Towards Bayesian Full-Waveform Source Inversion using Simulation-Based Inference

A. A. Saoulis[1,2], A. M. G. Ferreira[2], B. Joachimi[1], A. Spurio Mancini[1,3], D. Piras[4]

[1]Physics & Astronomy, University College London, [2]Earth Sciences, University College London,
[3]Mullard Space Science Laboratory, University College London, [4]University of Geneva

1. Motivation – Improving Bayesian Inversion

Given a forward model $f(\boldsymbol{\theta}) \rightarrow \boldsymbol{D}$ over parameters of interest $\boldsymbol{\theta}$ (e.g. source location):

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{D})}_{\textit{Posterior}} \propto \underbrace{p(\boldsymbol{D}|\boldsymbol{\theta})}_{\textit{Likelihood}} \times \underbrace{p(\boldsymbol{\theta})}_{\textit{Prior}}$$

- **Gaussian** likelihood function is often used to perform Bayesian inversions on seismic data.
- This methodology can introduce **bias** in the presence of non-Gaussian noise.

2. Simulation-Based Inference (SBI)

- **Learned** likelihood function replaces user-specified (e.g. Gaussian) likelihood function.
- This likelihood function is modelled by a Machine Learning (ML) model known as a Neural Density Estimator (NDE), which is trained on simulated data.

