Enabling deep-learning-based uncertainty quantification at scale for post-stack UHR seismic inversion

Authors: Gabrio Rizzuti, Rob Telling, Ivan Vasconcelos

Affiliation: *Shearwater Geoservices*

Summary: An accurate characterization of the shallow subsurface is crucial for evaluating the load-bearing capacity at offshore wind farm sites. It is widely recognized that seismic methods, particularly inversion, can significantly enhance traditional geotechnical approaches with three-dimensional information that is otherwise unavailable. A comprehensive geotechnical analysis, however, relies on several scenarios with different confidence levels to establish robust and trustworthy geotechnical parameter ranges. Deterministic seismic inversion is inherently ill-suited for this purpose because it provides only the most likely subsurface characterization. Here, we discuss a variational-inference-based Bayesian framework for UHR post-stack inversion, explicitly characterizing the uncertainties in acoustic impedance estimates. A notable feature of this approach is the ability to express these uncertainties in terms of credible intervals with arbitrary credibility. Our methodology leverages advanced deep-learning techniques, such as generative modeling via normalizing flows, to enable efficient and scalable implementations.

Keywords: UHR, acoustic impedance, uncertainty quantification, deep learning

Enabling uncertainty quantification at scale for post-stack UHR seismic inversion

Advancing quantitative seismic inversion for offshore windfarm site assessment

For shallow subsurface characterization at offshore windfarm sites, geotechnical data such as cone penetration testing (CPT) and ultra-high-resolution (UHR) seismic data are typically collected. Seismic surveys - providing input for volumetric 3D subsurface information - are relatively efficient to acquire and can effectively complement CPT measurements, which in turn are spatially limited and expensive. However, seismic methods are mostly limited to structural imaging and shallow hazard identification, overlooking their potential for deriving quantitative subsurface properties, particularly in a geotechnical context.

Achieving advanced quantitative inversion of seismic data requires specialized processing workflows commonly used in hydrocarbon exploration. However, these are rarely employed for windfarm site assessments and require notable adaptations to geotechnical requirements. These workflows include steps like de-signature, source-/receiver-side de-ghosting, and static corrections. Applying such techniques to UHR data poses challenges due to the lack of precise calibration measurements and the significant impact of sea-surface wave perturbations, but recent works demonstrated that this preprocessing can be adapted to UHR data as well (see, for instance, Telling et al., 2024).

In this study, we focus on a probabilistic analysis of UHR data. We aim to assess the uncertainties associated with acoustic impedance inversion from migrated stacks, as impedance estimates can be related to volumetric CPT properties (Klinkvort et al., 2024). This uncertainty estimation is critical in wind farm site assessment where conservative hazard estimates are required. We base our analysis on Bayesian uncertainty quantification, leveraging the approach discussed by Rizzuti and Vasconcelos, 2024, whose feasibility at scale is crucially enabled by modern deep-learning advancements.

Bayesian uncertainty quantification with deep “invertible” variational inference

In this section, we describe in broad terms the goal of Bayesian uncertainty quantification for post-stack acoustic seismic inversion. A linear model for seismic post-stack data $\boldsymbol{y}$ given acoustic impedance $\boldsymbol{x}$ (more precisely, the logarithm thereof) is often assumed, together with normally distributed additive noise:

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{D}_{\boldsymbol{t}}\boldsymbol{x} + \boldsymbol{n}, \qquad \boldsymbol{n} \sim N(\boldsymbol{0}, \sigma^2).$$

The linear operator is given by a time derivative followed by the (known) wavelet convolution. This describes the data *likelihood* model. Moreover, we also assume a *prior* distribution $p(\boldsymbol{x})$. According to the Bayes' rule, the *posterior* distribution of $\boldsymbol{x}$ knowing $\boldsymbol{y}$ is then

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}).$$

Bayesian uncertainty quantification is the characterization of the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$.

The traditional go-to tool of choice for uncertainty quantification in inverse problems has been Markov chain Monte Carlo sampling (McMC, Robert and Casella, 2010). Unfortunately, McMC methods do not scale effectively to large-sized inverse problems (especially in 3D), which has effectively limited Bayesian analysis for field data.

