Large-scale industrial deployment of machine learning workflows for seismic data processing

Julien Oukili[1]*, Jyoti Kumar[1], Jon Burren[1], Steve Cochran[1], Martin Bubner[1], Denis Nasyrov[1] and Bagher Farmani[1] discuss the benefits of implementing deep neural networks for certain steps of seismic data processing on data examples from around the world.

Introduction

Seismic data processing is often thought of as a non-deterministic journey where signals are incrementally separated from noise, with the inherent challenge of the noise being, in a lot of cases, distinctly different and separable from the signal but in some cases remarkably similar and difficult to distinguish. At the numerous steps of data preparation, migration and post-processing, the opportunities to improve quality are many, and so are the risks of harming the desired signal. Even evaluating the results remains a challenge, especially in the early steps of an imaging flow. Machine learning (ML) applications have caught the interest of many, with the desire for faster and more thorough quality control (QC), more reliable processes, or simply to automate some of the more mundane and highly repetitive tasks of the geophysicists working on increasingly larger amounts of data. At the same time, the accelerated energy transition has put increased pressure on geophysicists to get the most out of each seismic dataset which are often used for multi-purpose subsurface investigations (e.g., both oil and gas exploration and carbon storage screening).

In recent years many case studies have demonstrated the potential of ML methods for processing, QC and interpretation of seismic datasets. However, often these have come with caveats about the use of the results, especially if the actual data being processed significantly differs from those used for training the ML algorithm. In this paper, we look at two different types of machine learning use cases where the neural network workflows have been deployed to many marine seismic processing projects, at large scale, and with special attention given to the challenge of highly varying data characteristics, geological and geographical environments. The application examples are focused on problems where signal and noise separation were critical, either for robustness of subsequent processing steps or for the quality of direct interpretation efforts.

Conventional data processing flows broadly follow a strategy of 1) parameter testing, often on a very limited subset of data, 2) production set up on the full dataset, and 3) QC, which potentially reveals the needs to change or even rerun current processing steps with different parameterisation. Often the final parametrisation choice represents a compromise that is reached after a considerable amount of time has been spent on trying to optimise the said processing steps on a limited assessment of the data.

Using ML methods has allowed us to speed-up specific processing steps and has allowed the geophysicists to concentrate on improving the resulting data quality rather than spending time on optimising processes and parameters.

We are sharing some of our main learnings from routinely using ML methods in a number of specific applications scenarios over the last two years.

Use case 1 – Noise removal in raw data prior to wavelet processing

The first use case focuses on the very early stage of seismic data processing: denoising of raw recorded data. Applying sufficient denoise prior to the first multi-channel processes is crucial to avoid spreading noise or enhancing it, making noise more apparent to desired signal and difficult to remove. Noise characteristics can also vary greatly between seismic surveys, as well as within surveys, sail lines, individual shot records and between the different types of recording sensors in the case of multi-sensor acquisition systems.

Pressure and particle motion data from multi-sensor streamer acquisition can be combined to separate the wavefield into up- and down-going components (Carlson et al., 2007). To guarantee the generation of high-quality up-going and down-going wavefields, the noise from both records must be attenuated before the data are combined. ML tools have increasingly become the method of choice in a drive to increase automation and improve output consistency.

Farmani et al. (2023) presented specific workflows for pressure and particle motion data that employ deep learning to suppress the noise in the records efficiently. Real image denoising network (RIDNet), a convolutional neural network, sits at the core of the proposed workflows. The method, presented here, uses a single RIDNet model with a specific structure for both pressure and particle motion recordings, as opposed to machine learning-based workflows presented by Farmani and Pedersen (2020a; 2020b; and 2022), that attenuate incoherent noise in

[1] PGS
* Corresponding author, E-mail: julien.oukili@pgs.com
DOI: 10.3997/1365-2397.fb2023101

