Machine Learning In Geophysics

A New Advances in Geophysics Conference

University of East Anglia 11th – 13th September 2023.

Abstracts

Sponsors

BGA
BRITISH GEOPHYSICAL ASSOCIATION

güralp

CGG

Geomatrix
EARTH SCIENCE LTD

SMARTSOLO

RSK

SILIXA

Raspberry Shake®

STRYDE

Invited Speakers

Machine Learning for understanding ocean bio-physical interactions

Dr Fatma Jebri
Senior Scientist
Marine Physics and Ocean Climate
National Oceanography Centre

Machine Learning (ML) is emerging as a powerful technique across geophysical science, from solid earth to ocean and atmospheric sciences. Unsupervised ML, a major class of ML, identifies relationships among geophysical parameters while reducing their large dimensionality and removing unstructured noise. Unsupervised techniques include k-means, Self-Organizing Maps (SOM), autoencoders and Generative Adversarial Networks. Unsupervised ML provides new ways in analysing complex ocean biological and physical interactions, particularly for understanding dynamics of productive upwelling regimes in Western Boundary Current systems. The SOM algorithm unravelled links between surface current variability and phytoplankton productivity during seasonal upwelling over the Agulhas Bank (South Africa) from daily surface ocean currents, sea surface temperature, and chlorophyll-a. The SOM patterns extracted over this dynamically complex region, revealed four topologies/modes of the Agulhas Current (AC) system. These AC topologies were found to influence the circulation and the phytoplankton productivity on the shelf. Consequently, the types of flows corresponding to productive upwelling regimes and those responsible for reduced levels of productivity were identified. The year-to-year variability of the SOM patterns indicates that the high/low productivity events seem to be linked to the occurrence of extreme phases in climate variability modes. SOM shows also promising results for estimating biophysical indicators of change in the North Atlantic. A recent application of unsupervised clustering led to the automatic detection of upwelling areas off the Somali coast and classification of extreme events. These ML approaches can be applied to other oceanic regions, or at the global scale, and extended to other datasets, including hindcast modelling outputs and future climate projections. This could provide new insights into the links between ocean parameters and how they are evolving as the Earth’s climate changes.

Imaging the Earth's Interior with Machine Learning

Dr Andrew Valentine
Assistant Professor
Department of Earth Sciences
Durham University

Earth imaging is one of the classic problems of geophysics: how do we "look inside" the planet and understand its structure, composition and processes? Confined as we are to the Earth's surface, we must extract information from geophysical datasets, and use these to construct and constrain our state of knowledge. However, doing so is challenging: we must often work with data that is limited in its scope, quantity or quality, and we typically cannot isolate the feature or process of interest from the wider Earth system.

As such, the growth of 'machine learning' (ML) offers much potential: the field promises new techniques and tools that are designed to extract information from complex data. Many studies have sought to introduce ML-inspired or ML-assisted methods into geophysical imaging, with applications targeting data processing, model parameterisation, regularisation, physical simulation, and the formulation of the imaging problem itself. In this talk, I will survey this range of opportunities, and show results from recent studies that draw on ML concepts and ideas. I will also seek to highlight some particular characteristics of the imaging problem that influence how and where ML ideas should be applied.

Training deep learning models with limited labelled data for seismic monitoring at volcanoes and glaciers

Dr Sacha Lapins
Leverhulme Early Career Fellow
School of Earth Sciences
University of Bristol

Training deep learning models for new monitoring settings and data types can be challenging. Manually curated or labelled target data are often scant or unavailable, and existing pre-trained models may be unsuitable or non-existent for the types of seismicity (e.g., non-tectonic) and instrument data (e.g., from fibre optic sensing) that require processing. In this talk, I present two case studies where effective deep learning seismic signal processing models have been trained using limited or no labelled data. First, I present the use of transfer learning to improve seismic phase arrival 'picking' performance on data recorded by a network deployed at Nabro volcano in Eritrea. Nabro erupted without warning in June 2011 and had never undergone any previous seismic monitoring prior to the deployment of this temporary network shortly after eruption onset. The transfer learning model exhibits significant performance improvements over its base model (GPD), as well as another comparable pre-trained deep learning model (PhaseNet) and an existing manually compiled seismic catalogue. Approximately 34,000 events are detected across the full 14-month deployment, revealing the presence of fluids, faults, and magma 'recharge' at this volcano. Following this, I present the use of so-called 'weakly supervised' learning to suppress strong random instrument noise in data recorded by a Distributed Acoustic Sensing (DAS) array deployed on an Antarctic glacier. DAS datasets are particularly challenging to analyse or label as they have high temporal and spatial sampling density, leading to very large data volumes that make manual curation infeasible. The demonstrated approach requires only noisy data and maps random noise to a chosen summary statistic, such as the distribution mean, median or mode, whilst retaining the true underlying signal, greatly enhancing the signal-to-noise levels of microseismic icequake events for subsequent detection (for which I present another model, trained using only 'sparse' labels).

Applications of Machine Learning techniques in Electrical Resistivity Tomography imaging and monitoring

Dr Paul Wilkinson
Theoretical Research Geophysicist
British Geological Survey

Paul Wilkinson[1]*, Wil Ward[2], Arnaud Watlet[3], Jim Whiteley[4], Cornelia Inauen[5], Luke Hawley-Sibbett[1], Michael Tso[6], Marco Iglesias[7], Andy Binley[8], Li Yue Yan[9], Chris Pain[9], Phil Meldrum[1], Oliver Kuras[1] & Jon Chambers[1]

[1]British Geological Survey, Nottingham, UK
[2]University of Sheffield, Sheffield, UK
[3]Université de Mons, Mons, Belgium
[4]Atkins, Birmingham, UK
[5]Alfred-Wegener-Institut, Germany
[6]UK Centre for Ecology & Hydrology, Lancaster, UK
[7]University of Nottingham, Nottingham, UK
[8]University of Lancaster, Lancaster, UK
[9]Imperial College London, London, UK
*pbw@bgs.ac.uk

The British Geological Survey has been developing electrical resistivity imaging and monitoring for a range of geoscience applications for 30 years. The workflow required to transform geoelectrical measurements into an understanding of the subsurface structure and processes under investigation comprises several key stages, all of which require some degree of expert input to achieve the best results. This is due to inherent aspects and limitations of the underlying physics, the inverse problems that must be solved, or the challenges of maintaining instrumentation in the field. During the last decade, working in collaboration with students and university partners, we have explored applications of machine learning and related techniques to assist with data preprocessing, inversion, and analysis and classification of the resulting geophysical images.

We have explored a variety of supervised and unsupervised clustering approaches to detect interfaces, classify heterogeneity, analyse complementary geophysical data, and identify hydrofacies in 4D resistivity models. We have also used clustering methods to identify problematic electrodes and demonstrated how principal component analysis control charts can detect short circuiting caused by connector damage. Other image segmentation methods combined with Kalman filters have been used to track tracers in monitoring experiments, and an inversion framework based on Ensemble Kalman filters has been used to image subsurface resistivity with uncertainty quantification. Solving the inverse problem in the presence of electrical anisotropy is the focus of our most recent work. The underlying discretised partial differential equation is recast in the form of weights of convolutional layers in a neural network. This allows it to be solved extremely quickly on GPUs / Artificial Intelligence processors using AI software libraries, which automatically calculate the sensitivities and provide optimisation methods to perform the inversion. We are using a latent diffusion model approach, which can be taught to incorporate representative geological priors.

