Predicting infrasound transmission loss using deep learning

Quentin Brissaud, NORSAR, Kjeller, Norway

Sven *Peter* Näsholm, NORSAR & Department of Informatics, University of Oslo, Norway

Antoine Turquet, NORSAR, Kjeller, Norway

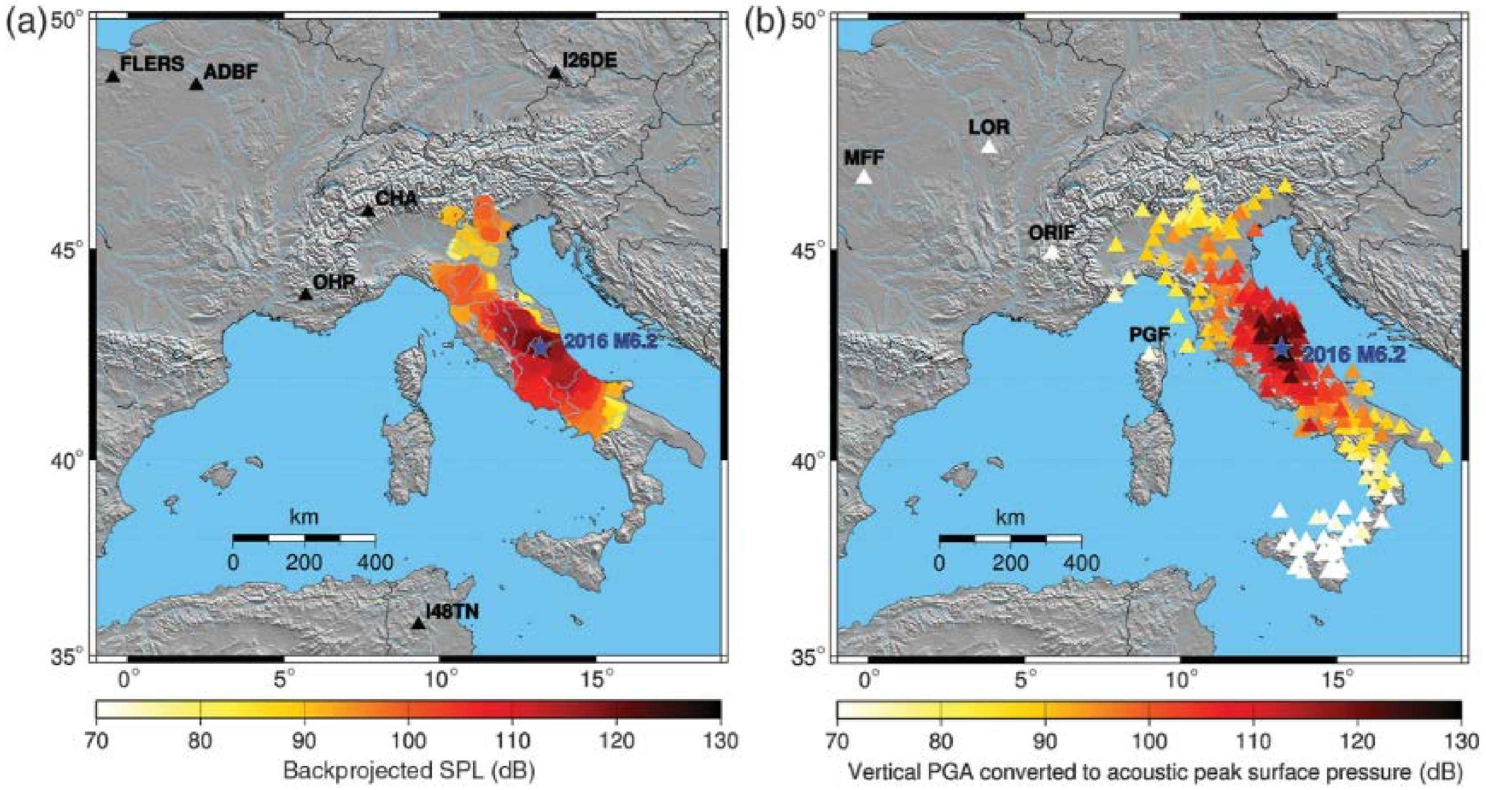
Alexis Le Pichon, CEA/DAM/DIF, Arpajon, France

March 22, 2022

Infrasound to retrieve source parameters

Accurate estimation of Transmission-Loss (TL), i.e., infrasound amplitude decay with distance
⇒ opportunity to complement seismic data with acoustic data for remote sensing of surface processes

Infrasound excited by surface sources can travel large distances and carry information about the source, e.g., surface pressure at the source after the 2016 Amatrice earthquake



Reconstructed & measured surface pressure

(left) Backprojected infrasound (SPL, dB)

(right) Acoustic peak surface pressure (PSP, in dB); triangle = seismic station.