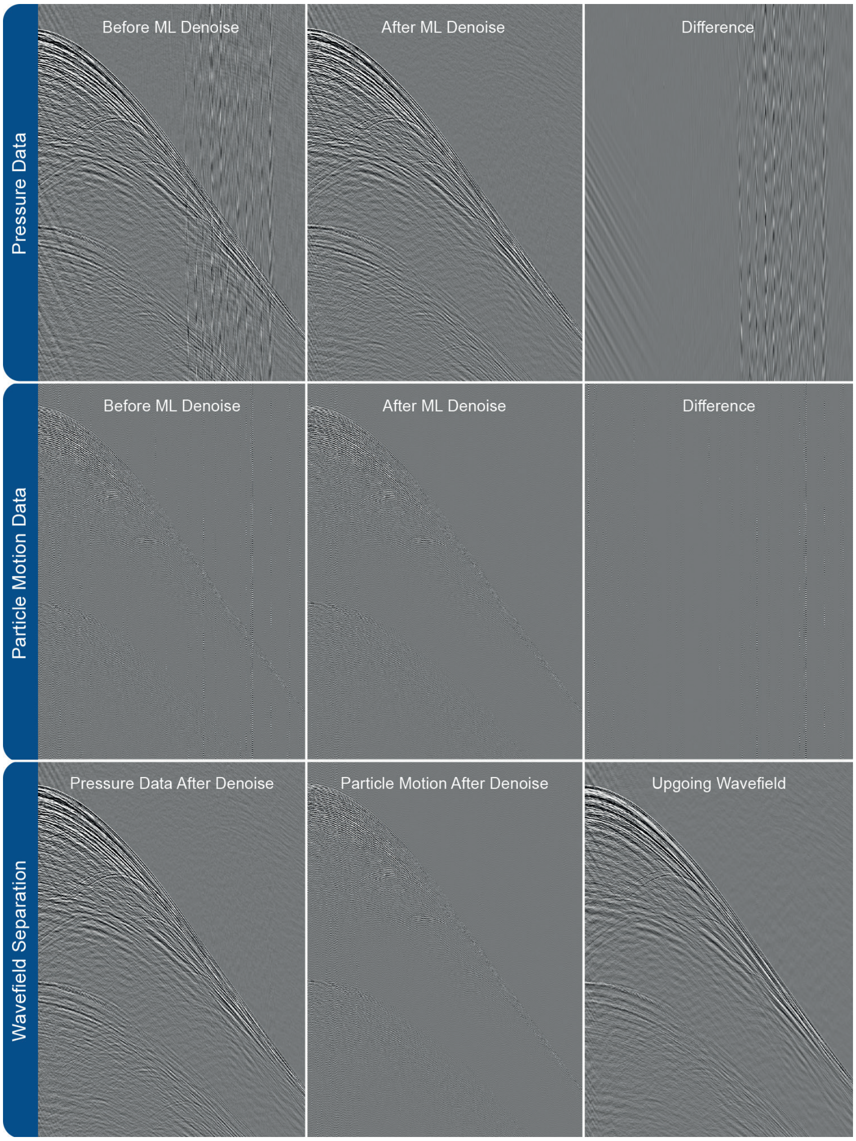

Figure 1 Shot gather from deep-water Eastern Mediterranean, offshore Egypt: RIDNet application on hydrophone data (top), particle motion sensor data (middle) and wavefield separation step (bottom). The raw hydrophone data are contaminated with towing noise (linear events at near offsets, left side) and turn noise at far offset (right side). The particle motion data have a 10 Hz low cut filter and initially shows visible noise towards the far offsets. The denoise was very stable in high- and low signal-to-noise areas. Furthermore, the difference of the particle motion data shows noise being removed at near offsets where signal and noise seem equally strong in amplitudes as well.

the bandwidth where most of the noise exists. As a result, the core components of the processes are greatly streamlined and strikingly comparable for the two types of sensor recordings. A network built on the RIDNet architecture is used to attenuate incoherent noise on pressure and particle motion records within the bandwidth of interest. Additional processing techniques can be used to attenuate noise that is present outside of the RIDNet application bandwidth.