An alternative class of methods and the focus for this work is variational inference (VI, Blei et al., 2017). The basic idea is to look only for an approximation of the target distribution that is easier to handle computationally:

$$p(\boldsymbol{x}|\boldsymbol{y}) \approx p_{\boldsymbol{\theta}}(\boldsymbol{x}|\boldsymbol{y}).$$

The parameter $\boldsymbol{\theta}$ may represent mean and covariance of a Gaussian distribution (as in the so-called “mean-field” approximation), or – as in our case- weights and biases of a powerful neural network, the optimization of which requires some notion of “distance” in the distribution space. Such deep-learning-based VI can provide expressive posterior representation together with computational feasibility.

In Rizzuti and Vasconcelos, 2024, we presented a variational inference approach where the posterior distribution is represented by an *invertible* neural network (also known as *normalizing flows*, Kobyzev et al., 2021). This class of generative models is especially well-suited for uncertainty quantification in seismic inverse problems: it allows a straightforward maximum likelihood estimation principle to optimize for the full candidate posterior distribution, enables training 3D networks at scale with cheaper requirements on HPC systems (compared to, e.g., VAEs, GANs, or diffusion models), and maintains fast inference (sample generation). In this paper, we apply this framework to UHR field data.

Field data analysis from the German North Sea

We analyze field data collected in the German North Sea for offshore windfarm development, focusing on a selected portion of a 2D seismic line located approximately 160 m from a projected CPT well-log site (see also Telling et al., 2024, for more information on the preprocessing workflow).

The background acoustic impedance model is derived directly from the migration interval velocity using Gardner's relation. A zero-phase wavelet was estimated by matching the background acoustic impedance with the post-stack seismic trace corresponding to the projected well-log location. The post-stack data and the background acoustic impedance model are shown in Figure 1.

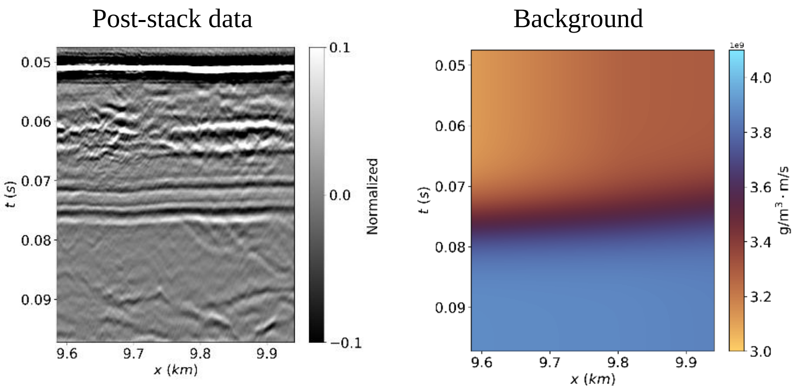

Figure 1 *Seismic post-stack data and background acoustic impedance.*

The results of the uncertainty quantification are represented by a stochastic generative network where the inputs to the model are randomly chosen. So, in principle, its most complete characterization is by means of output samples, as shown in Figure 2. While this may provide a qualitative sense of the posterior variability, summary statistics such as point-wise mean or covariance help simplify the analysis more succinctly.