Self-supervised seismic denoising: Deep learning without labels

Dr Claire Birnie
Research Scientist
Physical Sciences & Engineering
King Abdullah University of Science and Technology (KAUST)

Deep learning has revolutionized almost every field of science. However, traditional deep learning procedures are often supervised requiring a ground truth label that is used to train the network. In the field of geophysics, this label is often unobtainable, with the ground truth remaining unknown. One approach to still leverage deep learning in geophysics is to implement self-supervised methods, where the available data represents both the input and label for training a neural network. Given the example of noise suppression, in this presentation, I will illustrate how blind-spot, and subsequent blind-mask, networks can be utilised for the suppression of both random and coherent noise. Furthermore, I will introduce how we can exploit methods from the field of explainable artificial intelligence in order to design the optimum blind-mask, resulting in tailored noise suppression algorithms that require no clean training targets.

Delegate Abstracts (Alphabetical)

Seismic modelling and monitoring of CO_2 storage in geologic reservoirs

Adewunmi Abdul[1], Tarje Nissen-Meyer[1] and Mike Kendall[1]
[1]University of Oxford, England, United Kingdom

Abstract

Seismic data collected in vast amounts over a long period of time enables detailed characterization of CO2 migration within reservoir units; such data lends itself well to machine-learning techniques. This research aims to combine the knowledge acquired from the numerical modelling of wave propagation with machine learning in order to understand how leakage of sequestered CO2 into ambient rock formations manifests itself in seismic data.

We outline a methodology that simulates the corresponding waveform from a large range of structural scenarios arising from variations in the clastic rock and fluid properties obtained from similar CO2 monitoring works done by several other authors and extract a training example from each simulation for neural network using conditional encoder-decoder design. Some of these examples will be kept as a validation set during training and another as a test set to assess the performance of the trained network. Ultimately, the sensitivity of seismic waveforms to proxies such as thickness (reservoir, overburden, CO2), reservoir heterogeneity, velocity and CO2 saturation for leakage scenarios will be analyzed. The knowledge gained from such techniques will then be extended to time-lapse seismic data such as in the Utsira formation of the Sleipner field, North Sea where a large amount of CO2 is being sequestrated; thereby converting signals into meaningful information. Our work will enable further research directions such as investigating the value of poroelastic over acoustic media (which will be used in this study) in imaging the presence and migration of injected CO2, using passive seismic method for monitoring and the injection of CO2 into reservoirs containing a mixture of gas, oil and brine as opposed to the single-fluid injection into saline aquifer of this study.

References

Moseley, B., Nissen-Meyer, T., & Markham, A. (2020). Deep learning for fast simulation of seismic waves in complex media, Solid Earth, 11, 1527–1549, https://doi.org/10.5194/se-11-1527-2020.

Carcione, J., Picotti, S., Gei, D., & Rossi, G. (2006). Physics and seismic modelling of monitoring CO2 storage, Pure Appl. Geophys.,163, 175–207.

Introducing Conceptual Geological Prior Information into Bayesian Tomographic Imaging

Hugo Bloem[1], Andrew Curtis[1], Daniel Tetzlaff[2]
[1]School of Geosciences, University of Edinburgh, Edinburgh, United Kingdom
[2]Westchase Software Corporation, Houston, United States of America
hugo.bloem@ed.ac.uk

Abstract Text

Geological process models typically simulate a range of dynamic processes to evolve a base topography into a final 2-dimensional cross-section or 3-dimensional geological scenario. In principle, process parameters may be updated to better align with observed geophysical or geological data; however, it is hard to find *any* process model realisations that give *good* fits to all observations if data sets are complex and sparse in space (and hence time) because the simulations are typically chaotic}. Alternatively, geophysical probabilistic tomographic methods may be used to estimate the family of models of a target subsurface structure that are consistent both with information obtained from previous experiments and with new data (the Bayesian posterior probability distribution). However, this family seldom embodies geologically reasonable images. We show that the posterior distribution of tomographic images obtained from travel time data can be fully geological by injecting geological prior information into Bayesian inference, and that we can do this near-instantaneously using trained Mixture Density Networks (MDNs). We invoke two geological concepts as prior information about the depositional environment of an imaged target structure: a braided river system, and a set of marine parasequences, each parameterised by a Generative Adversarial Network. Data from a target structure can then be used to infer the image parameter values using either geological concept using MDNs. Our MDN solutions closely resemble those found using expensive Monte Carlo methods, and while the use of incorrect geological conceptual models provides less accurate results the mean structures still approximate the target. We also show that a classifier neural network can infer the correct geological conceptual model. So in summary, geological prior information significantly enhances geophysical tomography, imposing even incorrect geological prior information may still find geophysical tomographic images that resemble the true image, and in principle the correct geological conceptual model for an area can be inferred directly from geophysical data.

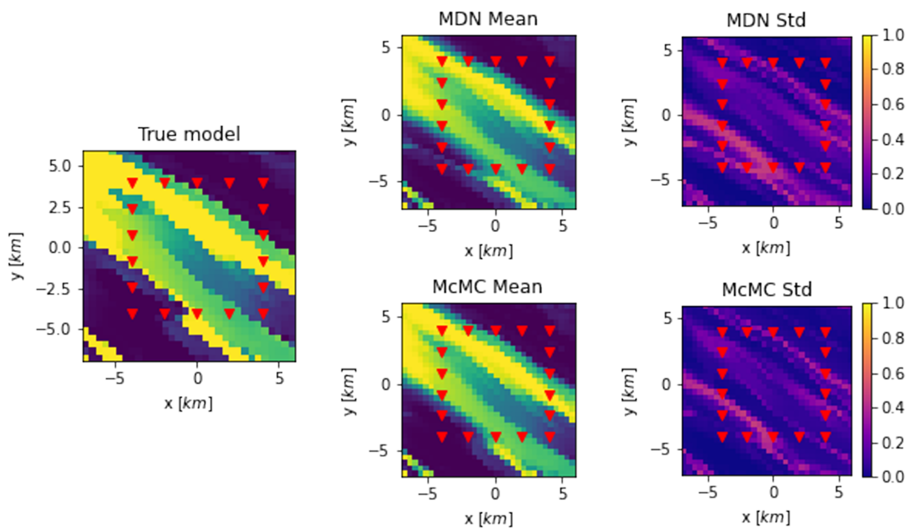


Figure 1: Inversion results using geological prior information. True model is shown on the left overlain with the source and receiver locations indicated by triangles. Posterior mean and posterior standard deviation are shown in the middle and right column. Top row represents results obtained using mixture density networks, bottom row using Markov chain Monte Carlo. Note that the former results took a fraction of the inference time compared to the latter.