The approach presented has been successfully applied to numerous datasets worldwide with high consistency and higher efficiency and has enabled us to produce denoised data for both pressure and particle motion sensors very quickly after the data have been acquired offshore.

Successful noise attenuation using RIDNet applications

The first field data example is from a 2023 seismic survey, part of multi-client campaign in the Eastern Mediterranean Sea, offshore Egypt. The data were acquired using a triple-source configuration and 12 10,000m-long multi-sensor streamers separated by 150 m. Although the sea was generally calm during the acquisition, some sail lines contained significant noise. Shot gathers from one of the noisy sail lines are shown in Figure 1. The top image displays the results of applying machine learning denoise to the pressure dataset using the RIDNet methodology. It is evident that noise has been efficiently attenuated, and the difference plots demonstrate that there has been no primary leakage during the process. The outcome of the particle motion dataset is shown in the middle image. Finally, the bottom row shows the upgoing wavefield generated through the wavefield separation process using pressure and particle motion input after ML denoising, as expected, without receiver ghost interference and without artifacts.

A second example is shown in Figure 2. This dataset was acquired in 2023 as part of multi-client campaign offshore Malaysia. Again, a triple-source configuration was used and 12 9000m-long multi-sensor streamers separated by 112.5 m. As observed in the previous example, the machine learning denoise workflow effectively attenuates the noise in pressure as well as particle motion datasets, which in turn manage to produce a high-quality upgoing wavefield. Even if minor residual noise can still be observed, it has not deteriorated through the wavefield separation process and hence can be dealt with effectively at a later stage.

The final example is from a dataset acquired in 2015 using also multi-sensor streamers in the Faroes Shetland Basin, offshore UK. This survey was acquired using a dual-source configuration and 10 7050m-long cables separated by 100 m. The data were reprocessed in 2023 as part of a regional multi-client rejuvenation project. Machine learning denoise once more worked admirably for both the pressure and particle motion datasets (Figure 3), demonstrating that even older datasets can benefit from the new denoising technology.

The machine learning denoise workflow described in these use-case examples has shown good and robust performance in attenuating common noise on datasets from all around the world and enabled generation of high-quality separated wavefields. And since the presented method using RIDNet is largely automated it only took a few days on each project to set up the respective wavefield separation workflow and apply it to several thousand square kilometres reliably despite the high variations in input data characteristics.

A similar RIDNet workflow based on a lighter network, more focused on signal and noise classifications, is currently being trialled to identify residual noise which might require further attention and investigation. Design work is continuing that should allow us to graphically visualise the results of this new residual noise detection workflow so that geophysicists can make quick and informed decisions.

Use case 2 — Noise attenuation in the post-migration image domain

In the following section we shift our attention from raw data denoise to the late stages of a standard seismic data processing flow, post-migration image denoise, where the starting assumption is fundamentally different from the previous application since the desired signal is expected to be already in focus (i.e., adequately imaged).

Seismic images are often contaminated by migration noise, sometimes referred to as migration smiles, swings, artifacts or defects. This noise is generated when assumptions made by the migration algorithm begin to break down, e.g., the midpoint, offset and azimuth sampling requirements of the input data are

