanic eruptions (Matoza et al., 2019; Marchetti et al., 2019)



Volcanic Information System (VIS)

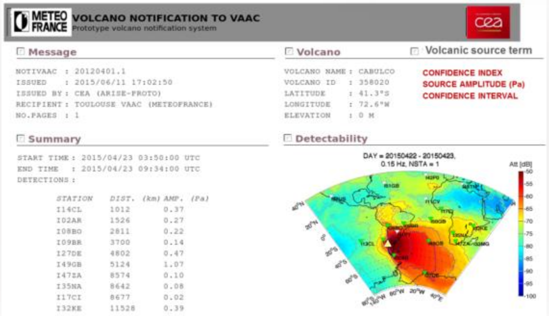
European Catalogues for Geohazard Analysis Infrastructure



A prototype system for monitoring volcanic eruptions

- Volcanic Information System (VIS) has been developed and evaluated through the ARISE EU project (H2020, 2015-2018)
- Use PMCC detection lists and consider long-range propagation effects
- Proposed approach tested with Toulouse VAAC mandated by the ICAO
- VIS was able to detect all the major eruptions (VEI > 3)

(Marchetti et al., 2019)



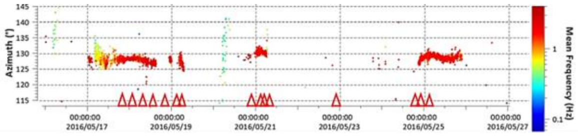
From science to operation

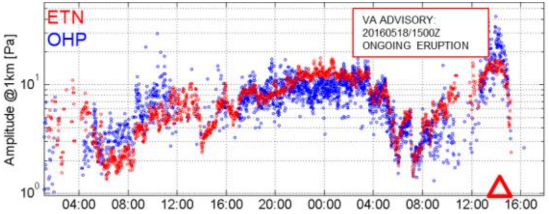
- Geo-Inquire (INFRA-SERV, 2022-2026)
- Data platform to reinforce synergy between different fields of geophysics for georisk and geohazard research
- VIS: service of the class "European Catalogues for Geohazard Analysis Infrastructure" capitalizing on multi-year archives of IMS observations

Etna test case: CEA / UNIFI / VAAC

CEA / UNIFI / VAAC

Observations at OHP (1050 km) / VAAs





Recent developments:
Multi-Channel Maximum-Likelihood (MCML) method

Collaborations with Telecom ParisTech

- 2011-2012: audit / improvement of PMCC algorithm
- 2014-2015: expert group meetings at CTBTO
- 2014-2017: PhD supervision (A. Nouvellet)
- 2019-2020: MLE in multi-band framework
- 2021-2024: PhD supervision (B. Poste) – High resolution method for source separation (e.g. Ouden et al., 2020)

Maximum Likelihood Estimator (MLE)

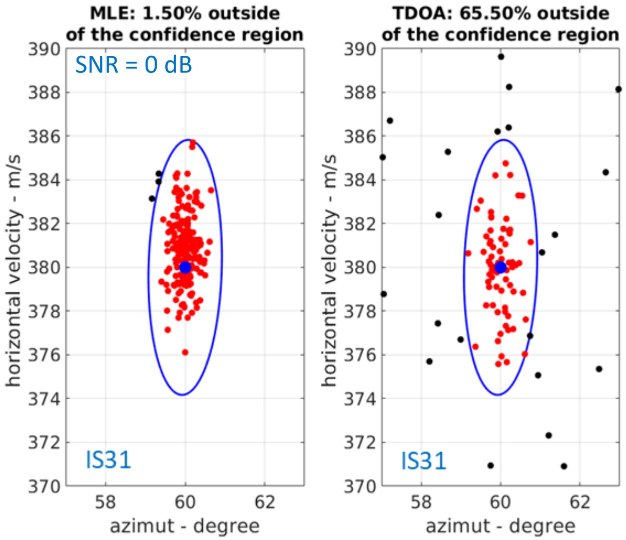
- Likelihood function derived from a multi-sensor statistical model detection (see poster by B. Poste)
- Estimation of the slowness vector, confidence intervals, SNR
- Detection criteria: Generalized Likelihood Ratio Test (GLRT, Burgess 1994)
- New asymptotic results: GLRT under the null hypothesis leading to the computation of the corresponding p-value (Gibbons, 1975)
  - *Benefits:* grid search approach, probability of null-hypothesis significance, SNR, asymptotic errors of the slowness vector, robustness
  - *Limitations:* computational cost