Hernandez, B., Le Pichon, et al. (2018). Estimating the Ground-Motion Distribution of the 2016 M w 6.2 Amatrice, Italy, Earthquake Using Remote Infrasound Observations. *Seismological Research Letters*, *89*(6), 2227-2236.

Challenges with existing inversion framework

Full-waveform modeling: computationally expensive ⇒ inversions typically using empirical regression equations (Le Pichon, 2012, referred in the following as: LP12)

LP12, optimized over an idealized set of Parabolic Equation (PE) simulations, predicts TL as function of range

Source frequency

$$\mathrm{A}_P(f, V_{eff-ratio}) = \frac{1}{R}10^{\frac{\alpha(f)R}{20}} + \frac{R^{\beta\left(\boxed{f}\,\boxed{V_{eff-ratio}}\right)}}{1+10^{\frac{\delta-R}{\sigma(f)}}}$$

Effective velocity ratio @ 50 km altitude

Equation: LP 12 regression equation

⇒ Neglects vertically and horizontally varying wind profiles

Generating models allowing for fast TL estimation

2 main approaches to incorporate atmospheric variability for fast TL estimation

Analytical fitting approach

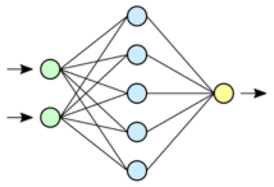
$$A_P(R, f, V_s) = A_0 R^{-\alpha(f,V_s)R}\, e^{-\beta(f,V_s)R}$$

Alexis Le Pichon, AGU21 →

Range-dependent analytical model

- Full control of predictive model parameters
- Explainability
- Simplicity
- Limited generalization for new data
- Difficult to introduce complexity in mapping function

Machine Learning (ML)



Machine learning

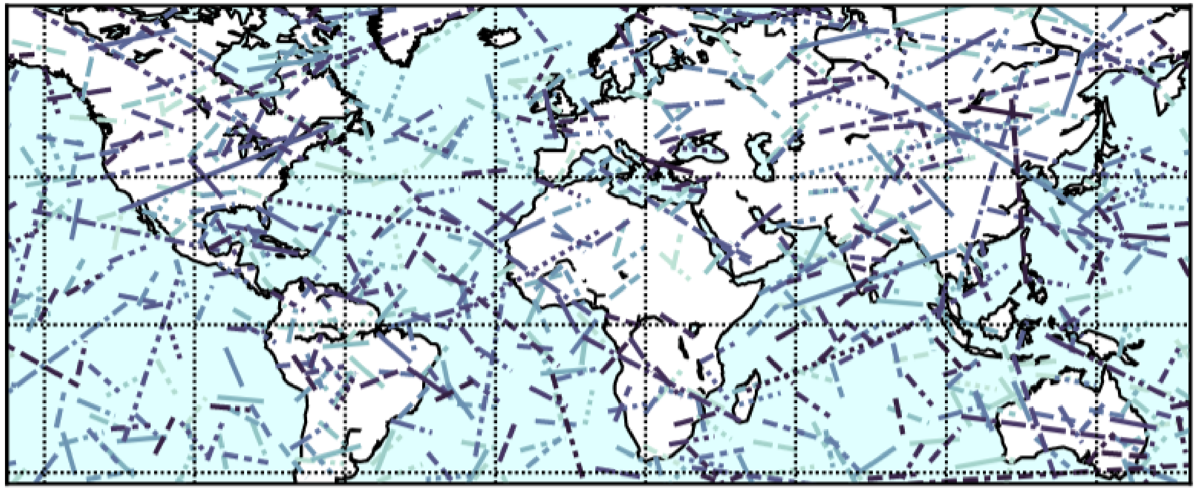
- Mapping with arbitrary complexity
- High accuracy
- "Black box"
- Costly training
- Tricky architecture optimization

