Seismic hazard analysis with a Fourier Neural Operator (FNO) surrogate model enhanced by transfer learning

Fanny Lehmann[*]
CEA/DAM/DIF, Arpajon, France
`fanny.lehmann@centralesupelec.fr`

Filippo Gatti
LMPS[*], Gif-sur-Yvette, France
`filippo.gatti@centralesupelec.fr`

Michaël Bertin
CEA/DAM/DIF, Arpajon, France
`michael.bertin@cea.fr`

Didier Clouteau
LMPS[*], Gif-sur-Yvette, France
`didier.clouteau@centralesupelec.fr`

Abstract

Seismic hazard analyses in the area of a nuclear installation must account for a large number of uncertainties, including limited geological knowledge. It is known that some geological features can create site-effects that considerably amplify ground motion. Combining the accuracy of physics-based simulations with the expressivity of deep neural networks can help quantifying the influence of geological heterogeneities on surface ground motion. This work demonstrates the use of a Factorized Fourier Neural Operator (F-FNO) that learns the relationship between 3D heterogeneous geologies and time-dependent surface wavefields. The F-FNO was pretrained on the generic HEMEW-3D database made of $30\,000$ samples. Then, a smaller database was built specifically for the region of the Le Teil earthquake (South-Eastern France) and the F-FNO was further trained with only 250 specific samples. Transfer learning improved the prediction error by $22\,\%$. As quantified by the Goodness-Of-Fit (GOF) criteria, $90\,\%$ of predictions had excellent phase GOF ($62\,\%$ for the enveloppe GOF). Although the intensity measures of surface ground motion were, in average, slightly underestimated by the FNO, considering a set of heterogeneous geologies always led to ground motion intensities larger than those obtained from a single homogeneous geology. These results suggest that neural operators are an efficient tool to quantify the range of ground motions that a nuclear installation could face in the presence of geological uncertainties. The HEMEW-3D database and the pretrained F-FNO model are publicly available to facilitate further developments and applications.

1 Introduction

The Le Teil earthquake (M_w4.9, 2019) is one of the major earthquakes that hit metropolitan France in the last decade [1]. Its occurrence in the Rhône valley, a dense industrial area with low-to-moderate seismicity, renewed the interest for seismic hazard analyses in this region, especially for nuclear installations. Seismic hazard analyses estimate the intensity of surface ground motion, depending on the regional seismicity and the geological context. However, geological features are generally poorly characterized, leading to large uncertainties in the range of ground motions that may impact the nuclear installation [2].

---

[*]LMPS - Laboratoire de Mécanique Paris-Saclay, Université Paris-Saclay, CentraleSupélec, ENS Paris-Saclay, CNRS

To quantify the impacts of geological uncertainties, the Monte-Carlo approach consists in running thousands of analyses with varying geologies. This method becomes unfeasible if the analyses are computationally demanding, which is the case when trying to reach high frequencies and high vertical resolution. Of special interest are three-dimensional (3D) simulations that use finite elements, finite differences, or spectral element methods to solve the wave propagation equation. Those numerical codes compute the propagation of seismic waves from the earthquake source to the surface where ground motion is synthetized by virtual sensors.

Representing 3D geological domains is crucial to ensure that site-effects are correctly accounted for in the seismic hazard analyses. Site-effects include sedimentary basins or topographic effects that can greatly amplify seismic waves and lead to extreme damage [3, 4]. To model site-effects, 3D numerical simulations are necessary [5–7] but computationally demanding. In fact, high-frequency (up to 10Hz) 3D simulations commonly amount to several tens of thousands of equivalent CPU (Computational Processing Unit) hours [8–10].

Therefore, there is a need to elaborate surrogate models that are precise enough to represent 3D geological domains and whose computational cost allows advanced uncertainty quantification. Several methods already exist to obtain ground motion features but they remain limited. Ground Motion Prediction Equations (GMPEs) or Ground Motion Models (GMMs) predict ground motion intensity measures, such as the Peak Ground Acceleration (PGA), by learning the relationship between earthquake parameters (e.g. depth, magnitude), geological conditions at the sensor position (e.g. $V_{S,30}$, the average velocity of shear waves to a depth of 30 meters), and recorded earthquake intensity measures [11]. Their applicability depends on the number of earthquakes recorded in contexts similar to the nuclear site under study. To alleviate this issue, Gaussian process regression models [12], polynomial chaos expansion [13], or orthogonal decomposition [14], among others, use numerical simulations to create synthetic databases. However, they generally model 1D or 2D geological domains [12, 13] and predict only some features of ground motion [12–14], not the entire timeseries.

The recent advances of artificial intelligence (AI) now make time- and space-dependent prediction of ground motion within reach [15, 16]. Convolutional Neural Networks (CNNs, [17]) and Physics Informed Neural Networks (PINNs, [18]) have shown great success to solve the 2D wave equation [19–21]. However, these models need to be retrained from scratch each time the configuration (e.g. geological domain) changes. Neural operators offer new perspectives to AI by designing deep learning models that are trained for a whole family of configurations [22, 23]. Indeed, neural operators learn the mapping between a functional space of geological configurations and the functional space of time- and space-dependent ground motion.