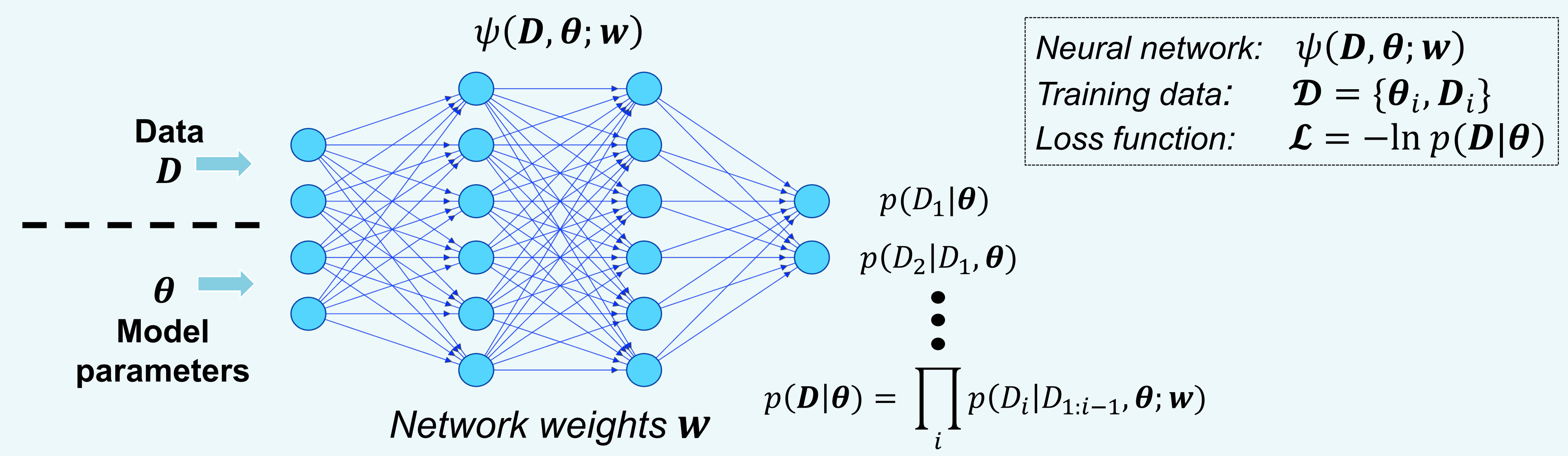


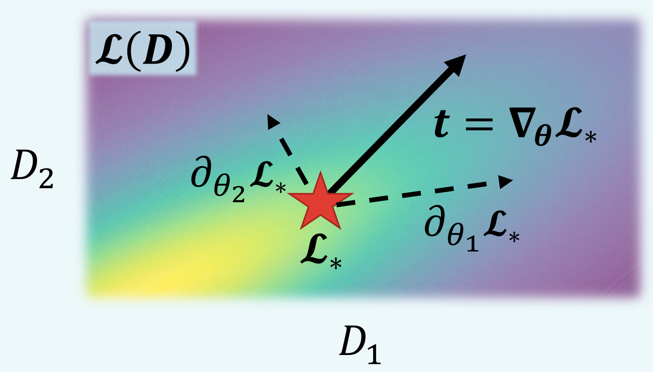
Fig. 1: Overview of a **M**asked **A**utoencoder for **D**ensity **E**stimation (**MADE**) [1], a class of NDE that has shown great success in modeling probability densities by ensuring its output is a normalised probability distribution.

3. Data Compression

- The very high dimensionality of the displacement data $\boldsymbol{D}$, which represents full-waveforms at each station, makes training a NDE **infeasible**.
- Data $\boldsymbol{D}$ must be compressed to $\dim(\boldsymbol{\theta})$ summary statistics $\boldsymbol{t}$, replacing $\boldsymbol{D}$ in *Fig. 1*.

This study investigates two compression methods ϕ for seismic data:

Optimal Score Compression
Expand about likelihood $\mathcal{L}_*$



ML-Based Compression

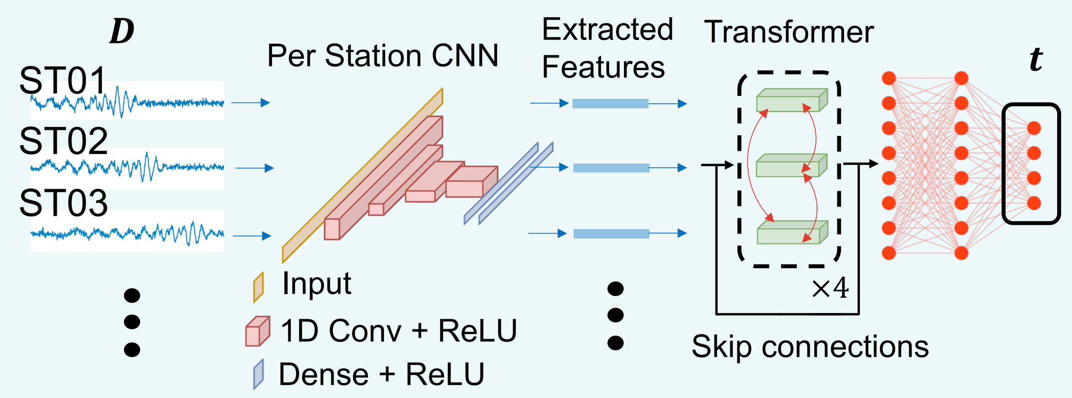


Fig. 2: Left: Classical technique for optimal compression using a first-order Taylor expansion, see [2]. *Right*: Neural network architecture, inspired by [3], trained to perform compression by learning to map $\phi : \boldsymbol{D} \rightarrow \boldsymbol{\theta}$.

4. Example Simulation-Based Study

- We demonstrate SBI on a simplified problem of source-location inversion, i.e. $\boldsymbol{\theta} = \{x, y, z, \Delta t\}$, a four parameter source inversion.
- Study region of the 2021 São Jorge crisis in the Azores, using the **UPFLOW 2021-22** [4] Ocean Bottom Seismometer (OBS) array.
- Entirely synthetic study, simulating events using the isotropic 1-D model PREM [5]. The simulations are accurate down to wave periods of $T{\sim}2$ s.