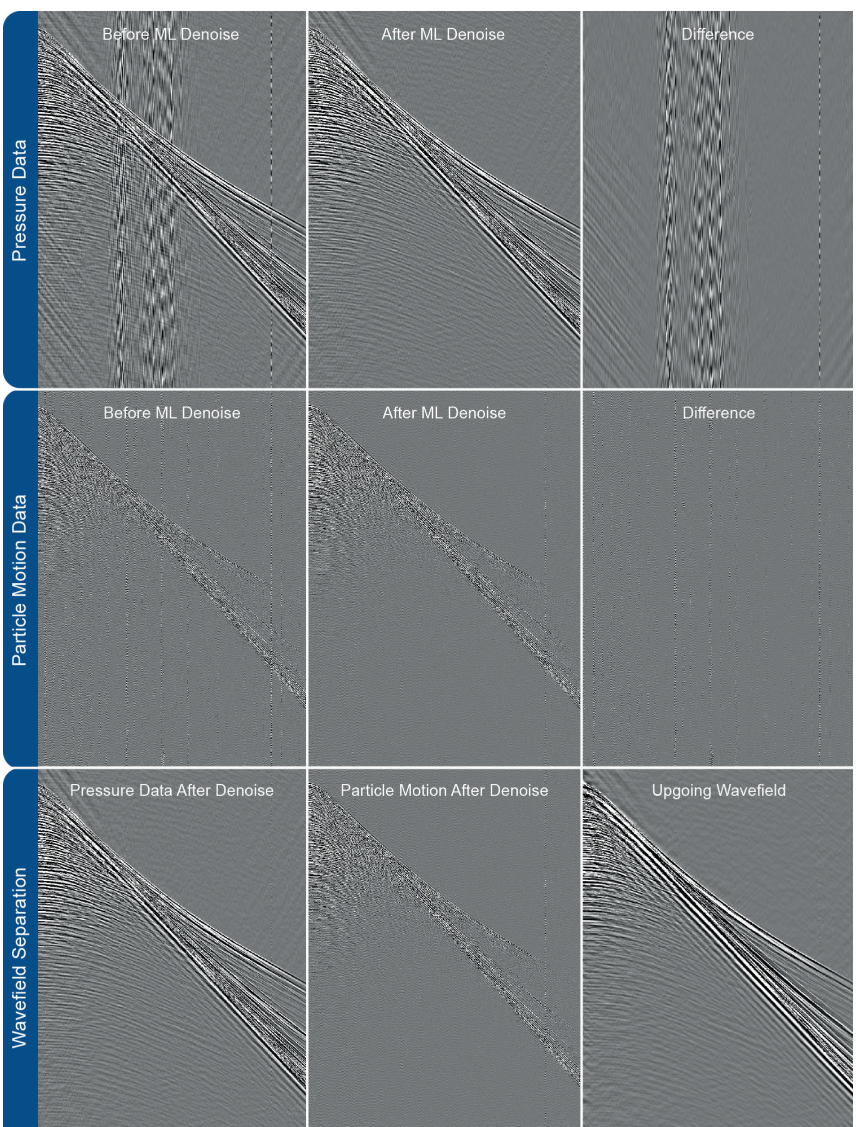

Figure 2 Shot gather from shallow water Sarawak, offshore Malaysia: RIDNet application on hydrophone data (top), particle motion sensor data (middle) and wavefield separation (bottom). Compared to the examples on Figure 1, the raw hydrophone data show more towing noise at far offsets (right side) and spurious noisy traces. The particle motion data show more noise in the form of vertical stripes, likely to be caused by more active birds (instruments which steer the streamers).