Among the numerous neural operators, the Fourier Neural Operator (FNO, [24]) has shown promising results, and one of its variants has been successfully applied to predict 3D seismic wave propagation [25]. Although geologies are inherently three-dimensional, predictions of the neural operators also depend on three coordinates: two spatial coordinates to describe the sensor position at the surface of the domain, and the temporal coordinate. Dealing with 3D inputs and outputs leads to several challenges for neural operators, in terms of memory requirements, model complexity, and data management [26]. 3D variables are rarely tackled by neural operators, and this work pioneers in this direction.

This work uses the recently introduced Factorized Fourier Neural Operator (F-FNO, [27]) to predict ground motion on a geological database specific to the Le Teil earthquake (Section 2). Due to the cost of acquiring datasets and training models, transfer learning is used to benefit from a previous training of the F-FNO on a generic database. Then, the predictions of surface wavefields are assessed using seismological metrics and the influence of geological uncertainties is quantified (Section 3).

2 Methods

2.1 The HEMEW-3D Database for training

The HEMEW-3D database contains $30\,000$ geologies with their associated velocity wavefields[2]. Each geology is a cubic domain of size $9.6\,\mathrm{km} \times 9.6\,\mathrm{km} \times 9.6\,\mathrm{km}$ representing shear wave (S-wave) velocity between $1071\,\mathrm{m/s}$ and $4500\,\mathrm{m/s}$. This dataset is meant to be as generic as possible to enable

---

[2]`https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/LAI6YU`

diverse applications. Therefore, geologies are not related to any known geological context. Instead, each geology contains horizontal layers of random thickness; and heterogeneities are added in each layer to create varying velocity values (Figure 1). It is important to note that geologies exhibit large heterogeneities (the mean coefficient of variation of heterogeneities being 20%).

A velocity wavefield is associated with each geology. It was created by propagating seismic waves originating from the seismic source of the Le Teil earthquake, described as a moment tensor of strike $= 48°$, dip $= 45°$, and rake $= 88°$, and seismic moment $M_0 = 2.47 \times 10^{16} \, \mathrm{N\,m}$, [28]. However, in the HEMEW-3D database, the source is located much deeper ($-7.8 \, \mathrm{km}$) than the Le Teil earthquake to allow a longer propagation of seismic waves inside the domain. The 3-component velocity wavefield is synthetized at a grid of 16×16 virtual sensors located at the surface of the domain, between $1 \, \mathrm{s}$ and $7.4 \, \mathrm{s}$ with a $0.02 \, \mathrm{s}$ sampling interval. Simulations of wave propagation were conducted with the earthquake simulator SEM3D[3] [10]. They are valid up to a $5 \, \mathrm{Hz}$ frequency, with numerical pollution appearing at higher frequency. Although the geological domains are intrinsically 3D (depending on the spatial coordinates x, y, and z), it is important to note that the ground motion velocity wavefields also depend on three variables: space x and y, and time t.

2.2 Factorized Fourier Neural Operator (F-FNO)

The core principle of the FNO is to view convolution operators as a multiplication of Fourier coefficients [24]. Each Fourier layer ℓ decomposes its inputs v_ℓ with the Fast Fourier Transform (FFT, $\mathcal{F}$) and multiplies the obtained Fourier coefficients with the FFT of the layer's weights R_ℓ:

$$v_{\ell+1} = \sigma \left(\mathcal{F}^{-1} \left(R_\ell \cdot \mathcal{F}(v_\ell) \right) + W_\ell v_\ell \right)$$

The computation is done efficiently by learning the Fourier layer's weights in the Fourier space. The Fourier weights contain 16 Fourier modes in the first and second dimensions, and 32 modes in the third dimension (excepting the first layer which has 16 modes).

The F-FNO replaces the 3D FFT in the Fourier layers by the sum of 1D FFT on each dimension, which significantly reduces the number of parameters [27]. Tran et al. [27] showed that the F-FNO error reduces when the number of F-Fourier layers is increased, which does not hold true for the original FNO. In this work, the F-FNO is implemented with 20 layers. Before the 20 factorized Fourier layers, the F-FNO includes an uplift sub-network that converts single-channel material inputs into 16-channel variables v_0 via two dense layers. Finally, three subnetworks project the 16-channel variable v_{20} onto 1-channel physical variables corresponding to the surface velocity. Each projection sub-network is specialized for one of the three wavefield components (East-West, North-South, Vertical) and comprises two dense layers.

The HEMEW-3D database was split in 27,000 training and 3,000 validation samples. The Adam optimizer was used with a learning rate of $6 \cdot 10^{-4}$ halved on plateau, a batch size of 16 and the relative mean absolute error as the loss function. The F-FNO was trained for 350 epochs, which took 24h on 4 Nvidia A100 GPUs.