3D and Time-Dependent Variational Bayesian Full Waveform Inversion

Andrew Curtis & Xin Zhang
University of Edinburgh, Edinburgh, U.K.
andrew.curtis@ed.ac.uk

Abstract

Seismic Tomography is a method to image the Earth's subsurface. In order to better interpret the resulting images it is important to assess imaging uncertainties, but this is hard to achieve. Monte Carlo random sampling methods are often applied for this purpose but the 'curse of dimensionality' makes them computationally intractable for high-dimensional parameter spaces. To extend uncertainty analysis to larger systems, variational inference methods developed in the machine learning community are introduced to seismic tomography. In contrast to random sampling, variational methods solve an optimization problem yet still provide probabilistic results.

Variational inference is applied to solve two types of tomographic problems: full waveform inversion (FWI), and time-dependent (known as 4D) FWI. Three different variational methods are tested: automatic differential variational inference (ADVI) and both deterministic and stochastic versions of Stein variational gradient descent (SVGD). ADVI provides a robust mean velocity model but biased uncertainties, whereas deterministic SVGD produces an accurate match to the results of Monte Carlo analysis, but at fraction of the computational cost. SVGD is significantly easier to parallelize, and for very large problems can be run in minibatch mode which is impossible using Monte Carlo methods without incurring probabilistic errors. Stochastic SVGD is shown to be the only method that may be capable of providing useable results for 3D FWI problems. This method is therefore extended to time-dependent monitoring problems of the type expected to be encountered in CO_2 of Hydrogen storage applications. Variational methods thus have the potential to extend probabilistic analysis to other Geophysical inverse problems and to higher dimensional tomographic systems than is currently thought possible.

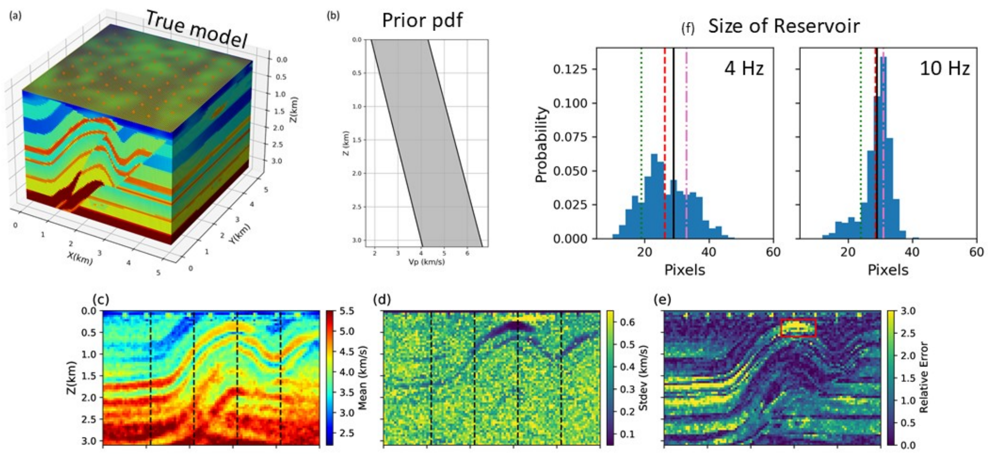


Figure: (a) 3D synthetic model. (b) Uniform prior probability density function. (c),(d),(e): Respectively the mean, standard deviation and relative error (difference between true and mean models, divided by the standard deviation) across central vertical slice in x-z plane. (f) Results of an interrogation problem to answer the question, "*How large is this storage reservoir?*" given probabilistic FWI results, using two different wavelet central frequencies: black line indicates the correct value.

Machine learning-based data-driven forecasting to improve earthquake predictability

Foteini Dervisi[a, b], Margarita Segou[a], Brian Baptie[a], Ian Main[b] and Andrew Curtis[b]
[a] British Geological Survey, The Lyell Centre, Edinburgh, United Kingdom
[b] The University of Edinburgh, School of GeoSciences, Edinburgh, United Kingdom
fdervisi@bgs.ac.uk

Abstract Text (Maximum 300 Words)

Recent advances in machine learning (ML)-based earthquake monitoring lead to the development of high-resolution earthquake catalogues that include at least a factor of ten more earthquakes compared to standard catalogues, while also considerably reducing the processing time needed compared to when catalogues are produced by human analysts (Tan et al. 2021). It has also been shown that physics-based and statistical forecasts are more informative when using deep learning (DL) catalogues compared to when standard catalogues are used (Mancini et al. 2022). This increase in the amount of available earthquake catalogue data indicates that ML techniques will uncover new relationships between earthquake catalogues (Beroza et al. 2021).

Here, we focus on the development of DL architectures to improve the predictive ability of earthquake forecasting models. We test the hypothesis that rich datasets paired with deep learning architectures can improve earthquake forecasting for earthquakes of all magnitudes and different characteristics. We employ convolutional neural network (U-Net) and transformer deep learning architectures and train the DL models using daily maps of earthquake activity from the 1980-2023 Southern California catalogue. We then evaluate the absolute and relative performance of the ML forecasts by comparing them against each other as well as against the persistence model, which assumes that no significant change occurs between consecutive days. The use of a persistence model, serving as the null hypothesis, is common in other predictive tasks today, such as weather forecasting. Preliminary results suggest that the trained DL models do not perform significantly better than the persistence model and that the transformer architecture achieves the best performance amongst the trained models. Our work has implications for the future of operational earthquake forecasting and earthquake predictability. Future work will focus on the validation of the DL architectures using the DL catalogues of the 2016-2017 Central Apennines earthquake sequence.

Abstract References in APA format (Maximum of three)

Beroza, G. C., Segou, M., & Mousavi, S. M. (2021). "Machine Learning and Earthquake Forecasting—next Steps." *Nature Communications* 12(1):4761. doi: 10.1038/s41467-021-24952-6.

Mancini, S., Segou, M., Werner, M. J., Parsons, T., Beroza, G., & Chiaraluce, L. (2022). "On the Use of High-Resolution and Deep-Learning Seismic Catalogs for Short-Term Earthquake Forecasts: Potential Benefits and Current Limitations." *Journal of Geophysical Research: Solid Earth* 127(11):e2022JB025202. doi: 10.1029/2022JB025202.

Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W., Michele, M., Chiaraluce, L., Beroza, G. C., & Segou, M. (2021). "Machine-Learning-Based High-Resolution Earthquake Catalog Reveals How Complex Fault Structures Were Activated during the 2016–2017 Central Italy Sequence." *The Seismic Record* 1(1):11–19. doi: 10.1785/0320210001.

Using neural networks to accelerate the single step ambient noise tomography forward problem enabling rapid 3D Monte Carlos Markov Chain inversions.