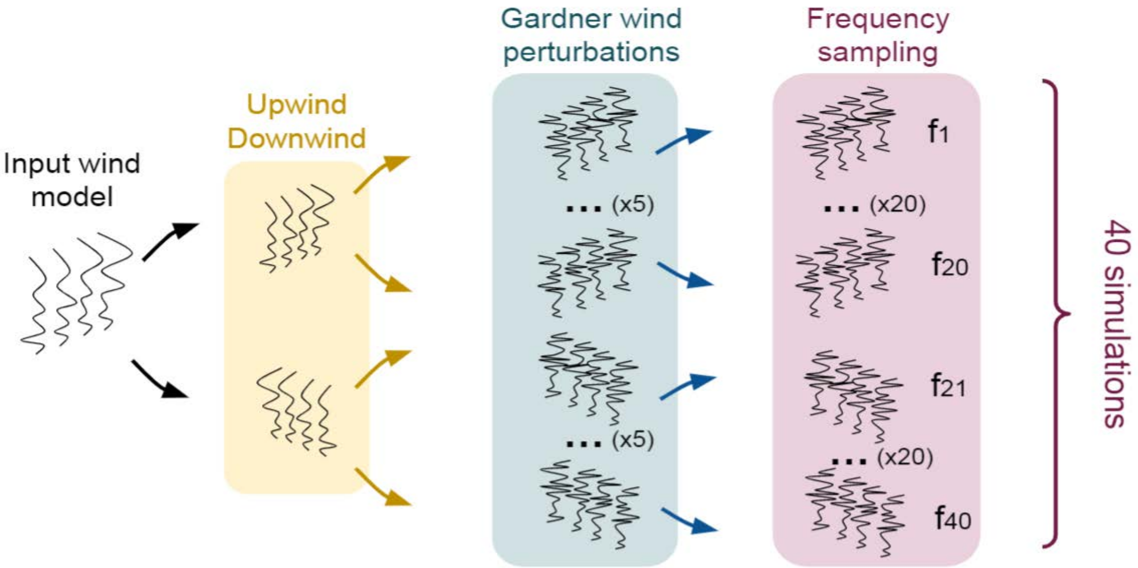
← **current approach**

Creating a “realistic” Transmission-Loss dataset

Accurate ML model requires training over a dataset representative of the variability in winds and TLs

- Synthetic simulations based on PE (NCPA ePape) ~25000 simulations
- Range-dependent (ERA5 & NRLMSIS-00/HWM13) & Gravity-wave perturbations
- Randomization:
    - Slice locations
    - Date/time (years, doy, hour)
    - Upwind/Downwind
    - Source frequency (0.1 – 3.2 Hz)





Learning TL from wind patterns using CNNs

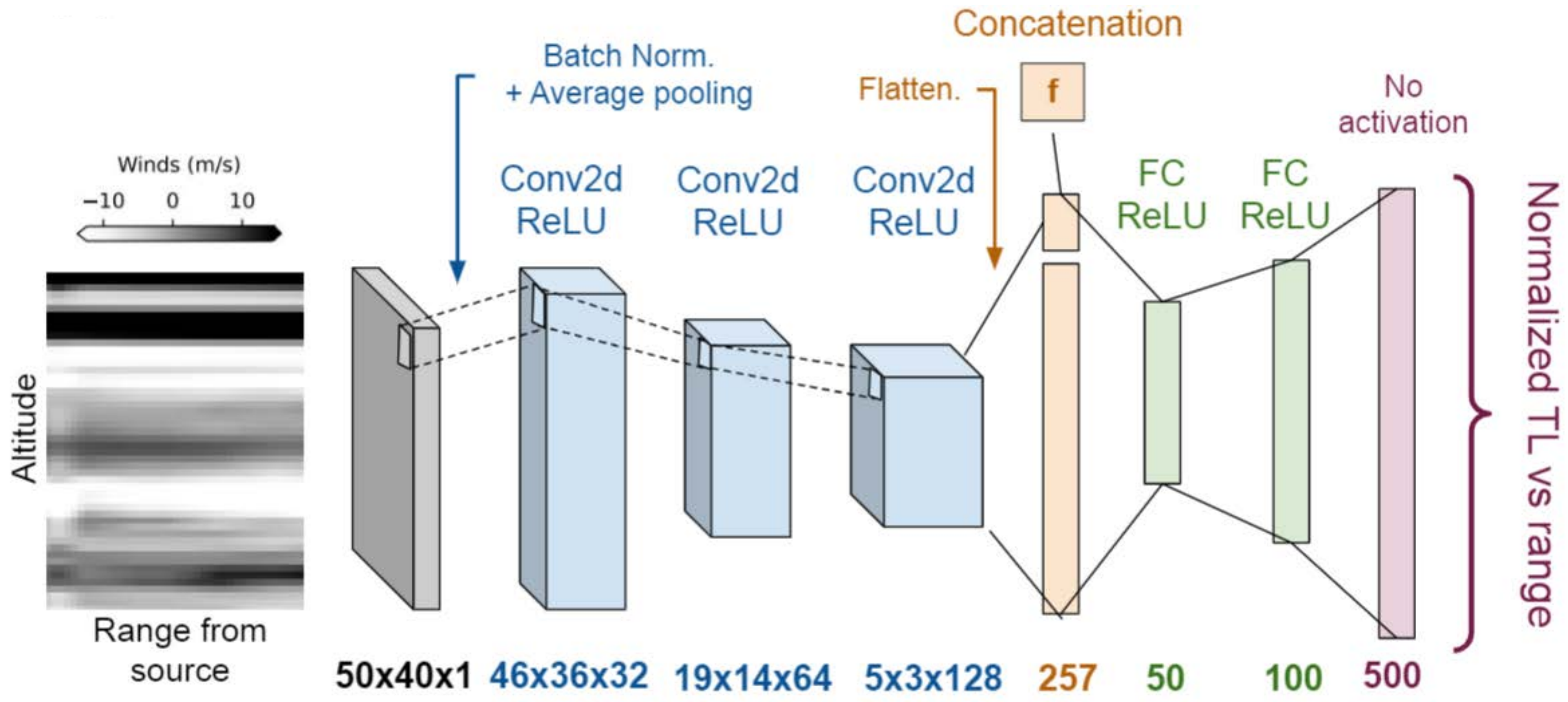
Small and large-scale wind variations + frequency control the acoustic wavefield structure at the ground