2.3 The Le Teil database for transfer learning

Since compiling a large database of physics-based simulations such as the HEMEW-3D database represents 1.6 million of CPU hours, it is not desirable to build new large databases when one wants to study a target region. With transfer learning, one benefits from a pre-trained model and continues training on a smaller, tailored database [29].

Heterogeneities were added to the reference geology to obtain a set of 4000 plausible geologies in the region. Heterogeneities are represented by random fields with the same parameters as the HEMEW-3D database: correlation lengths between $1.5 \, \mathrm{km}$ and $6 \, \mathrm{km}$, coefficients of variation normally distributed around $20 \, \%$ with a standard deviation of $10 \, \%$, and a von Karman correlation model [2]. To enable an easier interpretation of results, all layers in the Le Teil database share the same parameters, which is not the case for the HEMEW-3D database (it does not mean that heterogeneities are located at the same location in all layers, they only have similar properties). Figure 2 shows the vertical distribution of V_S. Velocity wavefields were obtained with the exact same method as the HEMEW-3D database.
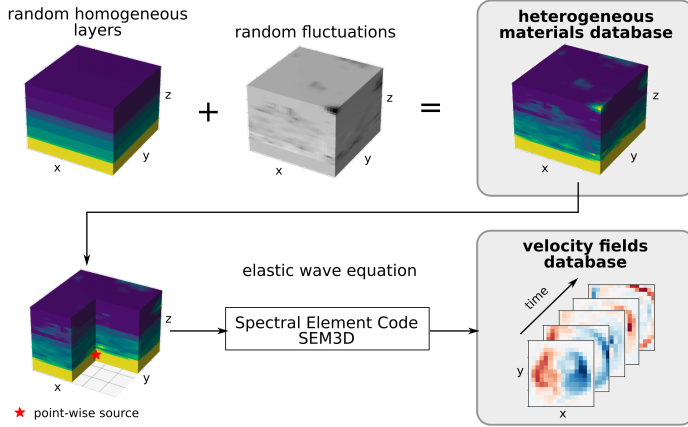
---

[3] https://github.com/sem3d/SEM

Figure 1: The HEMEW-3D database contains geological materials with their associated ground motion wavefields obtained after the propagation of seismic waves from a fixed seismic source up to the surface.
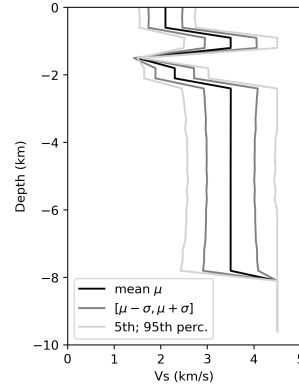
Figure 2: V_S distribution in the Le Teil database. For each of the 32 vertical levels, the mean μ, standard deviation σ, and percentiles are computed over all horizontal points and samples.

Inspired from the 1D velocity model used to reproduce the Le Teil earthquake [30], we designed a database of 4000 geologies based on this model. The *reference geology* is a 3D domain with six homogeneous horizontal layers and the S-wave velocities described in Table 1. The ratio of P- and S-waves velocities (V_P and V_S) is assumed constant and equal to 1.7 Density and attenuation factors are computed from V_P and V_S [31].

Thickness (m)	V_S (m/s)
600	2100
600	3500
300	1200
600	2300
5700	3500
1800	4500

Table 1: Reference 1D geological model for the LeTeil region. S-wave velocity (V_S) and thickness of each horizontal layer (adapted from [30]).

Among the 4000 pairs of geologies and velocity wavefields, up to 3000 were used for training, 300 for validation, and 700 for testing. With transfer learning, the neural operator's weights are initialized with the weights of the F-FNO trained on the HEMEW-3D database. Then, the model is trained solely on samples from the Le Teil database, forcing it to specialize to the Le Teil context, and hopefully improve its predictions. The number of samples used for the second training is called the number of *transfer learning samples* and is denoted N_{TL}.

3   Results

3.1   Transfer learning advantages

Figure 3 illustrates the benefits of transfer learning, even with a small number of specific training samples. For comparison purposes, the F-FNO was firstly trained without pre-training, meaning that the weights were randomly initialized. At the end of the training with 3000 samples from the Le Teil database, the rMAE was $40.4\,\%$ (grey dashed line). With transfer learning starting from weights initialized with the pretrained F-FNO, the rMAE reduced to $31.3\,\%$ with only 250 transfer learning samples (blue dashed line).