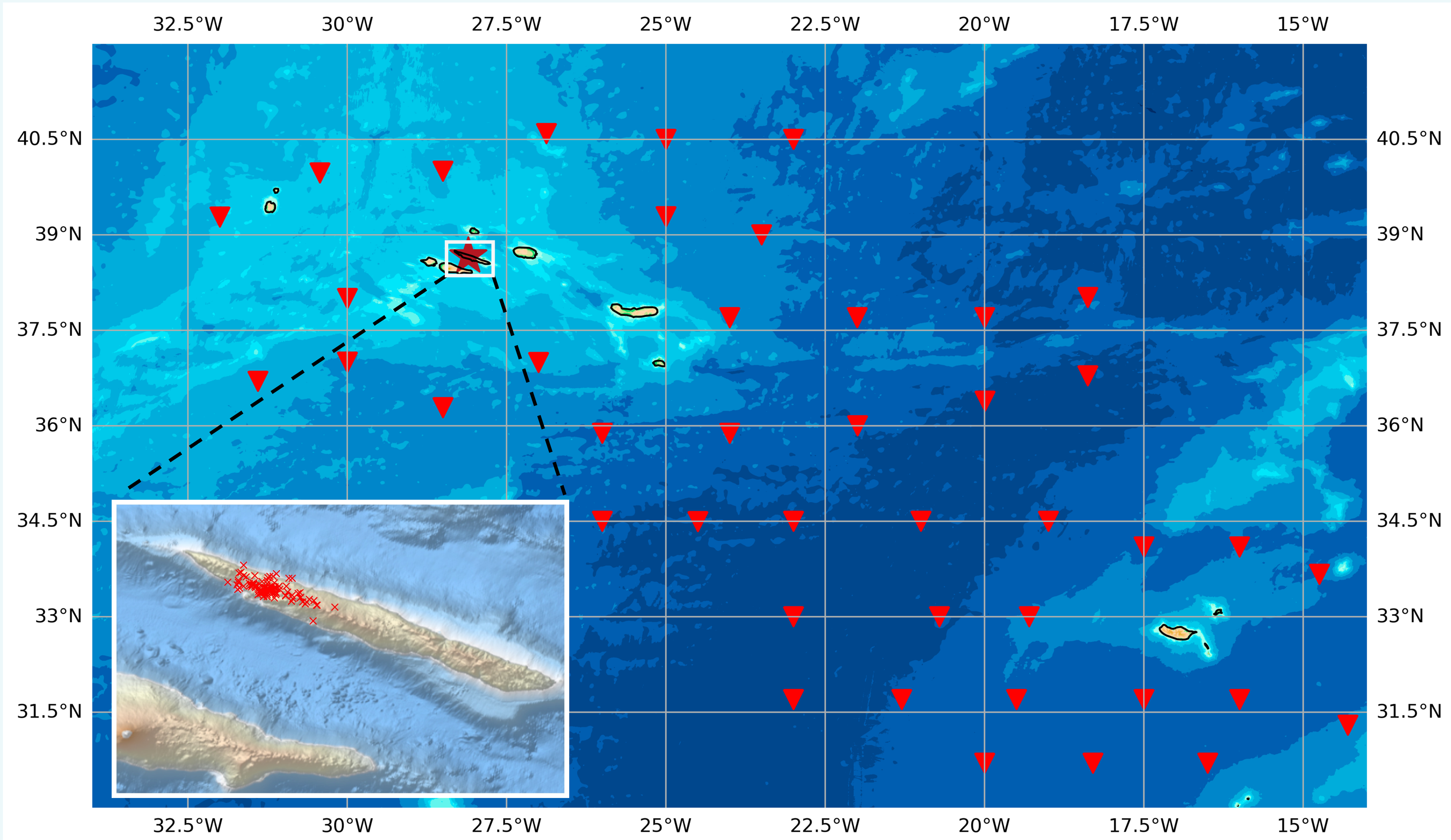


Fig. 3: A schematic showing the experimental setup for this simulation-based study. We use OBS stations from the UPFLOW array (*red triangular markers*) to study the recent seismic swarm crisis located on São Jorge island (*brown star*). *Bottom left*: zoom in on the São Jorge crisis, with a small selection of events from the *IPMA* [6] catalogue.