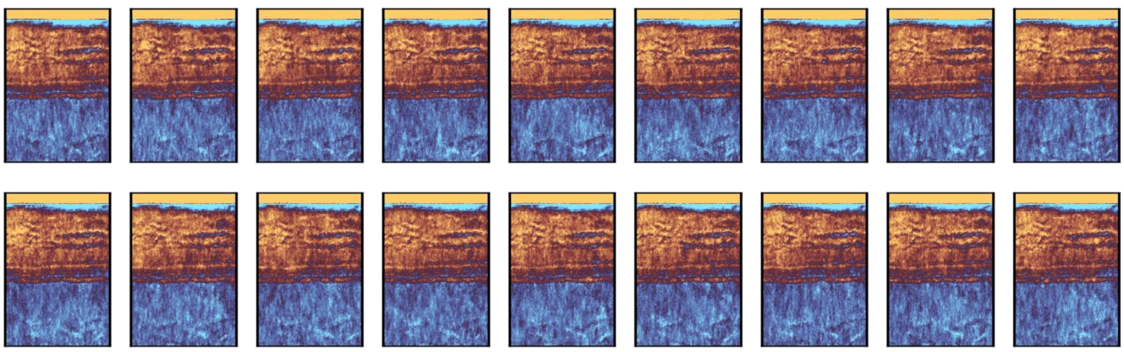

Figure 2 *Random samples from the estimated posterior distribution for post-stack inversion from which a qualitative sense of the posterior variability can be gleaned.*

In Figure 3, we display a conventional way to summarize the posterior distribution in terms of conditional mean, standard deviation, and covariance. These are computed from random posterior samples, and, for Gaussian prior and likelihood, they even provide a complete characterization of the posterior. However, this characterization is only partial when these conditions are not met.

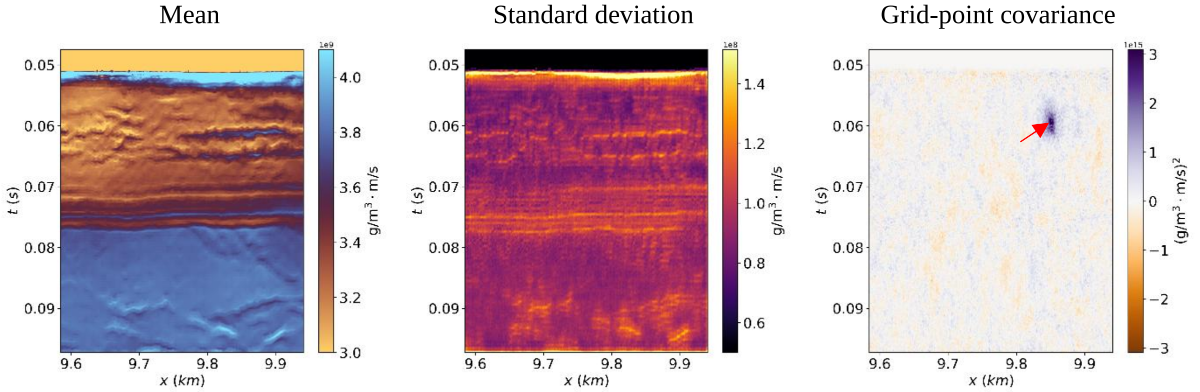

Figure 3 *Point-wise conditional mean, standard deviation, and grid-point covariance obtained by averaging the acoustic impedance samples generated by a neural network (see Figure 2). The grid-point covariance panel highlights in violet the spatial locations that are positively correlated with a grid-point of choice (indicated by the red arrow). The correlation region is mostly determined by the chosen regularization, which consists of structural smoothing.*

An alternative representation of the posterior distribution is by means of credible intervals, which provide ranges of the acoustic impedance parameters with a prescribed level of “confidence”. As an example, the inter-quartile and inter-decile intervals are presented in Figure 4. For imaging problems, we display those as a pair of images, each of which is the point-wise endpoint of said interval. For risk-averse applications, conservative ranges (such as the inter-decile interval) may be more indicated. Their use must be, in general, tailored to the problem and needs at hand.