Although the rMAE is pertinent to train the neural operator, it is not informative about the physical accuracy of the trained F-FNO. The Goodness-Of-Fit (GOF) criteria give a better assesment of the F-FNO accuracy. The GOFs were computed at each spatial location and for each testing sample. They
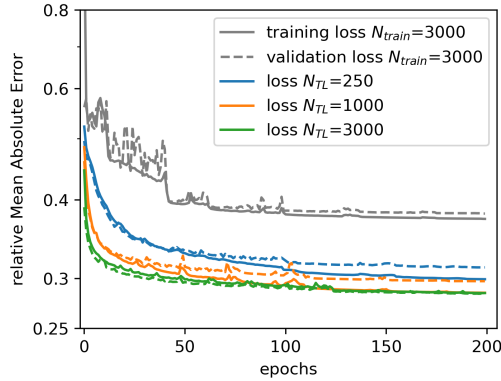
Figure 3: Evolution of the training loss (solid lines) and validation loss (dashed lines) computed as the relative Mean Absolute Error when the F-FNO is trained on the Le Teil database. Four trainings are shown: random initialization with no pre-training (grey lines), transfer learning with 250 samples (blue lines), 1000 samples (orange lines), 3000 samples (green lines).
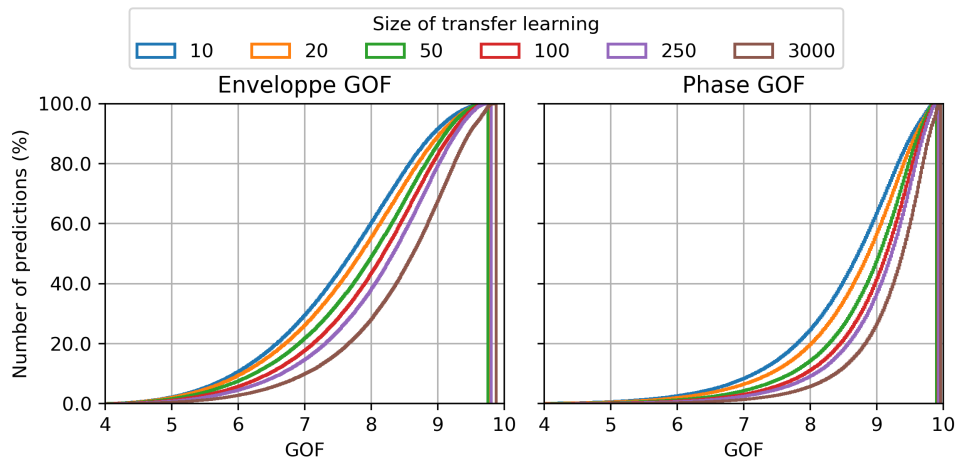

Figure 4: Cumulative distribution of Goodness-Of-Fit (GOF) for enveloppe (left) and phase (right), as a function of the number of transfer learning samples (colored lines). Results are shown for 700 testing samples.

evaluate the agreement between the predicted timeseries and the simulated timeseries considered as *ground truth*, in terms of enveloppe and phase [32]. GOFs are reported on a scale varying from 0 to 10, where 10 means a perfect agreement, and 8 is generally considered as excellent.

Figure 4 shows the cumulative distribution of enveloppe GOF and phase GOF when the number of transfer learning samples increases. One can firstly notice that the distributions are more shifted towards the right for the phase GOF than the enveloppe GOF, meaning that the F-FNO is better at predicting wave arrival times than perfectly accurate amplitudes. With 250 transfer learning samples, 91% of predictions have a phase GOF larger than 8. This is very close to the performance achieved with 3000 transfer learning samples (94% of predictions having a phase GOF larger than 8). Concerning the enveloppe, 62% of predictions achieve a GOF larger than 8 with 250 transfer learning samples. Using less than 250 samples from the Le Teil database is possible but degrades the F-FNO accuracy (43% of predictions have an enveloppe GOF larger than 8 when only 20 transfer learning samples are used).
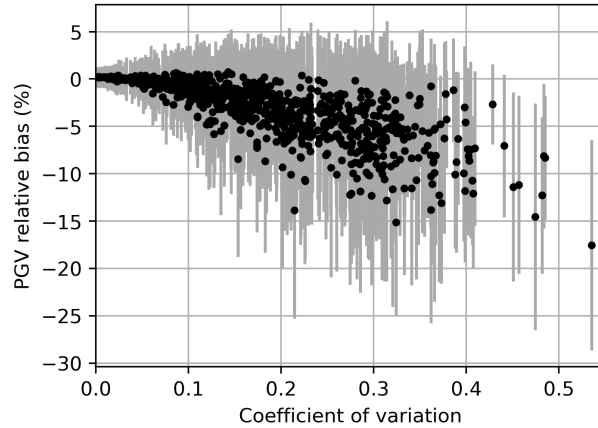
Figure 5: Each test sample is represented by the geological coefficient of variation (x axis). For each sample, the PGV bias is given as the mean of all sensors (black dot) and the standard deviation across samples (grey error bar).