Our approach: [input] 2D wind maps ⇒ ground TL [output]

(1) extract wind patterns using 2D Convolutional Neural Networks (CNNs)

(2) find frequency-dependent TL relationship with wind models using a Fully-Connected layer

CNNs are employed to extract local and global patterns in multi-dimensional images with several layers of convolutions



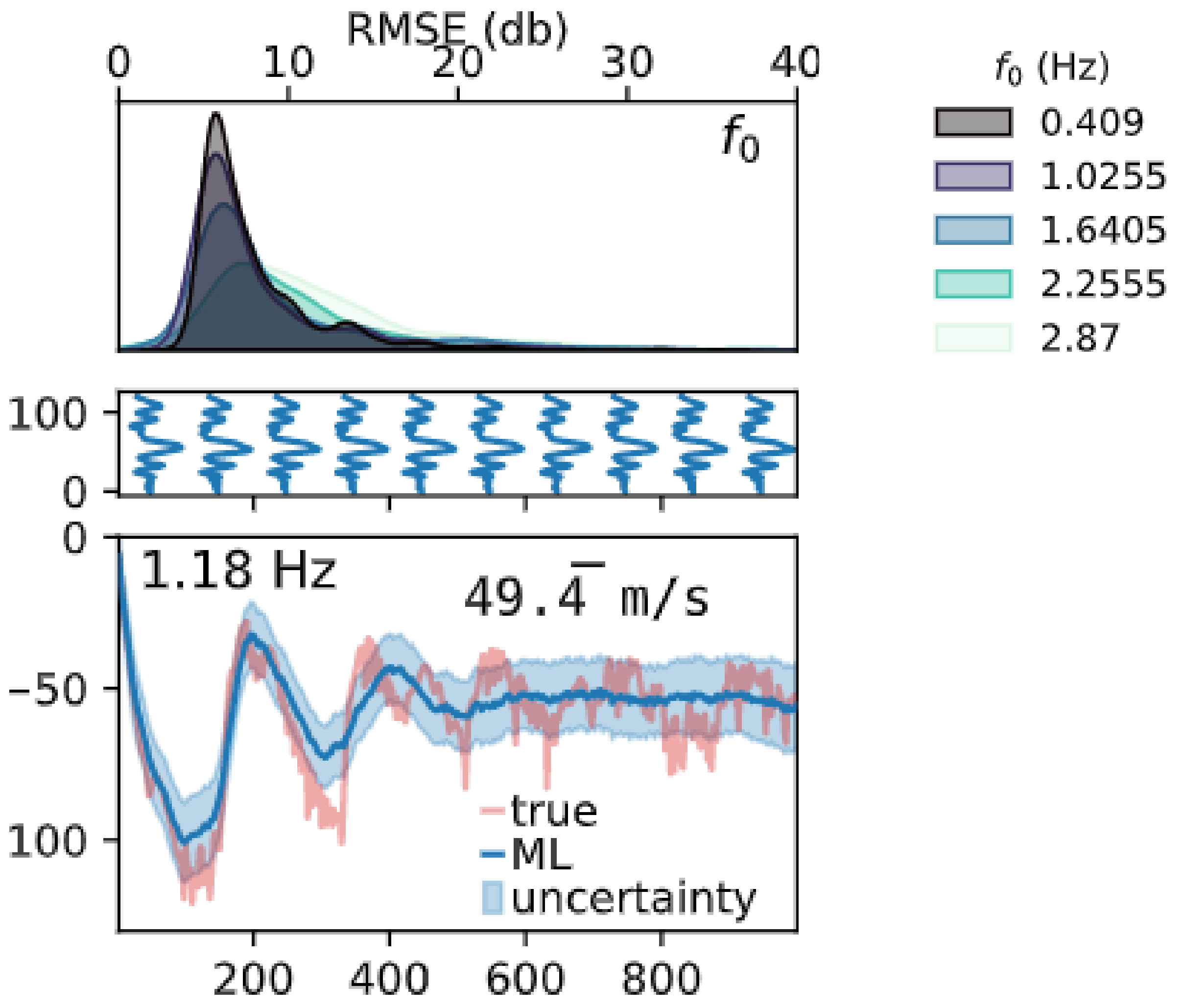
Training & validation

Training:

- Dataset split between Training (75%) & validation (25%)
- Mini-batch size 32
- Early stopping and reduce learning rate on plateau

Once trained

- 5 dB average accuracy over testing dataset
- ML-based TL estimate takes 0.05 s (vs. 10 to 150 s with PE simulations)
- Uncertainty estimates by computing error vs. range made by the ML model over the testing dataset
- Frequency and range independent errors
- Neglects “high-frequency” variations