Joseph Fone[1], Nicholas Rawlinson[1]
[1]Bullard Labs – Department of Earth Sciences, University of Cambridge, Cambridge, UK
jwf39@cam.ac.uk

Abstract Text

Ambient noise tomography (ANT) is a powerful tool for passively imaging regional scale structures in the crust and upper mantle. It exploits the ambient noise wavefield to produce measurements of surface wave dispersion between seismic stations. Once inter-station dispersion measurements are obtained the inversion for ANT is conventionally done in two steps. Firstly, 2D tomography is performed to produce sets of phase or group velocity maps. Then these maps are sampled at discrete points to obtain local pseudo-dispersion curves which are then individually inverted for and then joined into a final 3D model. The issue with the two-step method is that in performing these inversions separately there is no correlation between the lateral inversions meaning that in certain locations of poor ray path coverage there can be very low correlation between the phase or group velocity pseudo-dispersion curves which can contain spikes and non-physical anomalies which will be reflected as artefacts in the final model. The single step ambient noise tomographic inversion deals with these problems by inverting directly for 3D shear wave velocity structure. The issue is that the forward problem is computationally expensive as it requires taking a 3D shear wave velocity model, doing dispersion calculations to get maps of phase or group velocity followed by multiple uses of an Eikonal solver to produce inter-station travel times. In this study we use a set of neural networks to approximate this forward problem for a network in northern Borneo training them on synthetic data. The new forward approximator shows a 100 time speed up which enables us to perform a fast 3D Monte Carlos Markov Chain inversion of real data from northern Borneo which is compared to a model from the two-step inversion. These results show promise that neural networks can help with some compute problems within seismology and geophysics.

Tracing the Central Italy 2016-2017 seismic sequence fault system: insights from unsupervised Machine Learning and Principal Component Analysis

Itahisa Gonzalez Alvarez[1], Margarita Segou[2], Brian Baptie[2]
[1]British Geological Survey, Keyworth, United Kingdom
[2]British Geological Survey, Edinburgh, United Kingdom
itaga@bgs.ac.uk

In recent years, we have witnessed the rise of Machine Learning (ML) in popularity and adoption across most scientific disciplines. The reasons behind this success are partly its versatility to adapt to different problems and types of data sets, the automatization of time-consuming repetitive tasks or its ability to learn complex relationships between observed variables. All of these make ML indispensable to the scientific discovery. In Seismology, ML has been applied to problems as different as earthquake detection and phase picking, signal classification, ground motion prediction or early warning systems development.

In this work, we investigate a rich deep learning seismic catalogue from the Central Italy 2016-2017 seismic sequence (Tan et al., 2021) with the aim of identifying active faults and study their distribution and evolution over the duration of the sequence. The catalogue, built using a deep-neural-network based phase picker, includes over 900 000 earthquakes with moment magnitudes ranging from 0.5 to 6.2, of which 72 000 contain focal mechanism information (p.c. Meier, 2023). For our analysis, we combine unsupervised clustering algorithms such as DBSCAN, HDBSCAN or OPTICS with Principal Component Analysis (PCA). Our preliminary clustering results of the full, year-long, catalogue, as well as extracted month-, and week-long catalogues, with and without focal mechanisms, reveal the presence of high-density clusters of earthquakes of varying extent within a cloud of diffuse seismicity. Through PCA, we associate some of these high-density clusters to individual faults, highlighting the complexity of the fault system and showing how a multitude of faults, often small-scale, became active at different points of the seismic sequence.

References:

Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W., Michele, M., Chiaraluce, L., Beroza, G. C., and Segou, M. (2021). Machine-Learning-Based High-Resolution Earthquake Catalog Reveals How Complex Fault Structures Were Activated during the 2016–2017 Central Italy Sequence. The Seismic Record. 1, 11–19, doi: 10.1785/0320210001.

Neural network-based modelling of hydrology in borehole strain observations

Jessica C. Hawthorne
Department of Earth Sciences, University of Oxford, Oxford, United Kingdom
jessica.hawthorne@earth.ox.ac.uk

Abstract Text

Borehole strainmeters were designed to record the tectonic deformation of the Earth with remarkably high precision: around 0.1 nanostrain. However, the instruments do not just record tectonics; they record all signals very precisely, and the largest signals in the data are usually not tectonic in origin. The largest signals are often tidal deformation, responses to an atmospheric pressure load, and responses to hydrological loads. Hydrological loads have been particularly challenging to model because the responses are nonlinear. For instance, the deformation caused by rainfall in June, when the ground is drier, may differ from the deformation caused by rainfall in January, when the ground is wetter and colder.

In this study, I model some of the nonlinear hydrology-induced deformation using neural networks, so that the induced deformation is a nonlinear function of past rainfall, pressure, temperature, and time of year. I test the approach with several Plate Boundary Observatory strainmeters in the Pacific Northwest, using ECMWF weather parameters as input. The current model capture more than 50% of the variation in deformation rate on timescales of a few to 10 days. As we must limit the number of free parameters, this model considers 10-20 linear combinations of past rainfall, allowing for variation in the load's timing. The model then multiplies the rainfall combinations by a nonlinear function of past rainfall, pressure, temperature, and time of year. The inferred weights provide some information about the physics of hydrological loading; more recent rain loads the strainmeter more, and strain recovers when the weather is dry. However, the moderate success of the model is likely more useful for tectonic observations. It suggests that we can remove at least half the non-tectonic "noise" with a quickly trained nonlinear model.

Machine Learning and the Mogi model: Improving the efficiency of ensemble-based methods for volcano deformation analyses

Matthew Head[1], Patricia M. Gregg[1]
Department of Earth Science & Environmental Change, University of Illinois at Urbana-Champaign, IL, USA
mshead@illinois.edu

Abstract Text

Geodetic observations are key for assessing the unrest status of volcanoes worldwide, providing critical information about magmatic systems and the potential for magma migration and eruption. Analysing these signals relies on a robust data-model framework. One such approach is the Ensemble Kalman Filter (EnKF; Evensen, 2003), a data assimilation method that has been adapted for analyses of volcanic deformation (Gregg & Pettijohn, 2017). The EnKF sequentially assimilates and inverts geodetic observations, 'nudging' model parameters to reduce the model-observation misfit with each iteration. We employ the Finite Element Method (FEM) to construct thermomechanical models of volcanic regions, providing the necessary flexibility to incorporate complex 3D geometries and material heterogeneity. However, these simulations are computationally expensive when incorporated into the EnKF workflow, with an ensemble of >200 model states taking several hours to evaluate.

Here, we aim to reduce the computational cost of the EnKF-FEM workflow by using regression machine learning algorithms (MLAs), focusing on reducing the number of model states that need to be evaluated by the FEM. We start by using the 'Mogi' deformation model (Mogi, 1958), a simple analytical expression that calculates the displacement field due to a point source. The Mogi model takes 5 input parameters and produces three-component deformation data (Ux, Uy, and Uz) on a 51 x 51 spatial grid, totalling 7803 observations. We employ a tuneable nearest-neighbour approach to identify model states that occupy a 'similar' parameter space, using MLAs to predict the resultant displacements. The MLAs are then updated with new model results after each iteration of the EnKF. While MLAs do not improve computational efficiency with the Mogi model, it has significantly reduced complexity compared to that of an FEM, providing a simple platform to test different approaches.