3.2 Prediction accuracy

The enveloppe GOFs indicate that it is challenging to predict timeseries with very accurate amplitudes for the entire duration of the signal. Since the amplitude of ground motion is correlated with the intensity of shaking, it is crucial to ensure that our F-FNO gives reliable estimates of ground motion. The Peak Ground Velocity (PGV) was computed as the maximum value over time of the geometric mean of the two horizontal components. Since all timeseries are synthetic, note that the PGV corresponds to a maximal frequency of $5\,\mathrm{Hz}$. For each sample and each spatial point, the PGV accuracy was defined as the relative bias between the prediction and the simulated ground truth.

Figure 5 shows the PGV relative bias as a function of the heterogeneities' amplitude for the associated geology. One can notice that predictions are excellent when heterogeneities have a low amplitude. Indeed, the PGV relative bias is lower than 3% for coefficients of variation smaller than 0.05. Then, in average, the F-FNO tends to slightly underestimate the PGV. Indeed, the mean PGV relative bias across sensors is between $1\,\%$ and $-4\,\%$ for coefficients of variation around 0.1 When heterogeneities have large amplitudes (coefficients of variation larger than 0.2), the underestimation generally worsens. However, when one considers a sufficiently large number of samples, there is no spatial point where the PGV is consistently underestimated. Therefore, by using the maximum PGV of a set of heterogeneous geologies, some F-FNO predictions will always be larger than the simulated ground truth. This also holds true when considering only geologies with a coefficient of variation smaller than 0.1.

Figure 6 gives the F-FNO prediction and simulated ground truth for four geologies at the same spatial location. Predictions are excellent when the geologies have small coefficients of variation (top figures). One can also notice that the phase differences are significant between samples and the F-FNO is still able to capture this variability. The bottom left panel exemplifies a slight overestimation of the PGV, even though the small variations are still very well captured. The bottom right panel (coefficient of variation equal to 0.2) illustrates a situation where heterogeneities created an amplification of ground motion around $3.7\,\mathrm{s}$, which may be a later arrival of S-waves. This singular case is not well predicted by the F-FNO, suggesting that geological coefficients of variation larger than 0.2 become very challenging.

3.3 Seismic hazard analysis

The influence of heterogeneities on ground motion can also be evaluated in terms of Pseudo-Spectral Acceleration (PSA) with a $5\,\%$ damping ratio. With the reference 1D geology, one gets a PSA value for each spatial point and each period considered ($0.3\,\mathrm{s}$, $1\,\mathrm{s}$, and $3\,\mathrm{s}$). In Fig. 7, the bottom right display represents the PSA map (16×16 grid) computed on the 1D reference geology for period
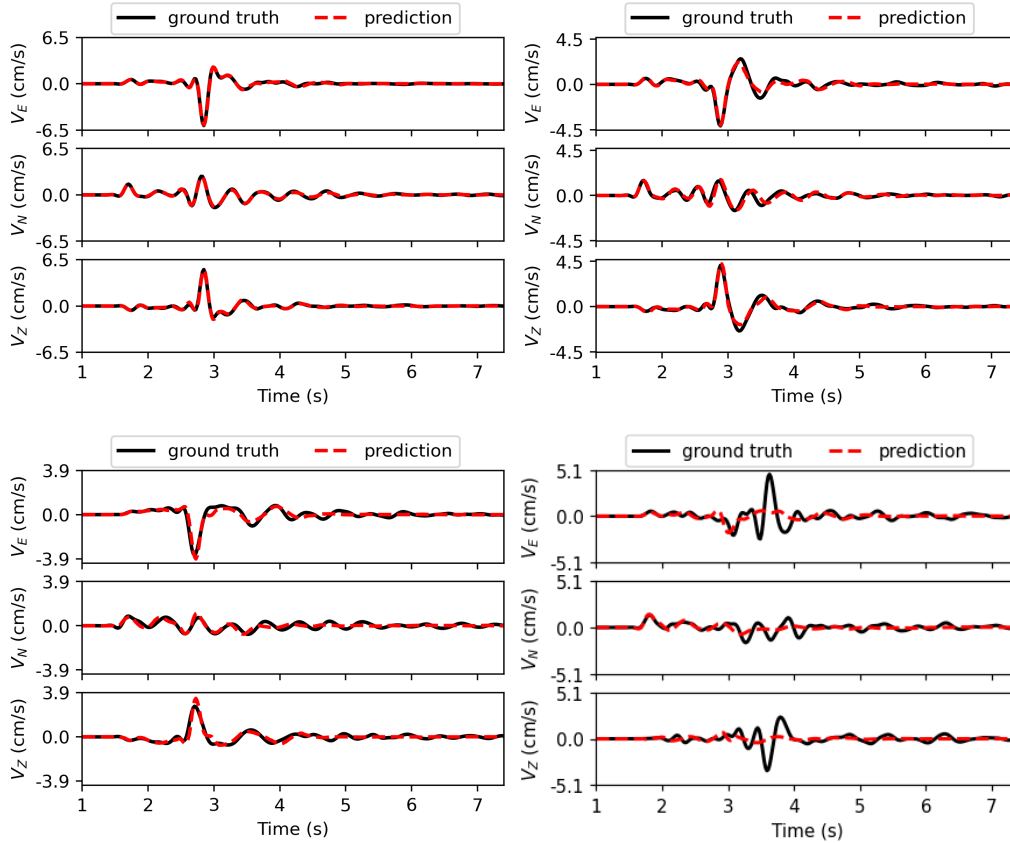
6

Figure 6: Three samples with coefficients of variation 0.05 (top left), 0.1 (top right), 0.2 (bottom left and right). Timeseries are shown at the same sensor for simulation results (black line) and F-FNO predictions (red dashed line).