5. Simulation-Based Inference Workflow

Generate a Compressed Dataset

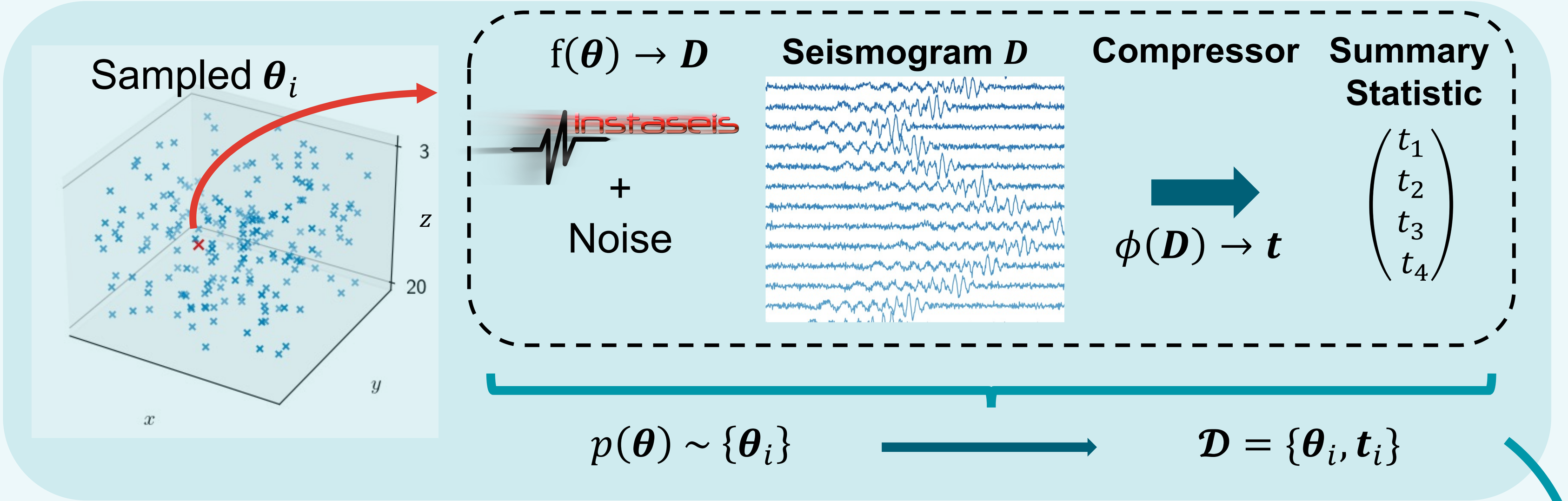


Fig 4. Events simulated with *Instaseis* [7]. Events are centred on the seismic swarm, sampled within a cube $(x, y, z) = 3\,\text{km} \times 3\,\text{km} \times 17\,\text{km}$, between depths $3 - 20$ km. The time shift Δt is sampled between $[-1,1]$ s. Synthetic noise is added to each waveform before compression.