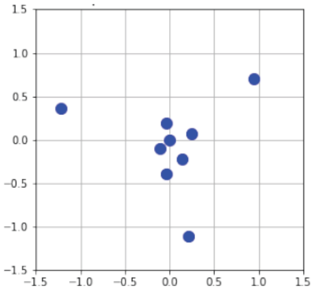
Asymptotic errors of the estimation

- Covariance matrix of the MLE estimator

$$R(a,v) = J\Gamma_0 J^T \quad \text{with} \quad J = \begin{bmatrix} -v\cos(a) & v\sin(a) \\ v^2\sin(a) & v^2\cos(a) \end{bmatrix}$$

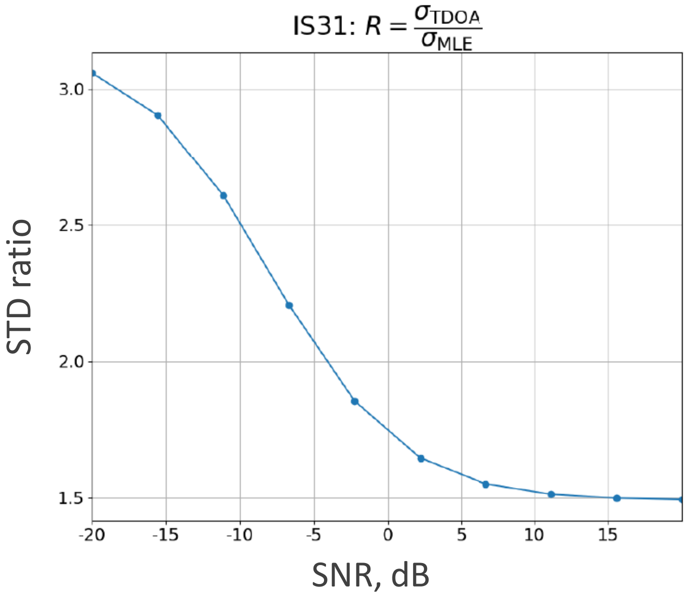
$$\Gamma_0 = \frac{\int 4\pi^2 f^2 |H(f)|^4 \, (M \times \mathrm{SNR}_\star + 1)\, df}{\left(\int 4\pi^2 f^2 M \times \mathrm{SNR}_\star \times |H(f)|^2 df\right)^2}$$





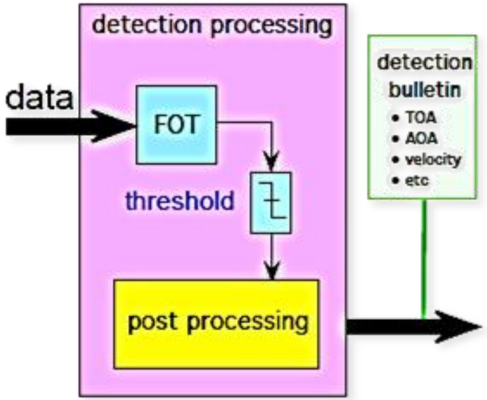
- Integrated expression of the covariance matrix

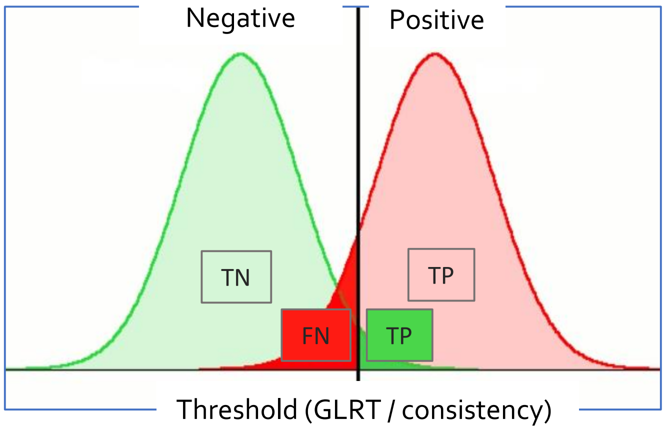
$$\bar{R} = \int_0^{360} \int_{V_{\min}}^{V_{\max}} R(a, v, \mathrm{SNR}, f_{\max} f_{\min}, X)\, da\, dv$$