Abstract References

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Gregg, P. M., & Pettijohn, J. C. (2016). A multi-data stream assimilation framework for the assessment of volcanic unrest. *Journal of Volcanology and Geothermal Research*, *309*, 63–77. https://doi.org/10.1016/j.jvolgeores.2015.11.008

Mogi, K. (1958). Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bulletin of the Earthquake Research Institute, 36, 99-134*

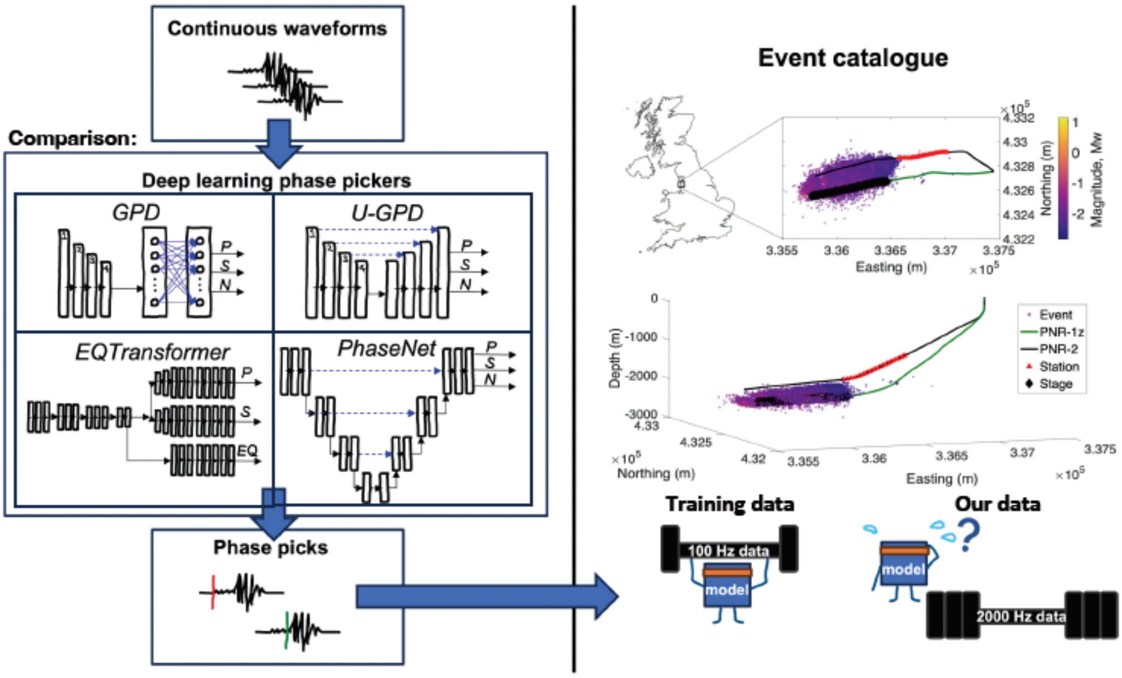
Deep learning phase pickers: can they detect induced seismicity?

Cindy Lim Shin Yee[1], Sacha Lapins[1], Margarita Segou[2], Maximilian Werner[1]
[1] School of Earth Sciences, University of Bristol, Bristol, UK.
[2] British Geological Survey (BGS), The Lyell Centre, Edinburgh, UK.
cindy.lim@bristol.ac.uk

Deep learning phase picking models have proven effective in processing large volumes of data produced during microseismic monitoring. These models have successfully detected earthquakes not catalogued by existing traditional methods (e.g., template matching, autocorrelation). Additionally detected earthquakes can affect the statistical properties of a catalogue (e.g., b-value) which is important for seismic hazard assessments. Deep learning enhanced catalogues may help us study different driving mechanisms of hydraulic fracturing induced seismicity (HFIS) by observing the spatiotemporal evolution of HFIS in greater detail.

We compared four established models (GPD, U-GPD, PhaseNet and EQTransformer) pre-trained on large volumes of regional earthquakes recorded on surface station datasets (100 Hz) and evaluated their ability to identify seismic phases in high-frequency (2000 Hz) borehole array data. We tested these models on the PNR-1z dataset, which comprises continuously recorded injection operations at a hydraulic fracturing site in Preston New Road, UK. Operators catalogued over 38,000 events using the Coalescence Microseismic Mapping (CMM) method. We generated earthquake catalogues for the PNR-1z dataset to compare (benchmark) against this initial catalogue.

The results demonstrated that PhaseNet detects seismic phases robustly within our data: it recovered up to 95% of the initial catalogue and detected over 15,800 additional events (36% increase). PhaseNet's robust application on our dataset could be due to its exposure to diverse instrument data during training, as well as its comparatively small model size which likely minimised overfitting to its initial training set. On the other hand, GPD, U-GPD and EQT require fine-tuning or re-training to improve microseismic event detection. We conclude that PhaseNet can be applied 'off-the-shelf' to detect HFIS in high frequency borehole data. Newly detected events could reveal new insights into the mechanisms controlling the spatiotemporal evolution of seismicity during fluid injections.



Earthquake detection and phase picking for nodal arrays in Indonesia with machine learning

Karen Lythgoe[1,2], Yukuan Chen[2], Erdong Huang[2], Zheng Yang[2], Chenyu Li[2], Muksin Umar[3], Shengji Wei[2]
[1]School of Geosciences, University of Edinburgh, Scotland
[2]Earth Observatory of Singapore, Nanyang Technological University, Singapore
[3]Syiah Kuala University, Banda Aceh, Aceh, Indonesia
karen.lythgoe@ed.ac.uk

Abstract Text

Machine learning detection and phase picking of earthquakes has vastly sped up the time taken to produce detailed earthquake catalogues at large seismic arrays at two locations in Indonesia. We use a deep learning model trained on global data – the EQTransformer algorithm (Mousavi et al 2020). Our first region is Lombok, where the month-long dataset contains multiple M6+ mainshock and aftershock sequences. The accuracy of the detection result, according to our manual evaluation, is estimated to be around 90%, which is acceptable because the majority of the false positive events are discarded during the subsequent association process. We find that EQTransformer finds over five times more events compared to the standard catalogue which was produced using a traditional STA/LTA detector, including detecting several M5+ aftershocks in the immediate coda of a M6.9 event which were originally missed. However, it misses the largest M6.9 events likely reflecting the relative lack of larger magnitude events in the training data. Pick quality is high, with a mean difference of 0s and standard deviation of 0.2s for both P and S, compared to picks by analysts. The new catalogue allows us to investigate earthquake triggering in more detail, and detect many more repeating event families.

We also use transfer learning to optimise the deep learning model for a dense nodal array dataset, comprising of over 120 stations, along the Sumatran Fault in Aceh, Sumatra. The 18-months of data is generally noisy and there are initially many false detections. We first apply the neural network to 5 weeks of data. These phase picks are then quality controlled, and the best picks are used to re-train the network. Using transfer learning approximately 10,000 high-quality events are detected and re-located. The detected events include local and regional earthquakes, in addition to low-frequency earthquakes that may be indicative of slow slip and have never been detected in this region before.

Abstract References

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, *11*(1), 3952.