T=0.3 s. Blue PSA curves in all other three displays are extractions of the PSA map along the horizontal blue line of sensors parallel to the x-axis.

One can then compute the PSA on this line for all geologies with a coefficient of variation smaller than 0.2. Figure 7 shows that the upper bound of the predicted PSA distribution (red shaded area) is significantly larger than the reference PSA (blue line). This implies that, even though F-FNO predictions are not perfect, using a set of heterogeneous geologies gives a considerable security margin compared to the usual method of considering a single homogeneous model.

One can additionally notice that the errors between predicted (grey areas) and true PSA (red areas) decrease when the period increases. This reflects earlier findings that small-scale patterns are more difficult to predict that large-scale evolutions [25].

4 Discussion and conclusion

We trained a Factorized Fourier Neural Operator (F-FNO) on the generic HEMEW-3D database to predict the relationship between geologies and surface ground motion. Acquiring the database of high-fidelity simulations is computationally demanding, and training a 3D neural operator on a large database also requires significant computational resources. However, the available pre-trained F-FNO can be used as the initialization for a second and much cheaper training thanks to transfer learning. We showed that as few as 250 transfer learning samples were necessary to obtain near-optimal accuracy for a database of ground motion based on the Le Teil earthquake. This training took only 26 min on one Nvidia A100 GPU, making it affordable for other site-specific applications.
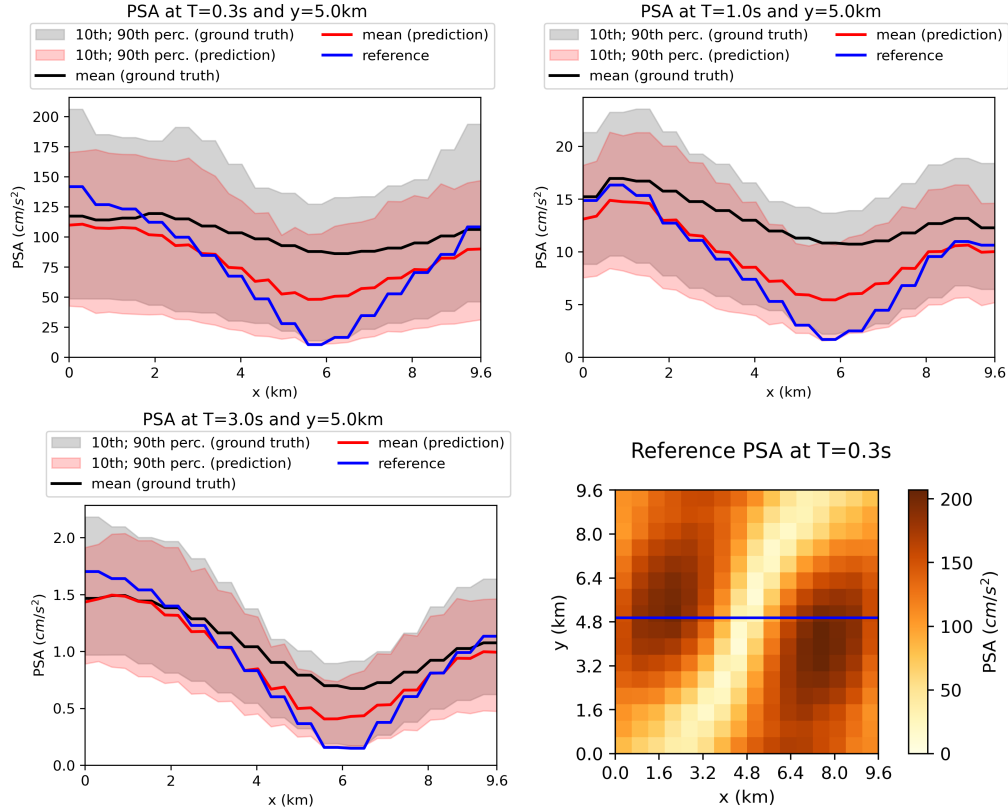
7

Figure 7: Pseudo-spectral acceleration (PSA) at periods $0.3\,\text{s}$, $1\,\text{s}$, and $3\,\text{s}$ for a sensor line parallel to the x-axis located at y=$5\,\text{km}$. The bottom right panel shows the PSA of the reference geology for all spatial points and indicates the line where the curves were computed (blue horizontal line). On PSA curves (three other panels), the blue line shows the PSA obtained with the reference geology, the black line shows the mean true PSA for geologies with a coefficient of variation smaller than 0.2, and the red line shows the mean prediction for the same geologies.