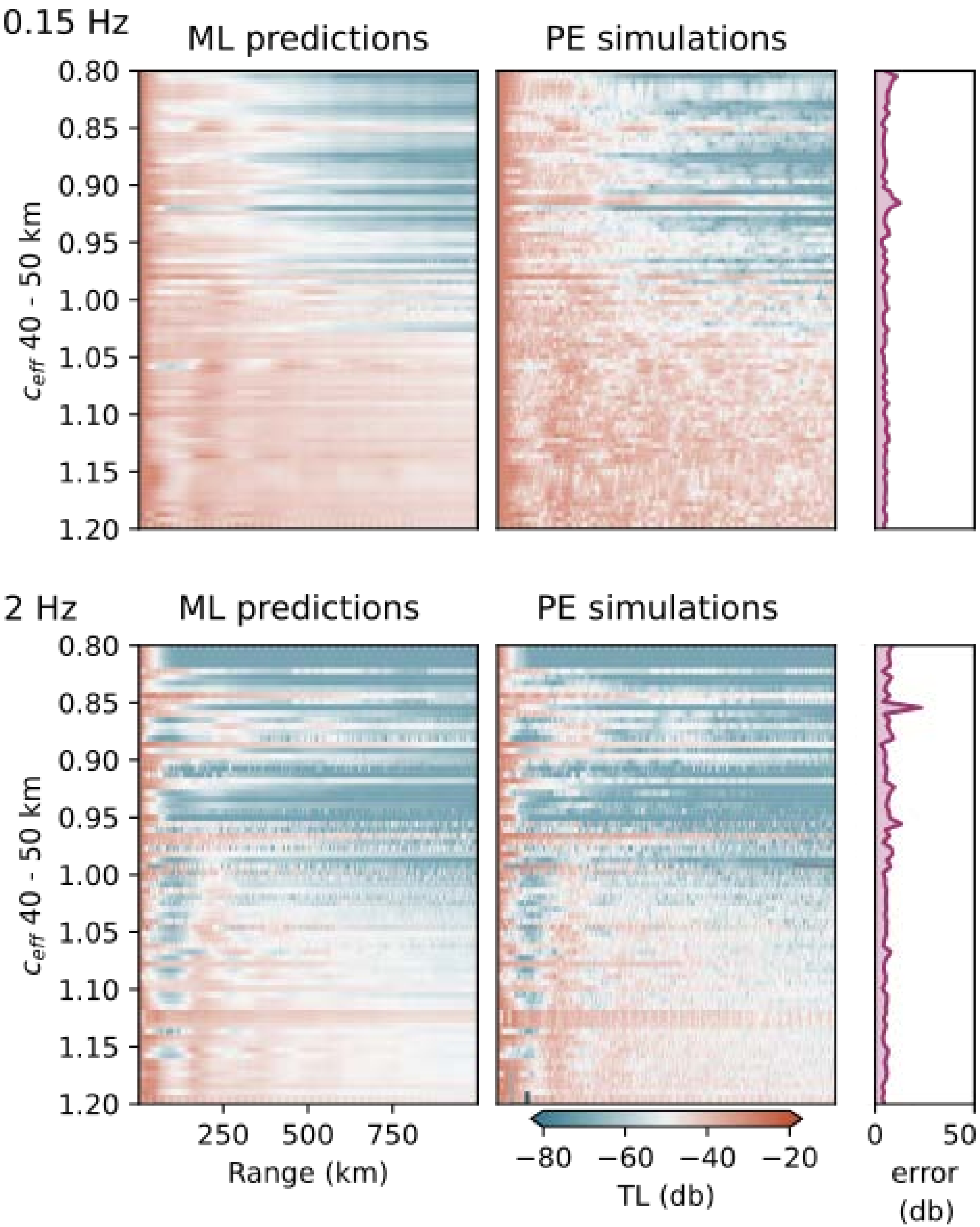


Resulting model

Effective velocity ratio maps show the variations of TL with variation in range from the source and "wind strength" in a given atmospheric layer (typically stratosphere)

ML captures the main features

- Multiple stratospheric shadow zones
- Tropospheric & thermospheric phases
- Low vs. high effective sound speed ratio



Perspectives

ML-based inexpensive (0.05 s) & accurate (around 5 dB) alternative to full simulations

Plenty of applications benefitting from rapid TL estimates: rapid atmospheric model inversions, (near-) realtime event characterization, microbarom modeling

- Rapid & efficient amplitude-based inversion procedures to retrieve source parameters (e.g., explosion yield, ground pressure levels)
- microbarom modeling requires greater propagation range (4 000 – 6 000 km) ⇒ build large-scale simulations to get new ground-truth & training ⇒ Currently being investigated by Edouard Forestier, intern at NORSAR