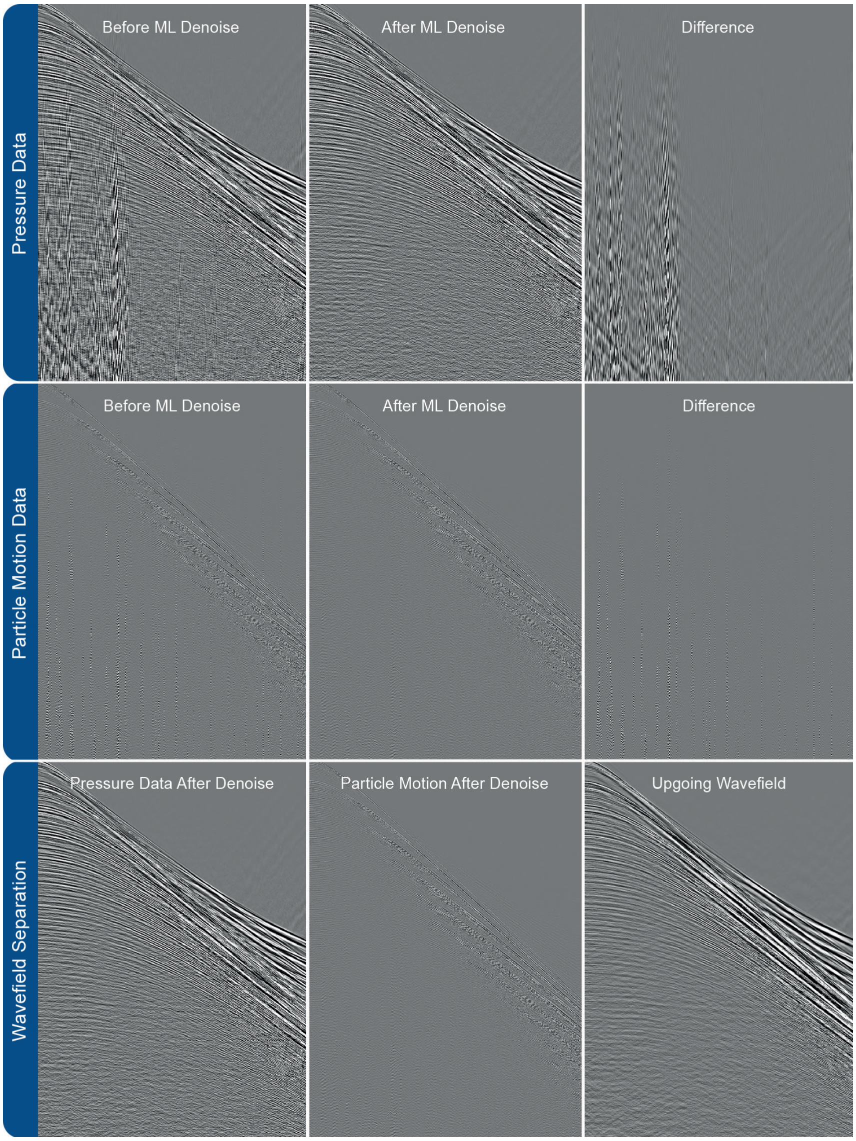

Figure 3 Shot gather from Faroe Shetland Basin, offshore United Kingdom: RIDNet application on hydrophone data (top), particle motion sensor data (middle) and wavefield separation (bottom). Very strong noise is present in the raw hydrophone data, as well as weaker linear noise, both well attenuated through the denoise, which reveals a lot of signal. The amplitudes of the noise on the particle motion data are also higher than in the examples of Figure 1 and 2.

not adequate (Long et al., 2006). The assumptions may break locally when the seismic wavefield propagates through complex media or is exposed to strong amplitude (reflectivity) contrasts. Data regularisation and filtering of aliased energy prior to migration can mitigate effects of the sampling challenges (e.g., Schonewille 2000; Chemingui and Biondi, 2002). The effects of under-sampled field data can be further compounded by subsurface complexity.

Preserving the amplitudes accurately is a key objective of any imaging exercise and care needs to be taken not to alter amplitudes in any noise removal process especially at the post-migration stage. Traditionally, applying any noise removal post migration has taken a lot of time and effort.

Klochikhina et al. (2021) described a machine learning method for tackling this problem. In their paper they demonstrate how a convolutional neural network (CNN) with U-net architecture was trained using synthetic data examples and the results demonstrated on some field data examples.

This ML-based denoise technique has now been widely adopted and applied to a wide range of different datasets as part of commercial imaging projects. It has proved very effective in removing or reducing migration noise, suggesting the training data were sufficiently representative.