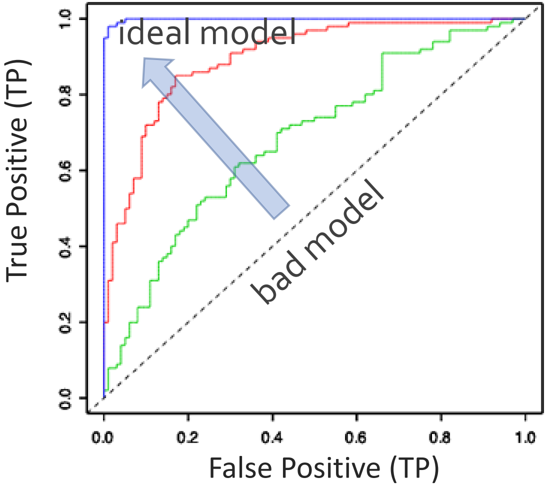


Evaluation protocol using synthetic dataset

- **Receiver Operating Characteristic** (ROC) curves: evaluation of TDOA/MLE according to **true positives** vs **false positives**
- Wave parameters randomly distributed within (θ; v): 0-360° and 300-800 m/s (5° and 20 m/s steps)
- Synthetic noise sequences and signals are generated:
  - H0 hypothesis (1000) with white and real wind noise
  - H1 hypothesis (2000): synthetic signals
- Various experiments: SNR, array geometry, number of sensors