Train NDE to Estimate Likelihood

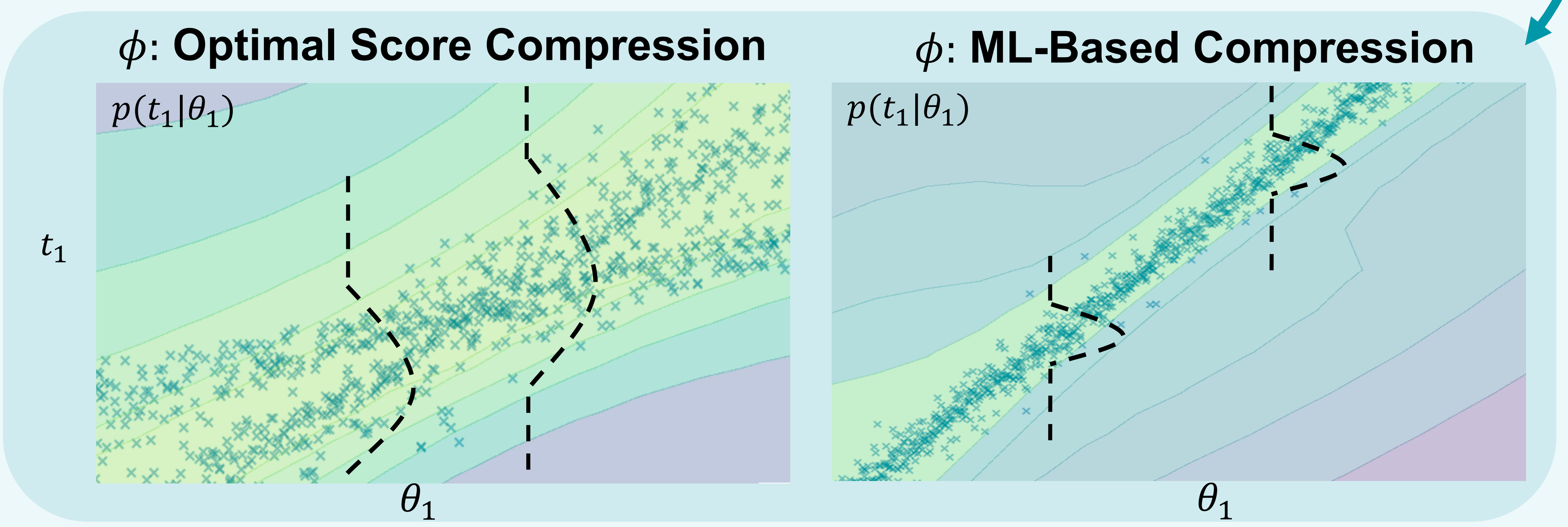


Fig 5. NDE likelihood contours for each compression method. Score compression performs poorly due to displacement $\boldsymbol{D}$ non-linearity with respect to location parameters $\boldsymbol{\theta}$. ML-based compression gives sharper likelihood contours, corresponding to less lossy compression.

6. Inversion Results

Samples from the posterior $p(\boldsymbol{\theta}\,|\boldsymbol{t}_{\text{obs}})$ are drawn using MCMC once the likelihood NDE is trained.

	Compression	
	Score	ML
Compression MSE ↓	$2.6{\times}10^{-2}$	$\mathbf{8.2{\times}10^{-4}}$
Posterior CRPS* ↓	$2.6{\times}10^{-2}$	$\mathbf{1.3{\times}10^{-2}}$
Calibration Error ↓	0.35	0.31

Table 1. ML-based method improves compression and yields tighter posteriors (lower CRPS). Both methods are relatively well calibrated.
**Continuous Ranked Probability Score*