Future work:

Currently: range-independent Gardner perturbations ⇒ unrealistic beyond a few 100 km ⇒ Build dataset with range-dependent Gardner perturbations

Explainable ML, e.g., Layer-wise Relevance Propagation (LRP) ⇒ relationship between specific atmospheric model regions & TL ⇒ sensitivity kernels

Ground truth from more expensive & accurate codes (spectral-element / nonlinear propagation / …) ⇒ taking 3d effects into account (e.g., cross-winds)

Preprint available at: https://doi.org/10.1002/essoar.10509609.1

Code & synthetic data will be available at: https://github.com/QuentinBrissaud

Thank you!

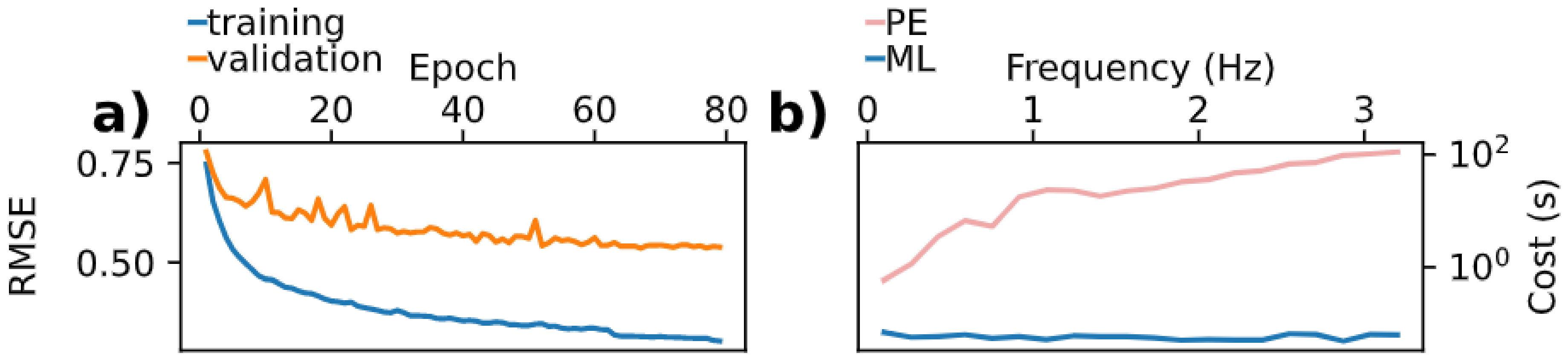
References

Le Pichon, A., Ceranna, L., & Vergoz, J. (2012). Incorporating numerical modeling into estimates of the detection capability of the IMS infrasound network. Journal of Geophysical Research: Atmospheres, 117(D5).