TerraPINN: solving the wave equation with approximately axisymmetric physics informed neural networks

Jack B. Muir, Tarje Nissen-Meyer
Department of Earth Sciences, University of Oxford, Oxford, UK
jack.muir@earth.ox.ac.uk

Abstract Text

Deep neural networks have revolutionised our ability to process and classify big data. In seismology in particular, neural network solutions now outperform human analysts in accuracy on standard observational tasks, while allowing orders of magnitude more data to be processed[1]. This revolution was only possible, however, because of the large, high quality labelled datasets accrued by researchers over many decades. As we move towards the next generation of geophysical machine learning, we are beginning to tackle problems where such labelled data is not available. One of the most pressing challenges is the solution of the forward seismic wave equation using deep learning. Ideally, by significantly accelerating existing solvers, we would be able to generate much larger ensembles of seismic wavefield data through realistic media. However, because the forward solution is intrinsically expensive, we do not have large volumes of training data. The physics informed neural network (PINN) framework offers a way to circumvent this problem[2]. Instead of training on synthetic data, we propose solutions and then penalise their misfit to the wave equation. Early investigations of PINNs have shown much promise, however they have so far struggled to solve multi scale problems, such as the seismic wave equation. In TerraPINN, we propose a hybrid approach between PINNs and traditional supervised machine learning. Recognising the approximate axisymmetry of seismic wave propagation and taking inspiration from the success of the factored form of the eikonal equation in calculating first arrivals[3], we first fit a reduced dimension radial wavefield in a laterally homogeneous and isotopic medium using a traditional supervised machine learning framework. We then expand azimuthally and train for a correction operator using PINNs. The combined network size has 2 orders of magnitude fewer parameters and trains 10x faster than an equivalent naive PINN formulation in 2D acoustic test cases.

Abstract References

[1]Mousavi, S.M. , & Beroza, G. C. (2023) *Machine Learning in Earthquake Seismology*. Annual Review of Earth and Planetary Sciences, Volume 51 pp. 105–129. https://doi.org/10.1146/annurev-earth-071822-100323

[2]Moseley, B., Markham, A., Nissen-Meyer, T. (2020) *Solving the wave equation with physics informed neural networks*. arXiv:2006.11894v1

[3]Fomel, S., Luo, S., Zhao, H. (2009). *Fast sweeping method for the factored eikonal equation*. Journal of Computational Physics, Volume 228, Issue 17, pp. 6440–6455. https://doi.org/10.1016/j.jcp.2009.05.029

Automatic waveform earthquake location using convolutional neural networks

Joanne Sheppard[1], James Ward[2], Andy Nowacki[2,*]
1 Advanced Research Computing, Durham University, Durham, UK
2 School of Earth and Environment, University of Leeds, Leeds, UK
* andy.nowacki@leeds.ac.uk

Abstract Text

Earthquakes pose risks and provide useful information about natural and human-induced processes. For these reasons, and due to ever-larger volumes of seismic data, it has become increasingly necessary and routine to detect and locate earthquakes automatically, usually using a variety of heuristics-based approaches and localisation techniques. There are limitations, however. For example, single- or multichannel triggers usually require parameters to be tuned to the dataset. Likewise, migration techniques are computationally expensive. Some of these limitations have been addressed over recent years using data-driven approaches such as deep learning, but often this relies on sufficient knowledge of the seismic wave speed of the subsurface. Here we present work where waveforms from multiple seismic recorders are used simultaneously and directly to locate earthquakes with no intermediate picking or location step, making use of the full waveform and without *a priori* knowledge of the regional velocity structure.

We use a convolutional neural network (CNN) which feeds into a multi-layer perceptron, similar to van den Ende & Ampuero (2020), to predict the 3D earthquake location by training its parameters on data from 17 three-component seismometers from 4 to 25 July 2019, recording the Ridge Crest earthquake sequence. We use a window of 10 s around the earthquake origin time and include the position of each channel as additional input to the network. This setup yields locations from the test dataset within a few km of the catalogue locations after only 20 epochs of training on a single CPU. We test the transferability of the network by applying it to a different dataset with different stations, and explore the ability of it to cope with missing stations, including by using a fully-convolutional network. The results are encouraging and demonstrate that data-driven methods to locate earthquakes are likely to become increasingly routine and important.

References

van den Ende, M.P.A., & Ampuero, J.-P. (2020). Automated seismic source characterization using deep graph neural networks. *Geophysical Research Letters*, 47, e2020GL088690. https://doi.org/10.1029/2020GL088690

Unsupervised machine learning to identify mantle plumes in geodynamic models.

James Panton[1,a], Alistair Campbell[1], Huw Davies[1], Jamie Ward[2]
[1]School of Earth and Environmental Sciences, Cardiff University, Cardiff, Wales
[2]School of Earth and Environment, University of Leeds, Leeds, UK
[a]pantonjc@cardiff.ac.uk

Abstract Text

As the field of numerical geodynamic modelling has advanced, we now routinely see simulations which reproduce Earth-like structures such as the deep, seismically observable Large Low Velocity Provinces (LLVPs), and upwelling features, namely mantle plumes. Naturally, we now want to interrogate these modelled structures to better understand their properties and allow for comparison against observations of their terrestrial counterparts. In particular, it is of interest to identify and interrogate mantle plumes to better understand their temperature structure, radial velocity, composition and longevity. As such we need a method by which mantle plumes can be objectively identified within geodynamic models to make studies reproducible and to allow efficient post processing of model results. Building on previously described methods (Hassan., et al, 2015) we present a workflow for identifying mantle plumes produced in geodynamic simulations carried out using the 3D mantle convection code TERRA. The workflow incorporates unsupervised machine learning algorithms such as K-means and DBSCAN to first identify regions with plume-like properties and then separate them into individual plumes. This work will be used as part of Mantle Circulation Constrained (MC^2) project to allow for comparisons against estimates of plume fluxes, mantle potential temperatures and other inferred properties (Matthews., et al, 2021). We demonstrate the effectiveness of the algorithms by tracking plumes over the course of a simulation and report on how they evolve.

Abstract References

Hassan,R., Flament, N., Gurnis, M., Bower, D., Müller, D. (2015). Provenance of plumes in global convection models. Geochemistry, Geophysics, Geosystems, 16 (5), 1465-1489. https://doi.org/10.1002/2015GC005751

Matthews, S., Wong, K., Shorttle, O., Edmonds, M., Maclennan, J. (2021). Do olivine crystallization temperatures faithfully record mantle temperature variability? Geochemistry, Geophysics, Geosystems, 22 (4), e2020GC009157, https://doi.org/10.1029/2020GC009157

Fields and Flows: Determining the dynamics of the core mantle boundary using Physics-Informed Machine Learning

Naomi Shakespeare-Rees, Philip Livermore, Christopher Davies, Ciaran Beggan, William Brown, Christopher Finlay
School of Earth and Environment, University of Leeds, Leeds, UK
eenmsr@leeds.ac.uk

Abstract Text

The Earth's geomagnetic field arises from the constant motion of the fluid outer core. By assuming that these motions are advection-dominated, rather than diffusion, one can relate this motion at the core surface to the secular variation of the geomagnetic field, providing an observational approach to understanding the motions in the deep earth. Existing methods predominantly employ global inversions, assuming large-scale solutions where all observed secular variations are attributed to the flow. In contrast, this work introduces a novel technique based on machine learning, specifically Physics-Informed Neural Networks, to perform local flow inversions. Our approach incorporates a loss function comprising of both data loss and physics-based loss, in which different flow assumptions can be swapped in and out when needed. This poster presents the set-up, underlying assumptions, and preliminary results of this methodology using Toroidal and Tangentially Geostrophic flow constraints. Furthermore, we discuss the technical and scientific next steps to advance this method as a powerful tool in understanding the dynamics of the Earth's core.