We proved that the F-FNO has an excellent phase accuracy for 91% of predictions and excellence amplitude accuracy for more than 62% of predictions (GOF larger than 8). Ground motion intensity was further assessed with the Peak Ground Velocity (PGV) and the Pseudo-Spectral Acceleration (PSA). Results indicate that the F-FNO tends to slightly underestimate intensity measures when comparing predictions and simulated ground truths for the same geology. However, when looking at a set of heterogeneous geologies, the F-FNO predictions are always significantly larger than the intensity measures obtained with a single simulation of the reference homogeneous geology. This indicates that neural operators are appropriate tools to quantify the ground motion uncertainties deriving from limited knowledge of the geology under a nuclear installation. Also, since predictions are entire timeseries of ground motion, any intensity measure can be computed for seismic hazard analyses.

It should however be noted that transfer learning was used between geologies of the same size and with the same seismic source. Extending the domains to larger areas and modifying the source's location and orientation was out of the scope of this study. Transfer learning may help in these directions but fundamental properties of neural operators should also be further explored.

Transfer learning has also been successfully applied on a database of geologies containing hard rock inclusions (not shown here). Although this database does not originate from a reference geology and hard rocks have very different shapes and locations inside the domain, the F-FNO predictions were accurate with only 500 transfer learning samples. Therefore, transfer learning on neural operators seems promising to reduce the computational costs associated with acquiring synthetic data and training large deep learning models.

References

[1] M. Causse et al. "Exceptional Ground Motion during the Shallow Mw 4.9 2019 Le Teil Earthquake, France". In: *Communications Earth & Environment* 2.1 (2021), p. 14. DOI: 10.1038/s43247-020-00089-0.

[2] F. Lehmann et al. "Machine Learning Opportunities to Conduct High-Fidelity Earthquake Simulations in Multi-Scale Heterogeneous Geology". In: *Frontiers in Earth Science* 10 (2022). DOI: 10.3389/feart.2022.1029160.

[3] K. Aki. "Local Site Effects on Weak and Strong Ground Motion". In: *Tectonophysics* 218.1-3 (1993), pp. 93–111. DOI: 10.1016/0040-1951(93)90262-I.

[4] F. J. Sánchez-Sesma, V. J. Palencia, and F. Luzón. "Estimation of Local Site Effects During Earthquakes: An Overview". In: *ISET Journal of Earthquake Technology* 39.3 (2002).

[5] P. Moczo et al. "Key Structural Parameters Affecting Earthquake Ground Motion in 2D and 3D Sedimentary Structures". In: *Bulletin of Earthquake Engineering* 16.6 (2018), pp. 2421–2450. DOI: 10.1007/s10518-018-0345-5.

[6] C. Smerzini, R. Paolucci, and M. Stupazzini. "Comparison of 3D, 2D and 1D Numerical Approaches to Predict Long Period Earthquake Ground Motion in the Gubbio Plain, Central Italy". In: *Bulletin of Earthquake Engineering* 9.6 (2011), pp. 2007–2029. DOI: 10.1007/s10518-011-9289-8.

[7] C. Zhu et al. "Seismic Aggravation in Shallow Basins in Addition to One-dimensional Site Amplification". In: *Journal of Earthquake Engineering* 24.9 (2020), pp. 1477–1499. DOI: 10.1080/13632469.2018.1472679.

[8] F. De Martin et al. "Influential Parameters on 3-D Synthetic Ground Motions in a Sedimentary Basin Derived from Global Sensitivity Analysis". In: *Geophysical Journal International* 227.3 (2021), pp. 1795–1817. DOI: 10.1093/gji/ggab304.

[9] A. Heinecke et al. "Petascale High Order Dynamic Rupture Earthquake Simulations on Heterogeneous Supercomputers". In: SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. New Orleans, LA, USA, 2014, pp. 3–14. DOI: 10.1109/SC.2014.6.

[10] S. Touhami et al. "SEM3D: A 3D High-Fidelity Numerical Earthquake Simulator for Broad-band (0–10 Hz) Seismic Response Prediction at a Regional Scale". In: *Geosciences* 12.3 (2022), p. 112. DOI: 10.3390/geosciences12030112.

[11] J. Douglas and B. Edwards. "Recent and Future Developments in Earthquake Ground Motion Estimation". In: *Earth-Science Reviews* 160 (2016), pp. 203–219. DOI: 10.1016/j.earscirev.2016.07.005.

[12] A. Alimoradi and J. L. Beck. "Machine-Learning Methods for Earthquake Ground Motion Analysis and Simulation". In: *Journal of Engineering Mechanics* 141.4 (2015), p. 04014147. DOI: 10.1061/(ASCE)EM.1943-7889.0000869.