Application scenario 1 — Structural Imaging and post-migration clean-up

A field data example from offshore Newfoundland demonstrates the ability of the neural network to attenuate noise on the migrated section, making the image more easily interpretable and improving the structural picture (Figure 4).

A second field data example from offshore Norway (Figure 5) represents a more challenging scenario where some of the migration noise, especially from the Top Chalk reflection, does interfere with shallower structures of similar dips in heavily faulted formations. In this application example the CNN noise removal had to be limited to a horizon-bound interval and was complemented by structurally conformable filtering. This resulted in a considerable improvement in signal-to-noise ratio without detrimental image distortion. We postulate that more traditional denoise tools would have failed in preserving the smaller details if they had been designed to achieve a similar noise reduction.

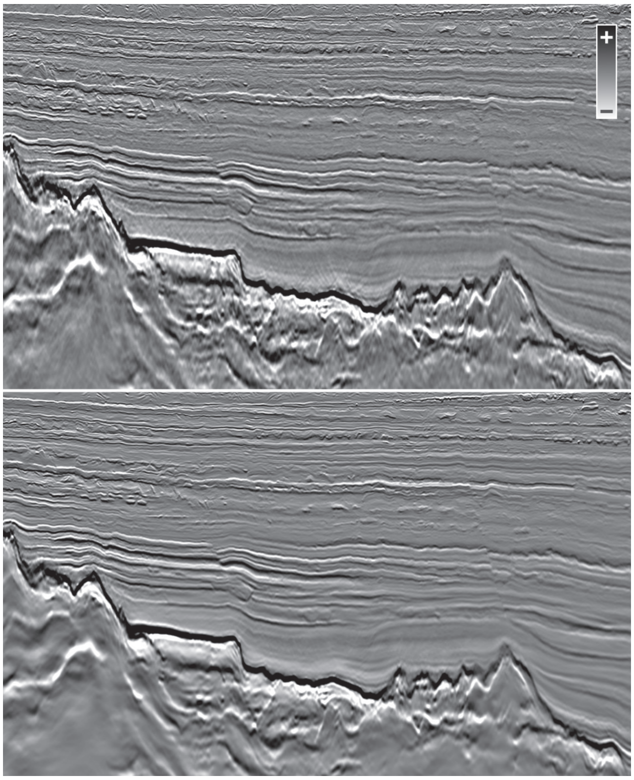

Figure 4 Migrated section before (top) and after (bottom) application of the neural network. The migration artifacts are more prominent above strong interfaces and could be in the worst case interpreted as faults.

Application scenario 2 — Supporting quantitative interpretation (QI) work

In a further example (Figure 6) we illustrate how the CNN denoising capabilities positively impact on QI workflow. The signal-to-noise ratio of the data influences the stability of any attributes derived from the data, so a typical QI workflow necessitates the inclusion of steps to precondition the pre-stack and stack-domain data to address unwanted noise. This can be a time-consuming process, depending on the quality of the input data. Including the neural network in the data preparation workflow provides the required noise suppression, to the benefit of the QI attribute derivation. Figure 6 shows the impact on the AVA gradient attribute; the level of noise is clearly reduced while the signal continuity is maintained. Considerable uplift is also seen on the relative P-wave impedance attribute computed from the same data.

Application scenario 3 — High-end imaging using least squares migration workflows

Incorporating the same neural network methodology into high-end imaging workflows has helped overcome challenges with the migration noise frequently observed in least-squares migration (LSM) results. LSM solutions can be divided into two main categories: data-domain and image-domain solutions. Both approaches are powerful techniques for overcoming challenges with subsurface illumination and image blurring and ensure more reliable amplitude preservation in the imaging process,