Abstract References

Roberts, P. H., & Scott, S. (1965). On analysis of the secular variation. Journal of geomagnetism and geoelectricity, 17 (2), 137–151. https://doi.org/10.5636/jgg.17.137

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https : //doi.org/10.1016/j.jcp.2018.10.045

Holme, R. (2007). 8.04—Large-Scale Flow in the Core. In G. Schubert (Ed.), *Treatise on Geophysics* (pp. 107–130). Elsevier. https://doi.org/10.1016/B978-044452748-6.00127-9

Variational Bayesian Experimental Design for Geophysical Applications

Dominik Strutz, Andrew Curtis
School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FE, UK
dstrutz@ed.ac.uk

Abstract Text

In geophysical experiments or surveys, recorded data are used to constrain target properties or dynamics of the planetary subsurface, oceans, cryosphere or atmosphere. The exact choice of experimental design controls how much, and precisely what information is transferred to target variables. Typical design parameters that can be varied are source and sensor types and locations, and the choice of modelling or data processing methods to be applied to the data. These may all be optimised subject to various cost constraints. Bayesian experimental design methods quantify and maximise information about targets of interest. This constitutes a macro-optimisation problem in which we design the set of Bayesian inference problems that we might encounter post-experiment.

We introduce novel variational design methods that leverage functional approximations to probability distributions and model-data relationships. These methods have gained prominence in the machine learning community to optimize the design of experiments. We show that they enable accurate estimation of model parameters, or allow experiments to be focused on answers to specific questions about the system under investigation.

Our variational methods rely on neural network approximations to probability functions, which need to be trained similarly to many other machine learning approaches. We will present how mixture density networks and mutual information lower bounds can be used for the design of focused experiments and surveys in geophysics.

To illustrate the advantages of these methods, we show that they enable passive seismological surveys to be designed to locate earthquakes optimally, and active seismic surveys to be designed specifically to constrain CO_2 saturation in subsurface storage scenarios. These applications demonstrate that optimal designs can vary substantially depending on the objectives of interest. Full details are available in Strutz & Curtis (2023).

Abstract References

Strutz, D., & Curtis, A. (2023). Variational Bayesian experimental design for geophysical applications. In *arXiv [physics.geo-ph]*. arXiv. http://arxiv.org/abs/2307.01039

Predicting Urban Traffic Patterns Using AI and Seismic Signatures: Insights for Sustainable Urban Planning

Rochelle Taylor[1], Luciano Gerber[1], David Healy[2]

1. Manchester Metropolitan University, Manchester, UK
2. University of Aberdeen, Aberdeen, UK

Rochelle.Taylor@mmu.ac.uk

A climate emergency was acknowledged in 2021 (IPCC, 2021) as greenhouse gas concentrations in the atmosphere are continually rising and are a catalyst for global warming. A key factor of the acceleration of atmospheric greenhouse gases is fossil fuel combustion; national estimates indicate that 38% of atmospheric CO_2 comes from road transportation (BEIS, 2019). To gain accurate local data and a better understanding of the sector's impact, a network of Raspberry Shake seismometers has been deployed across Greater Manchester, as part of a UKRI NERC-funded project "Listen to Manchester" (Twitter: @listen2mcr). These seismometers possess high sensitivity to high frequency anthropogenic noise, making them suitable for capturing local seismic signatures.

The urgent need to comprehend and predict urban traffic patterns stems from the alarming levels of atmospheric CO_2, which have reached their highest point in the last 650,000 years (Lüthi et al. 2008), and the visible air quality improvements seen during the 2020 global pandemic. Understanding factors influencing air quality and traffic volume is essential for sustainable urban planning and the development of effective transportation management strategies.

In recent years, the application of artificial intelligence (AI) techniques in geosciences has gained attention and has had positive impacts on geoscience research. In this study, we employ AI algorithms to detect seismic signals associated with anthropogenic noise. By extracting features from the power spectrum of the frequency domain, we predict urban traffic volumes along Manchester City Centre's Oxford Road corridor.

Our research aims to provide valuable insights into the local dynamics of CO_2 emissions and the influence of road travel, particularly when combined with air quality data. These insights will be vital for smart city development and advancing our understanding of climate change. The integration of AI techniques with geoscientific data holds immense potential to facilitate evidence-based decision-making to mitigate its detrimental effects.

References

1. BEIS. (2019). Greenhouse gas reporting: conversion factors 2019. Retrieved from https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2019
2. Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., & Stocker, T. F. (2008). High-resolution carbon dioxide concentration record 65,000-800,000 years before present. Nature, 453. https://doi.org/10.1038/nature06949
3. IPCC. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (V. Masson-Delmotte, et al., Eds.). Cambridge University Press. (In Press)

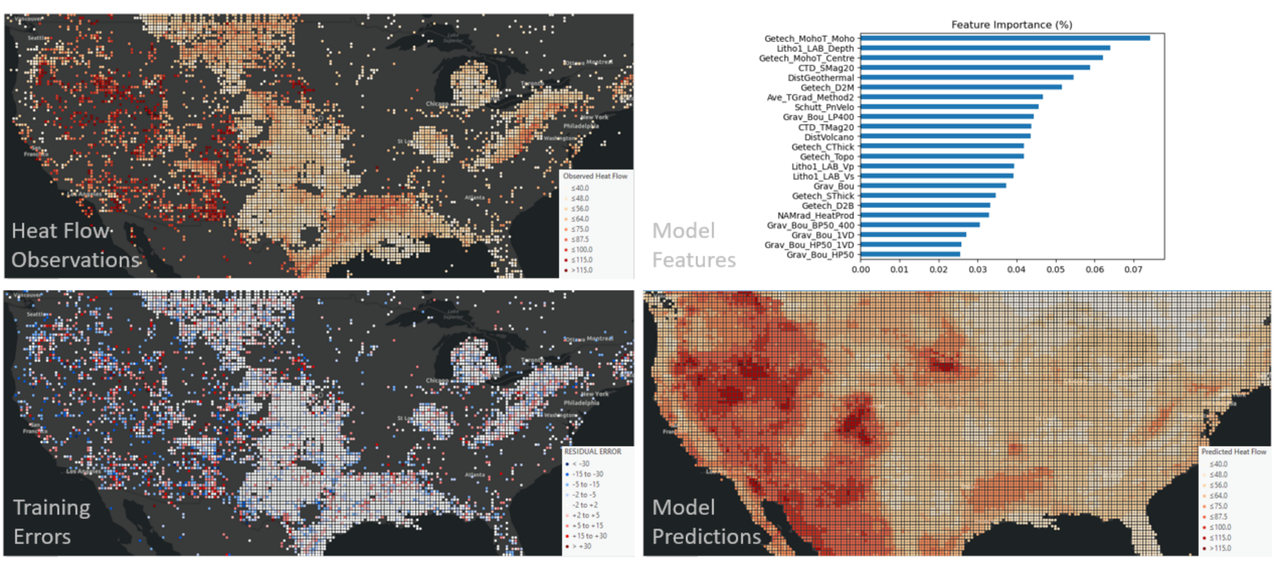
Predicting Heat Flow for Resource Exploration using Random Forest Regression

Peter Webb[1], Samuel Cheyney[1], Sheona Masterton[1], Chris Green[1,2]
[1] Getech Group, Kitson House, Leeds, UK
[2] School of Earth and Environment, University of Leeds, Leeds, UK
peter.webb@getech.com

Heat flow is an important physical observation in traditional and transitional resource exploration. The availability of direct measurements of heat flow is inconsistent across the globe, with relatively dense measurements in North America and Europe and sparser measurements elsewhere. Several geophysical and geological data sets and observations may act as proxies for heat flow, however variability on small and large scales makes it challenging to correlate individual data sets to the heat flow measurements. Therefore, a multivariate regression approach is sensible. Here we investigate the use of random forest regression techniques to predict terrestrial heat flow across the globe.