[13] P. Sochala, F. De Martin, and O. Le Maître. "Model Reduction for Large-Scale Earthquake Simulation in an Uncertain 3D Medium". In: *International Journal for Uncertainty Quantification* 10.2 (2020), pp. 101–127. DOI: 10.1615/Int.J.UncertaintyQuantification.2020031165.

[14] J. M. Rekoske, A. Gabriel, and D. A. May. "Instantaneous Physics-Based Ground Motion Maps Using Reduced-Order Modeling". In: *Journal of Geophysical Research: Solid Earth* 128.8 (2023), e2023JB026975. DOI: 10.1029/2023JB026975.

[15] S. Bloemheuvel et al. "Graph Neural Networks for Multivariate Time Series Regression with Application to Seismic Data". In: *International Journal of Data Science and Analytics* 16.3 (2023), pp. 317–332. DOI: 10.1007/s41060-022-00349-6.

[16] Z. Zhang and Y. Wang. "A Spatiotemporal Model for Global Earthquake Prediction Based on Convolutional LSTM". In: *IEEE Transactions on Geoscience and Remote Sensing* (2023), pp. 1–1. DOI: 10.1109/TGRS.2023.3302316.

[17] Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning". In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539.

[18] M. Raissi, P. Perdikaris, and G. Karniadakis. "Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations". In: *Journal of Computational Physics* 378 (2019), pp. 686–707. DOI: 10.1016/j.jcp.2018.10.045.

[19] Y. Ding et al. "Self-Adaptive Physics-Driven Deep Learning for Seismic Wave Modeling in Complex Topography". In: *Engineering Applications of Artificial Intelligence* 123 (2023), p. 106425. DOI: 10.1016/j.engappai.2023.106425.

[20] D. Jozinović et al. "Rapid Prediction of Earthquake Ground Shaking Intensity Using Raw Waveform Data and a Convolutional Neural Network". In: *Geophysical Journal International* 222.2 (2020), pp. 1379–1389. DOI: 10.1093/gji/ggaa233.

[21] B. Moseley, A. Markham, and T. Nissen-Meyer. *Solving the Wave Equation with Physics-Informed Deep Learning*. 2020.

[22] Z. Li et al. *Neural Operator: Graph Kernel Network for Partial Differential Equations*. 2020. arXiv: 2003.03485. URL: http://arxiv.org/abs/2003.03485.

[23] L. Lu et al. "Learning Nonlinear Operators via DeepONet Based on the Universal Approximation Theorem of Operators". In: *Nature Machine Intelligence* 3.3 (2021), pp. 218–229. DOI: 10.1038/s42256-021-00302-5.

[24] Z. Li et al. "Fourier Neural Operator for Parametric Partial Differential Equations". In: *International Conference on Learning Representations*. 2021. DOI: 10.48550/arXiv.2010.08895.

[25] F. Lehmann et al. "Fourier Neural Operator Surrogate Model to Predict 3D Seismic Waves Propagation". In: 5th ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering. Athens, Greece, 2023. DOI: 10.7712/120223.10339.20362.

[26] P. Witte, R. Hewett, and R. Chandra. "Industry-Scale CO2 Flow Simulations with Model-Parallel Fourier Neural Operators". In: *NeurIPS 2022 Workshop Tackling Climate Change with Machine Learning*. 2022. URL: https://www.microsoft.com/en-us/research/publication/industry-scale-co2-flow-simulations-with-model-parallel-fourier-neural-operators/.

[27] A. Tran et al. "Factorized Fourier Neural Operators". In: *The Eleventh International Conference on Learning Representations*. 2023. URL: https://openreview.net/forum?id=tmIiMPl4IPa.

[28] B. Delouis et al. "Constraining the Point Source Parameters of the 11 November 2019 Mw 4.9 Le Teil Earthquake Using Multiple Relocation Approaches, First Motion and Full Waveform Inversions". In: *Comptes Rendus. Géoscience* 353.S1 (2021), pp. 1–24. DOI: 10.5802/crgeos.78.

[29] J. Quiñonero-Candela, ed. *Dataset Shift in Machine Learning*. Neural Information Processing Series. Cambridge, Mass: MIT Press, 2009. 229 pp.

[30] C. Smerzini et al. "Regional Physics-Based Simulation of Ground Motion within the Rhône Valley, France, during the Mw 4.9 2019 Le Teil Earthquake". In: *Bulletin of Earthquake Engineering* 21.4 (2023), pp. 1747–1774. DOI: 10.1007/s10518-022-01591-w.

[31] F. Lehmann. *Physics-Based Simulations of 3D Wave Propagation: A Dataset for Scientific Machine Learning*. 2023. DOI: 10.57745/LAI6YU.

[32] M. Kristeková, J. Kristek, and P. Moczo. "Time-Frequency Misfit and Goodness-of-Fit Criteria for Quantitative Comparison of Time Signals". In: *Geophysical Journal International* 178.2 (2009), pp. 813–825. DOI: 10.1111/j.1365-246X.2009.04177.x.