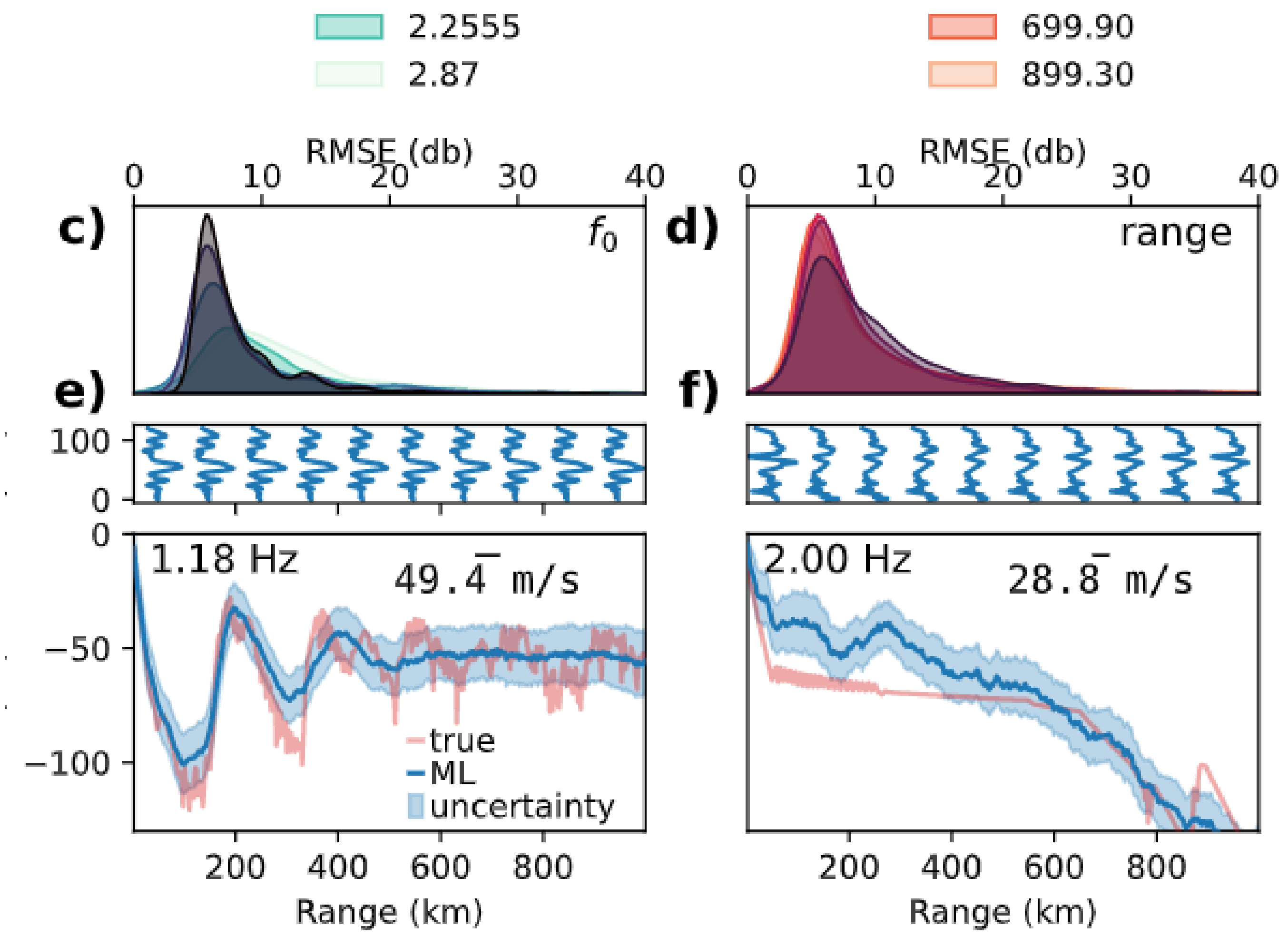
Hernandez, B., Le Pichon, A., Vergoz, J., Herry, P., Ceranna, L., Pilger, C., ... & Bossu, R. (2018). Estimating the Ground-Motion Distribution of the 2016 M w 6.2 Amatrice, Italy, Earthquake Using Remote Infrasound Observations. Seismological Research Letters, 89(6), 2227-2236.

Waxler, R., C. Hetzer, J. Assink, and D. Velea (2021), chetzer-ncpa/ncpaprop-release: Ncpaprop v2.1.0, doi:10.5281/zenodo.5562713, last accessed on 29 October 2021.