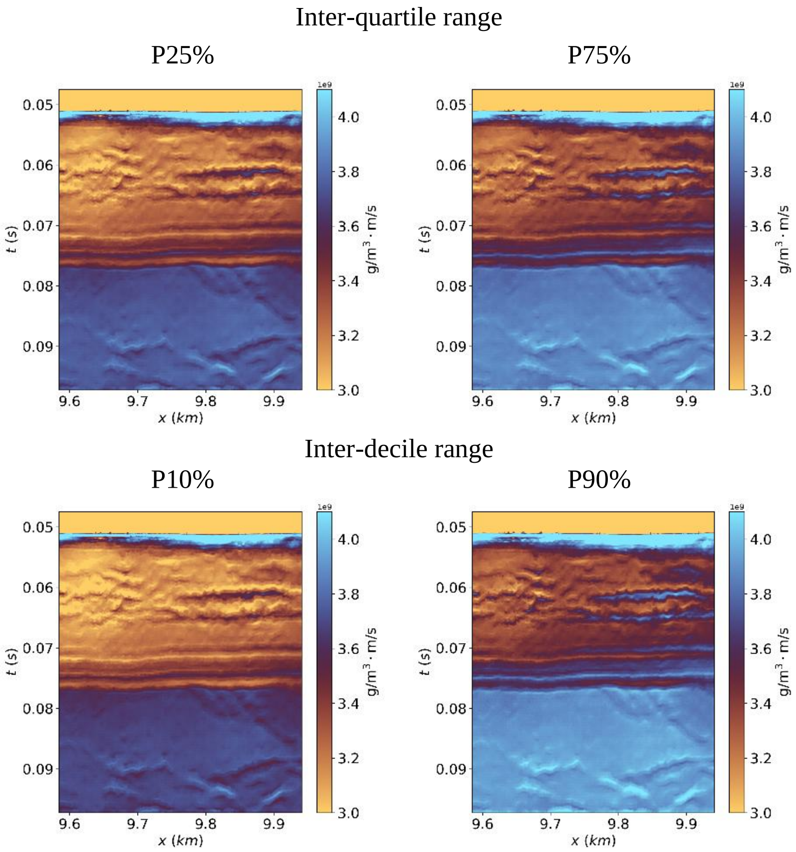

Figure 4 *Examples of credible intervals for the acoustic impedance: the inter-quartile and inter-decile ranges. Each of these is represented by a pair of values (images) representing the endpoints.*

Future directions for probabilistic UHR quantitative inversion

The deep-learning workflow presented here offers a flexible and computationally efficient approach for uncertainty quantification in UHR post-stack seismic data. This method holds significant potential to inform decision-making processes in offshore windfarm development. While our example focuses on a 2D seismic line, the approach is inherently scalable to 3D post-stack inversion problems. Although UHR data acquisition is currently predominantly limited to 2D, the proposed framework is well-prepared for future advancements in 3D acquisition technologies or the use of more complex imaging such as FWI.

A key challenge in a quantitative inversion of UHR data remains the integration of CPT or seismic-CPT well-log information, particularly for accurate wavelet calibration and background acoustic impedance estimation. Seismic-CPT measurements often provide well-log data sampled at lower resolutions than seismic data, with geo-mechanical properties like density inferred indirectly. Additionally, seismic data characteristics and local measurements can exhibit significant variability over short distances, even within tens of meters from a reference 2D line. Looking ahead, we aim to refine the quantitative analysis of UHR data by incorporating seismic-CPT well-logs and accounting for wavelet effects in the uncertainty framework. This methodology can also be readily extended to elastic impedance inversion.

Acknowledgements

We acknowledge the German Federal Maritime and Hydrographic Agency (Bundesamt für Seeschifffahrt und Hydrographie, BSH) for providing the data analyzed in this study.

References

Blei, D.M., Kucukelbir, A. and McAuliffe, J.D. [2017] Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, **112**(518), 859–877.

Klinkvort, R.T., Sauvin, G., Dujardin, J., Griffiths, L., Vardy, M.E., Vanneste, M., 2024, *Cone Penetration Testing prediction using seismo-acoustic data*, *85th EAGE Annual Conference & Exhibition*, Extended Abstracts

Kobyzev, I., Prince, S.J.D. and Brubaker, M.A. [2021] Normalizing Flows: An Introduction and Review of Current Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**(11), 3964-3979.

Rizzuti, G., and Vasconcelos, I. [2024] Multiscale Uncertainty Quantification for Post-Stack Seismic Inversion with Wavelet Flows, *85th EAGE Annual Conference & Exhibition*, Extended Abstracts

Robert, C.P., and Casella, G. [2010] Monte Carlo Statistical Methods. *Springer Texts in Statistics*

Telling R., JafarGandomi, A., and Perratt, O. [2024] Towards quantitative interpretation of UHR marine seismic: an example from the North Sea, *85th EAGE Annual Conference and Exhibition*, Extended Abstracts