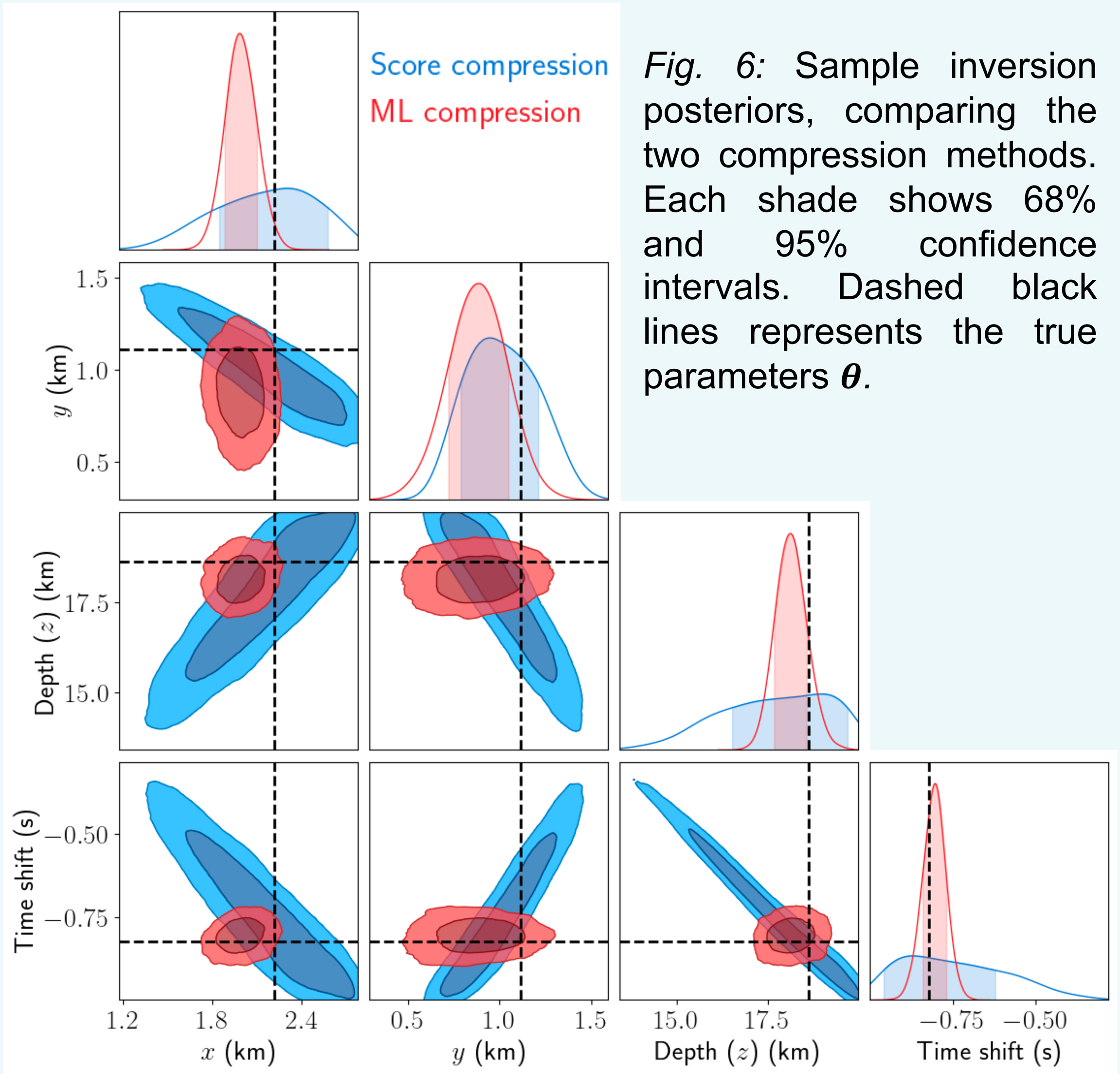


Fig. 6: Sample inversion posteriors, comparing the two compression methods. Each shade shows 68% and 95% confidence intervals. Dashed black lines represents the true parameters $\boldsymbol{\theta}$.

7. Conclusions

- SBI can incorporate full-waveform information and account for non-Gaussian noise effects in the posterior distribution. Future work will address and quantify these advantages.
- Samples are generated in the compressed space $\{t, \theta\}$, foregoing the forward model and giving $\sim 60\times$ speed-up over MCMC using *Instaseis.*
- The choice and tuning of the compression technique is important.
- More work is needed to avoid failure in the presence of modelling errors.

References:
[1] Germain et al. (2015), *CoRR* - **[2]** Alsing & Wandelt (2018), *MNRAS* - **[3]** Münchmeyer et al. (2021), *GJI* - **[4]** UPFLOW project, *upflow-eu.github.io*
[5] Dziewonski & Anderson (1981), *Phys. Earth Planet. Inter.* - **[6]** IPMA, *www.ipma.pt* - **[7]** van Driel et al. (2015), *Solid Earth.*

Correspondence:
a.saoulis@ucl.ac.uk

EGU23-7939
This presentation participates in OSPP
Outstanding Student & PhD candidate Presentation contest