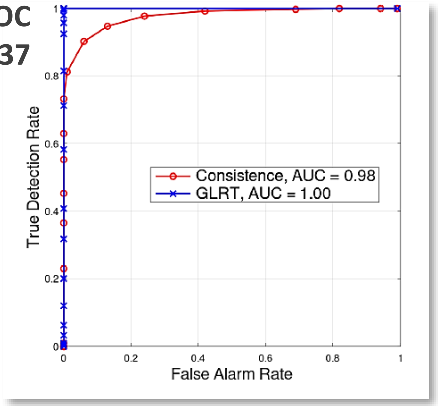




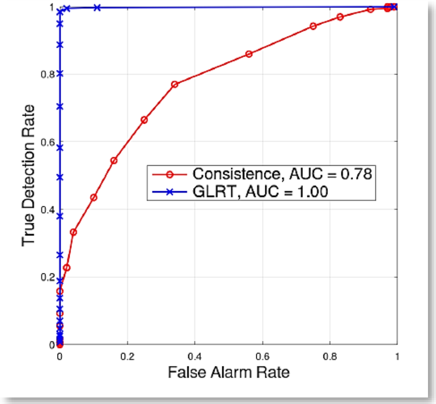


Evaluation results using synthetic dataset

ROC
IS37

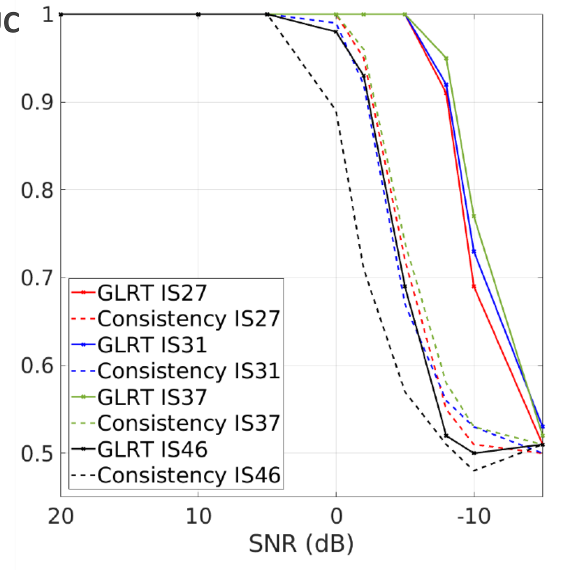


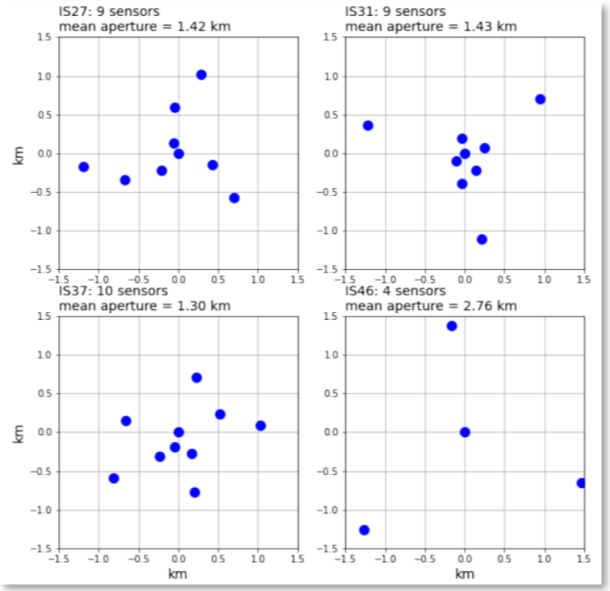
SNR=0 dB



SNR=-5 dB

AUC



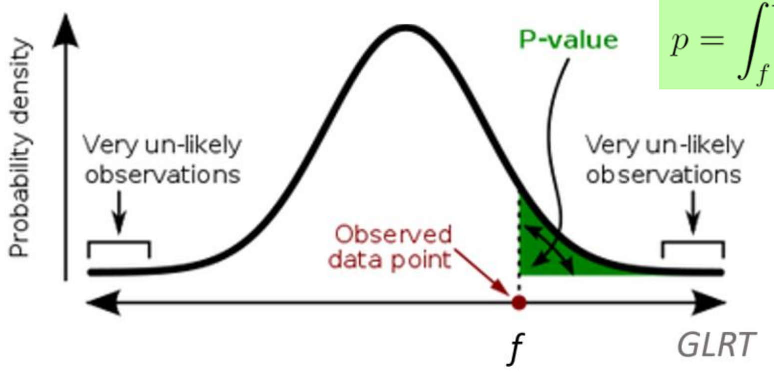


Area Under the ROC Curves (AUC): probability of **true positives** versus **false positives** (Hanley and McNeil, 1982)

- AUC=1 for SNR>5 dB; overall decrease of the AUC with decreasing SNR
- At low SNR (<0 dB), GLRT performs better than the consistency
- Comparable AUC for weak number of sensors (IS46, 4 elements)

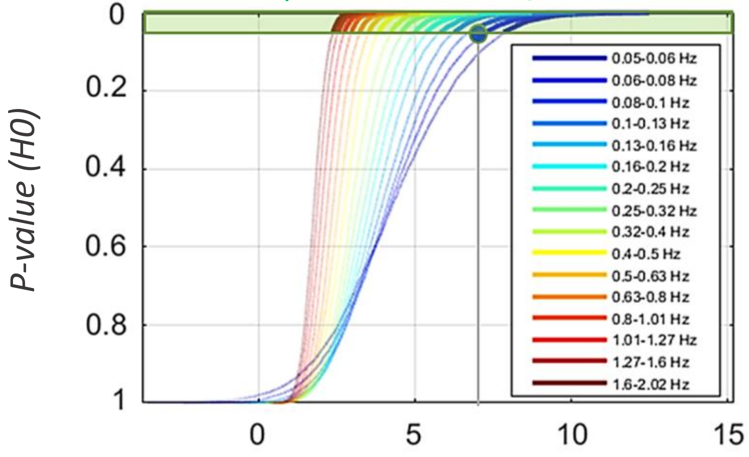
P-value in statistical hypothesis tests
Statistics derived from GRLT under null-hypothesis