We use random forest regressions to estimate heat flow globally on a quarter-degree grid. Our training data set is made up of nearly 75000 measurements of heat flow provided in the International Heat Flow Commission (IHFC) compilation. The training uses up to 25 independent variables including depths to basement, Moho and asthenosphere, crustal type, P_n velocity (a proxy for Moho temperature), seismic velocity at the top of the asthenosphere, and proximity to surface thermal features. Our early random forest regression models were for the United States, where heat flow observations are plentiful, and the method was assessed using high-resolution independent variables before we expanded the analysis to the rest of the globe.

Owing to the high-quality data available, random forest regression models for the United States were the most successful, achieving R^2 values of around 0.7 for hold-out model validation data. Global models achieved R^2 values between 0.55 and 0.65 depending on the data used for training. The difference in performance can be partly attributed to the absence of some variables (e.g., the high-quality seismic based variables obtained from US Array studies) at the global scale. The most valuable independent variables for the random forest regression tended to be the more smoothly varying parameters; this is reflected by lower precision model predictions where observed heat flow is highest. These results suggest short-wavelength or local geological factors are still important in areas with the highest heat flow. Given the repeatable method, ever-improving data compilations and well-understood limitations these results have value for resource exploration, particularly where direct observations are limited.



Hyperbola detection with Retinanet in archaeological GPR datasets

Tina Wunderlich, Dennis Wilken, Bente S. Majchczack, Clemens Mohr, Wolfgang Rabbel
Institute of Geosciences, Kiel University, Kiel, Germany
tina.wunderlich@ifg.uni-kiel.de

Abstract Text

The use of machine learning methods for the automatic detection of features in Ground Penetrating Radar (GPR) data became popular in the last years and is mainly applied to urban infrastructure or tree roots. These infrastructures such as pipes or rebar produce hyperbolic pattern in the radargrams which can be located with e.g. convolutional neural networks (CNN). Hyperbolic pattern are not limited to infrastructure datasets but appear also in data from archaeological sites, originating from small stones or wooden structures (either natural or artificial). In order to detect these features automatically, we trained a CNN (Retinanet, Lin et al. 2017a,b) with GPR data from an archaeological site (Wunderlich et al. 2022). The average precision was 0.58 and more than 38000 hyperbola were detected. Subsequent automatic velocity analysis of the detected hyperbola results in a 3D velocity model, which can be used e.g. for migration. The spatial distribution of objects helps to improve the archaeological interpretation. Tests with other datasets showed high potential for the transfer of the trained Retinanet to new data as long as the aspect ratio of the figures stays the same.

Abstract References

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017a). Focal loss for dense object detection. Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2980-2988. https://doi.org/10.48550/arXiv.1708.02002

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017b). Feature Pyramid Networks for Object Detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 936-944. https://doi.org/10.1109/CVPR.2017.106

Wunderlich, T., Wilken, D., Majchczack, B. S., Segschneider, M., & Rabbel, W. (2022). Hyperbola Detection with RetinaNet and Comparison of Hyperbola Fitting Methods in GPR Data from an Archaeological Site. Remote Sens. 14, 3665. https://doi.org/10.3390/rs14153665

Automatic crater detection using semi-supervised machine learning

Jingyi Zhao, Gareth S. Collins, Navjot Kukreja
Imperial College London, London, UK
Imperial College London, London, UK
jz1121@ic.ac.uk

Abstract Text

Crater populations are an important tool for understanding subsurface properties and the geological history of planetary bodies. Identifying, characterising and measuring impact craters on solid planetary surfaces, however, can be a difficult and time-consuming task. Automatic crater detection algorithms (CDA) based on machine learning are effective alternatives to manual crater counting, especially for small craters (La Grassa et al 2023). The laborious process of hand labelling makes it challenging and time-consuming to build a fully labelled benchmark dataset for CDAs. Recent research on deep self-training offers a potent method for unsupervised or semi-supervised domain adaptation, which involves an iterative process of predicting on the target domain and then taking the confident predictions as pseudo-labels for retraining (Rosenberg et al 2005). However, as pseudo-labels can introduce unwanted noise, the self-training process poses a risk in assigning overconfident labels to incorrect detections, resulting in error accumulation over iterations. This paper proposes an auto-iterative, self-training system for automatic crater detection based on YOLO (Redmon et al 2016) to detect unlabelled small craters. The pseudo-label selection metric includels a confidence threshold that varies with with crater size and a novel IOU-based ensemble learning module that combines detection results from multiple models [explain the key difference between models]. Validated on a manually labelled dataset of all-size craters in THEMIS images of Mars, the proposed method outperforms a state-of-the-art CDA (Benedix et al., 2020) and excels in standard self-training systems in terms of precision and recall. In general, the proposed self-learning approach generates reliable results in small crater detection in a user-friendly and efficient manner.

Abstract References

La Grassa, R., Cremonese, G., Gallo, I., Re, C., & Martellato, E. (2023). YOLOLens: A Deep Learning Model Based on Super-Resolution to Enhance the Crater Detection of the Planetary Surfaces. Remote Sensing, 15(5), 1171.

Rosenberg, C., Hebert, M., & Schneiderman, H. (2005). Semi-supervised self-training of object detection models.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Bayesian Inversion using Boosting Variational Inference

Xuebin Zhao and Andrew Curtis
School of GeoSciences, University of Edinburgh, United Kingdom
Xuebin.zhao@ed.ac.uk

Abstract Text

Geophysical inversions use observed data to estimate properties of the Earth's interior, but they often pose non-linear and non-unique challenges. Bayesian inference provides a probabilistic framework for solving inverse problems and enables quantification of uncertainties in inversion results. Recently, variational inference has emerged as an efficient alternative to expensive Monte Carlo sampling methods, particularly for high-dimensional problems. By seeking the closest distribution to the unknown posterior distribution within a family of distributions, variational inference can yield a fully probabilistic solution. However, defining expressive variational families can significantly increase optimization complexity. In this paper, we introduce a new method called *Boosting Variational Inference* (BVI) to geophysics, which constructs a flexible approximating family comprising all possible finite mixtures of simple component distributions. Specifically, we use the Gaussian distribution as the mixture component due to its ease of training and fully parametric nature. Each component is sequentially trained using a greedy algorithm. We apply BVI to seismic travel time tomography and full waveform inversion, comparing its performance with Monte Carlo and other variational methods. The results demonstrate that BVI achieves both efficiency and accuracy while enabling the construction of an analytic posterior expression. Consequently, importance samples can be obtained directly from each component. These samples represent part of the uncertainty information and can be used to interrogate the subsurface structures efficiently.