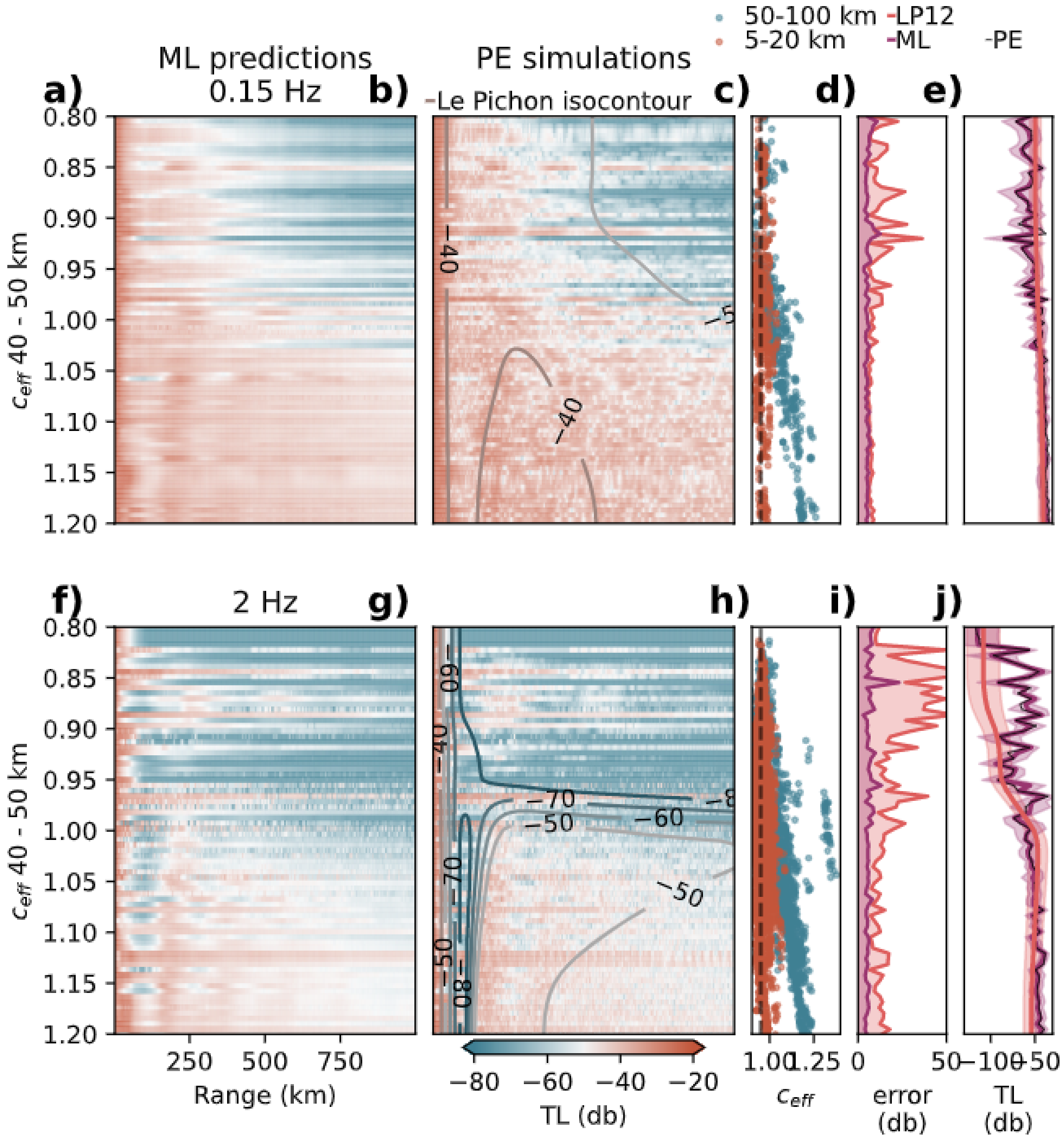
Training & computational time



Errors over test dataset



ML vs. LP12



LP12 reproduces the main features

- First stratospheric shadow zone
- Low vs. high effective sound speed ratio

Creating a realistic Transmission-Loss dataset

Accurate ML model requires training over a dataset representative of the variability in winds and TLs

- Similar to LP12:
  generate synthetic dataset from PE simulations (NCPA ePape)
- Atmospheric range-dependent models:
  ERA5 & NRLMSIS-00/HWM13
- Randomly sample:
  - Slice locations
  - Year
  - Day