- Quantification of the statistical significance of the detection
- ***P-value*: objective decision fixed by the analyst** in terms of probability of false alarm
- Asymptotic distribution of the GLRT test under the noise only hypothesis
  - Cumulative distribution of *f* under H0
  - Probability to observe H0 for a test larger than *f*
  - For *p* < 1%, the null-hypothesis is unlikely
- For each band: invert GLRT threshold from *p-value*



$$p = \int_f^{+\infty} p_{F|H_0}(t)dt$$

Probability of false alarm <5%

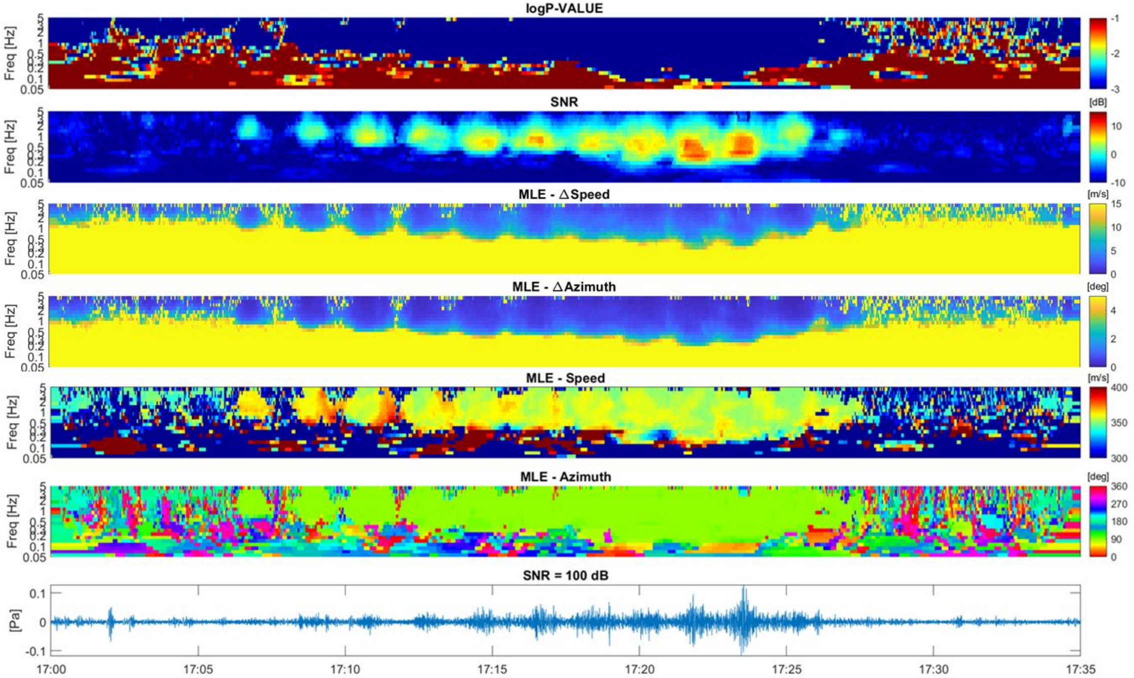


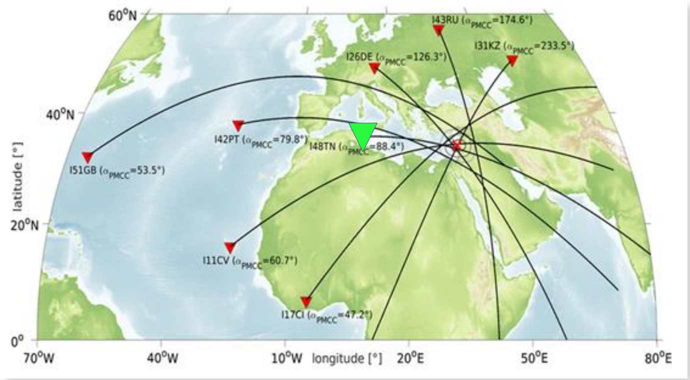
$$GLRT = \max_{(a,v)} \left(\frac{S(a,v)}{T(a,v)} - 1 \right)_+ \in [0, M-1)$$

$$\begin{aligned} S_N(\theta, v) &= \mathbb{1}_M^T R_N(\theta, v) \mathbb{1}_M \\ T_N(\theta, v) &= \operatorname{trace}(R_N(\theta, v)) \end{aligned}$$

