SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

P2.1-797

# On the use of metamodels for Bayesian localization and characterization of infrasound events

Alexandre Goupy[1*], Christophe Millet[1,2], Pierre Sochala[1]

[1]CEA, DAM, DIF, F-91297, Arpajon, France
[2]ENS Paris-Saclay, F-91190 Gif-sur-Yvette, France
*Actually at EDF Lab, F-91120 Palaiseau, France

## INTRODUCTION AND MAIN RESULTS

**This work investigates the use of generalized Polynomial Chaos (gPC)-based metamodels, embedded in a Bayesian framework, to achieve more robust and cost-efficient source localization.** Application to the 2018 Bering Sea bolide event: (1) reduced localization error from ~220 km (IDC estimate) to less than 50 km, consistent with NASA satellite observations; (2) highlighted the critical impact of arrival-time extraction uncertainty, comparable to atmospheric uncertainties; (3) revealed the **ambivalent role of station IS59**, whose contribution improves localization but also introduces significant uncertainty, illustrating the need of adaptive station weighting.

SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

# On the use of metamodels for Bayesian localization and characterization of infrasound events

Alexandre Goupy[1], Christophe Millet[2,3], Pierre Sochala[2]
[1]EDF Lab, F-91120 Palaiseau, [2]CEA, DAM, DIF, F-91297, Arpajon, [3]ENS Paris-Saclay, F-91190 Gif-sur-Yvette

P2.1-797

## From forward modelling to inference

Source localization remains challenging due to atmospheric variability, uncertainties in signal picking, and the computational cost of full-wave propagation models. Bayesian inference provides a principled framework for combining detections with prior knowledge (here provided by IDC), but its practical use requires fast and accurate forward models.

**Our goal:** leverage **Polynomial Chaos Expansions (PCE)** as an efficient surrogate model for Bayesian localization. The arrival time at station $\#j$ is emulated by:

$$t_j(\mathbf{x}) \simeq \sum_{k=1}^{K} a_{jk} H_k(\mathbf{x}),$$

$\mathbf{x}$ denotes the source location, $\{H_k\}$ is the PCE basis, and $\{a_{jk}\}$ is the set of coefficients computed using least square regression from simulated observations

$$\mathbf{a}_j = \mathrm{argmin}_{\mathbf{a}_j \in \mathbb{R}^p} \left\| \mathbf{t}_j - \mathbf{H}\mathbf{a}_j \right\|_2,$$

where $\mathbf{a}_j = [a_{jk}]$; $\mathbf{H} = [H_k(\mathbf{x}_m)]$ and $\mathbf{t}_j = [t_j(\mathbf{x}_m)]$ are computed from a sampling of source locations.

Assuming that arrival time measurement errors at distinct stations **are independent**, the joint likelihood $\mathrm{P}(\mathbf{t}|\mathbf{x})$ factorizes so as to give

$$\prod_j \mathrm{P}\left(t_j|\mathbf{x}\right) \propto \prod_j \sigma_j^{-1} \exp\left[-\epsilon_j^2(\mathbf{x})/(2\sigma_j^2)\right],$$

where $\epsilon_j$ is the difference between the observed arrival time and the PCE prediction.

## Application to the Bering sea bolide

On Dec. 18, 2018, a small asteroid entered Earth's atmosphere over the Bering Sea. We apply the Bayesian method using 6 IMS stations to test its ability to reproduce the reference locations (NASA [1]).

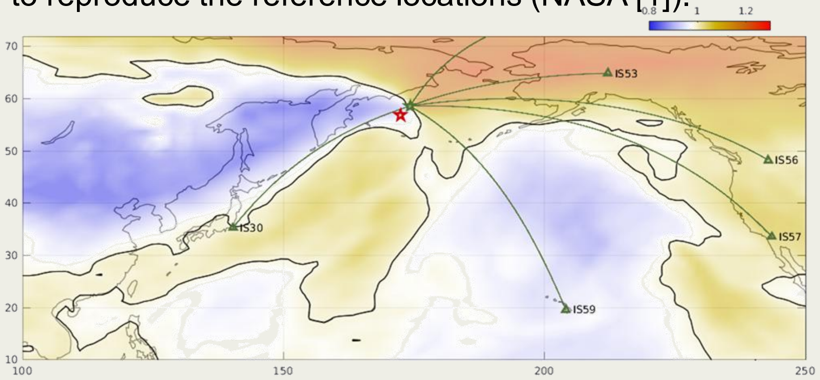



**Fig. 1 :** Eff. sound speed ratio; source locations, IDC: ★, NASA: ★.

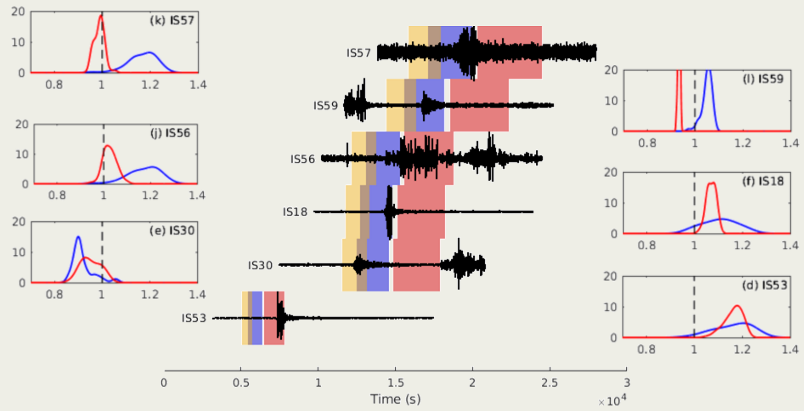
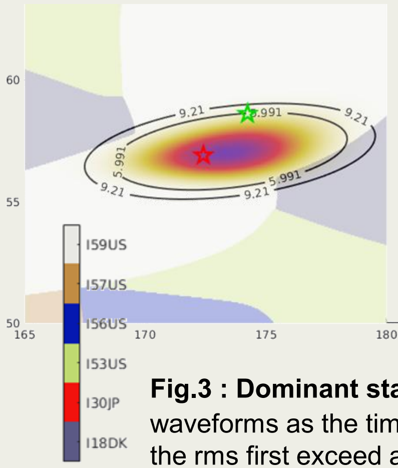



**Fig. 2 :** Signals; **W**, **S**, and **T** phases using an empirical model [2]. Eff. sound speed ratios: source-level and path-averaged profiles [3].

## On the role of each station in the localization: likelihoods & posteriors

Each IMS station contributes differently to the likelihood surface. By examining partial likelihoods $\mathrm{P}\left(t_j|\mathbf{x}\right)$, we identify which stations **dominate the localization**.




**Dominance of IS59 and its limitations:** posterior map shows that IS59 apparently dominates the localization, but using the Eff. Sound speed profile at source level (see **Fig. 2**), which overestimates the efficiency of ducting.

**Fig.3 : Dominant station** - $t_j$'s are extracted from waveforms as the times for which the amplitudes of the rms first exceed a fixed threshold of 30%.

## Future work

Accurate uncertainty estimation using MCMC will target both arrival time extraction and the impact of background atmospheric variability. The MCMC scheme should **adaptively downweight** stations whose contributions are uninformative.

[1] https://cneos.jpl.nasa.gov/reballs/
[2] Blom et al., Improved Bayesian Infrasonic Source Localization for regional infrasoundGJI, 2015.
[3] A. Goupy, A Spectral Approach of Multi-Scale Metamodelling Applied to Acoustic Propagation, PhD these, ENS Paris-Saclay, 2021.