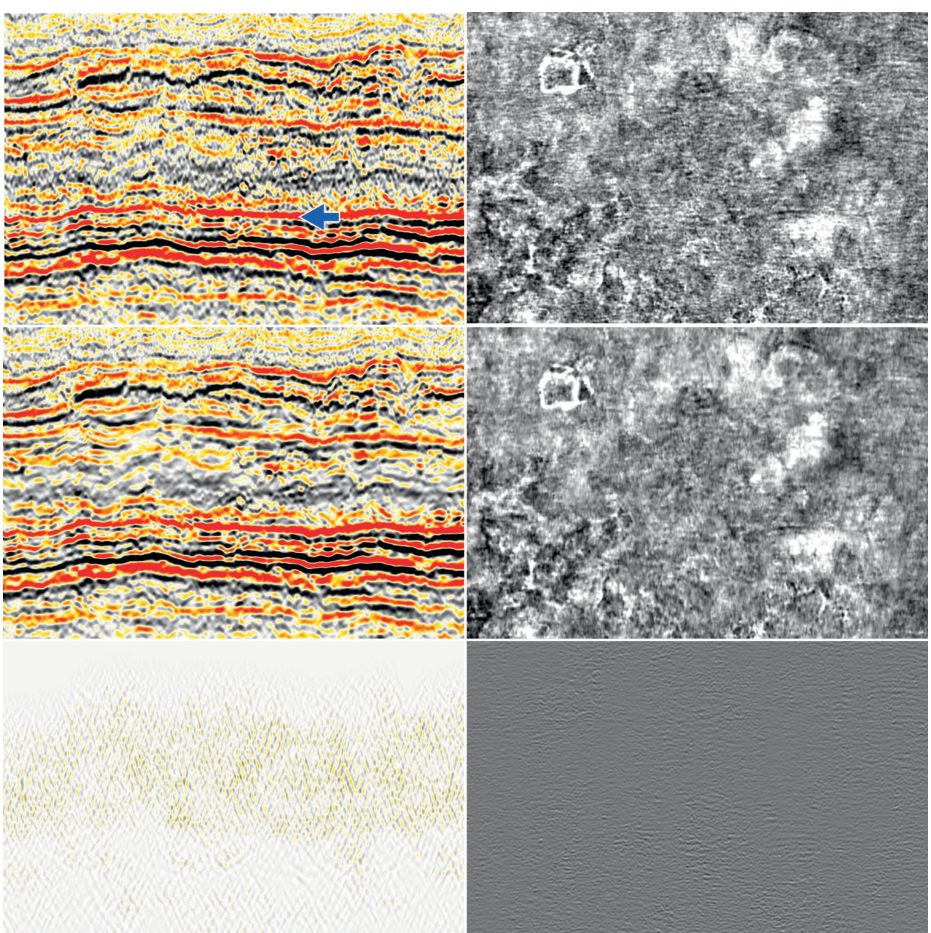

Figure 5 Migrated cross-section (left column) and time slice (right column), before (top row) and after (middle row) targeted image denoise and difference (bottom row). Contrary to the example of Figure 4, the migration artifacts above the Top Chalk (blue arrow) are not easily distinguished from the shallower complex geology. However, the artifacts have a more distinct lineation pattern on time slices which is taken just above the Top Chalk. Features which are parallel to the noise pattern on time slices have been well preserved.