Example of MCML application: Beirut explosion

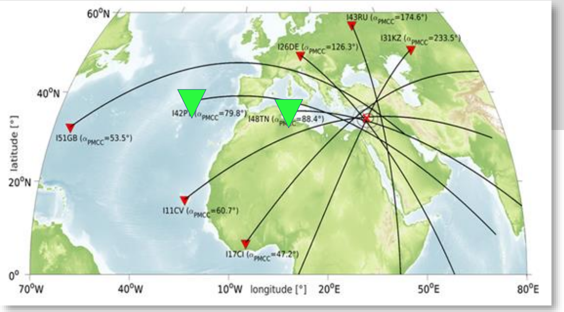
IS48 – 2455 km

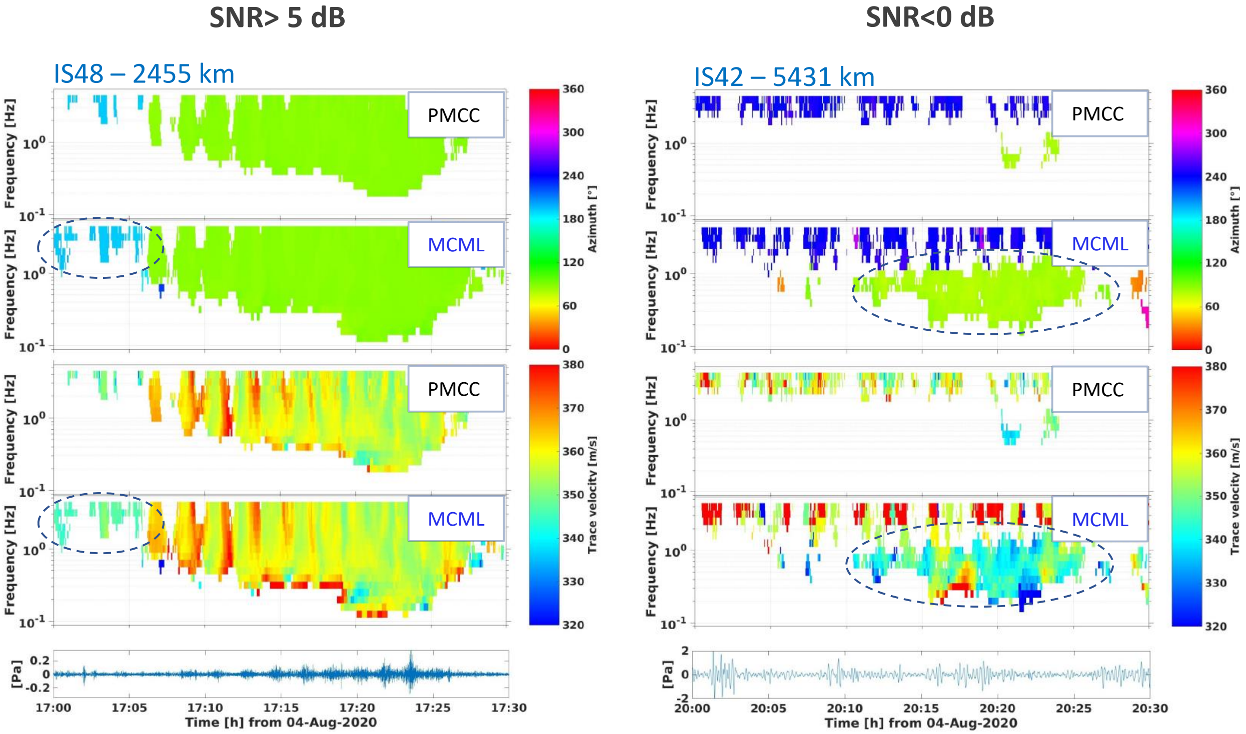




- PMCC and MCML implemented in the same processing framework
- Signal separation into at least 11 successive stratospheric propagation paths characterized by p-values lower than 0.01
- Errors of ~5 m/s and ~1° (SNR>5 dB)

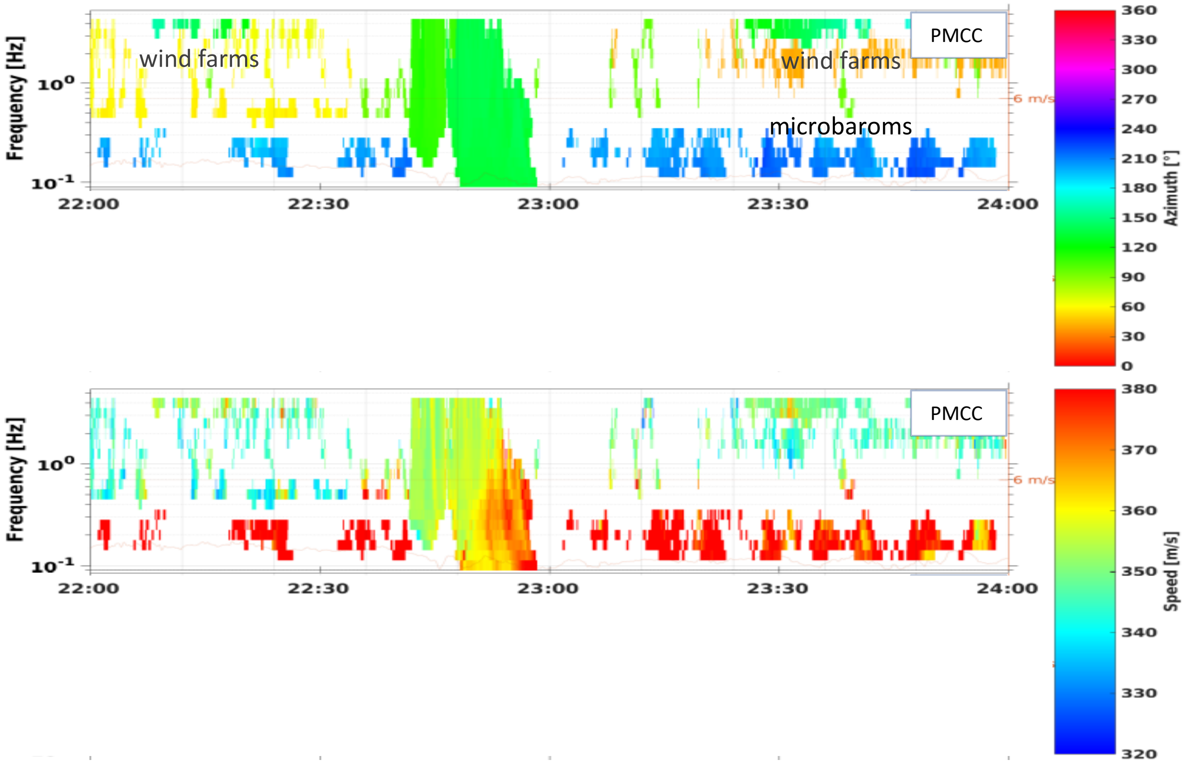
MCML vs. PMCC: Beirut explosion

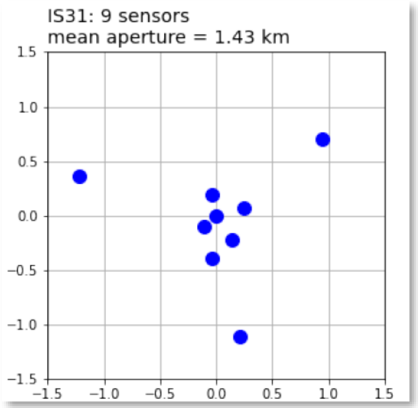




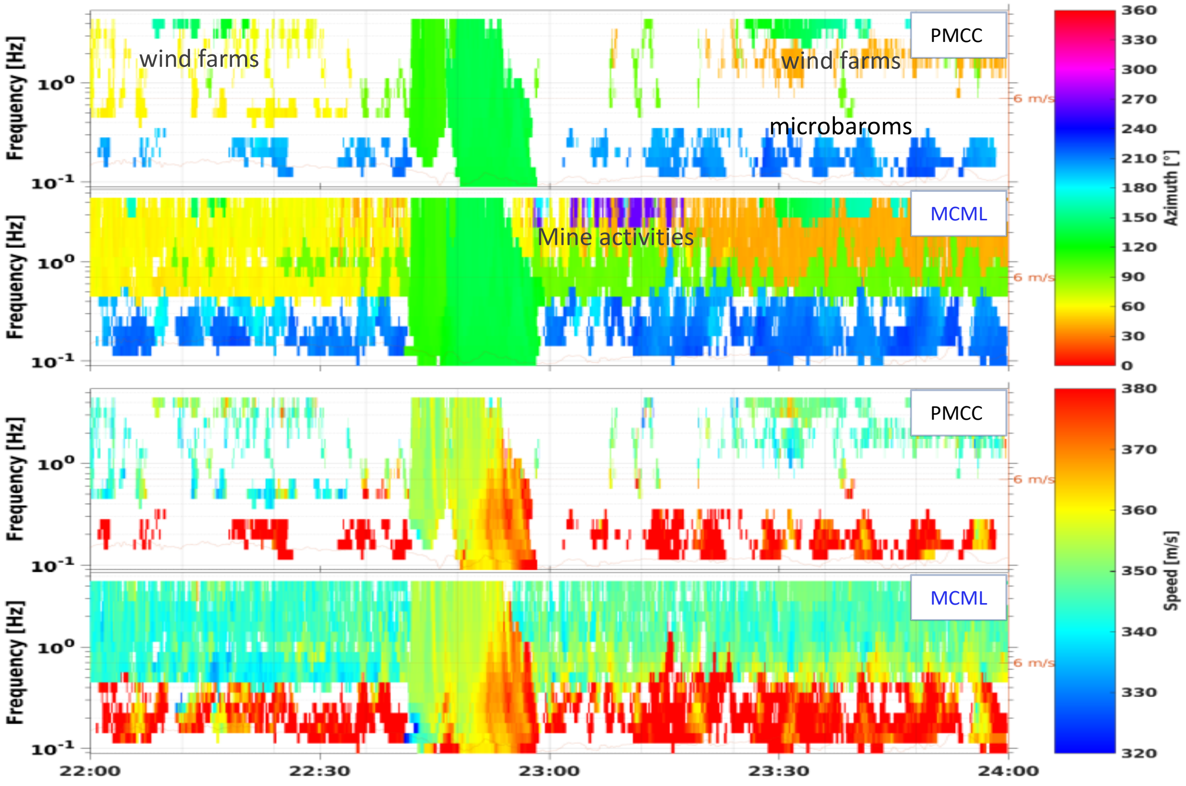
- 26 log-scaled bands (0.01-4 Hz)
- Detection results (*p-value*<0.01)
- Comparable results in high SNR conditions
- MCML yields enhanced detection in the lower part of the frequency domain where the SNR decreases

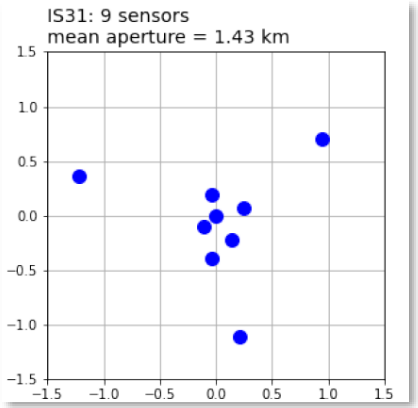
MCML vs. PMCC: a typical soundscape at IS31





MCML vs. PMCC: a typical soundscape at IS31





- Comparable estimates at high SNR (>5 dB)
- MCML improves detection capability at low SNR
- The consistency fails to detect SOI when permanent signals overlap

Summary & Perspectives

- MCML method has been implemented in the PMCC time-frequency domain framework
- Using the likelihood expression of MLE, SNR and asymptotic errors are calculated
- From the GLRT, the p-value allows to fix an objective decision in terms of probability of false alarm
- MCML vs. TDOA evaluated using synthetics and real signals
  - MCML performs better than TDOA, especially at low SNR values
  - In 0.01-4 Hz band, MCML yields enhanced detection results in low SNR conditions more especially when multiple signals interfere
- Limitations: computational cost is reduced, parallelized grid search approach
  - e.g.: ~30 s for one hour of signals recorded at a 10-element array, sampled at 20 Hz (DELL PowerEdge workstation, 56 cores)
- A promising method; further optimization needed for a real-time implementation