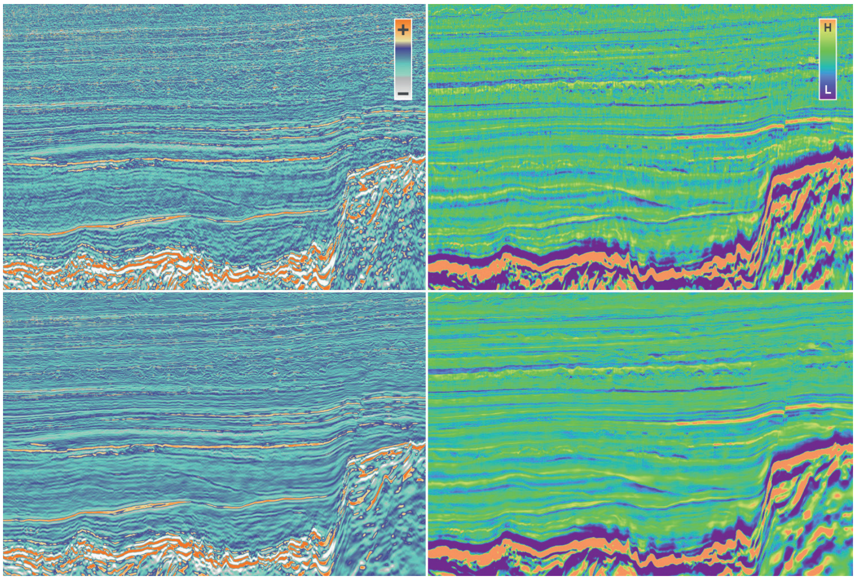

Figure 6 Left side: AVA gradient attribute calculated on data before (top) and after (below) migration noise attenuation applied to angle-stacks, using the neural network. Right side: relative Ip attribute calculated on data before (top) and after (below) migration noise attenuation applied to angle-stacks, using the neural network.

important for AVA studies. Data-domain solutions have a lot in common with full waveform inversion (FWI), involving forward modelling, followed by comparison between modelled and observed data and the back-propagation of residuals, resulting in the estimation of reflectivity instead of velocity. This approach to LSM can prove computationally expensive, especially for higher frequencies.

Image-domain LSM solutions derive and apply corrections to a conventional migration. If the signal-to-noise ratio of the conventional migration is poor, the LSM result will be noisy, therefore appropriate data preconditioning and regularisation (as part of the inversion) is critical for this methodology.

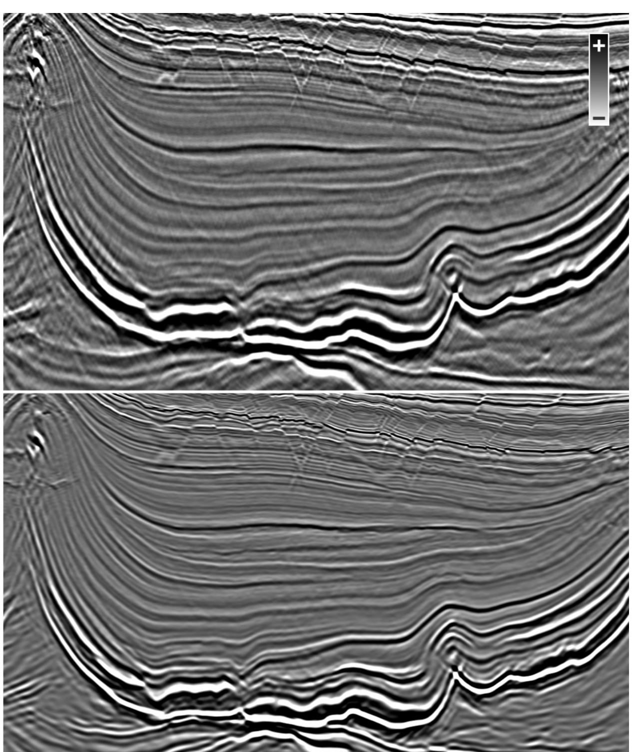

Figure 7 Raw migration (top) showing clear evidence of migration noise. The output of the least-squares migration workflow (bottom), incorporating the neural network, shows improved horizontal and vertical resolution with minimal contamination from the migration noise on the raw migration.

The example from the Campos Basin, Brazil (Figure 7) shows the results of an image-domain LSM. The inclusion of the neural network-based denoise in the least-squares workflow has ensured that the enhanced image is not contaminated by migration noise.

Some takeaways for post-processing of seismic images

The application examples above illustrate the potential of an uplift in image quality that benefits the subsequent processing and use of the data. An equally important consideration is the time needed to address the migration noise; using the neural network significantly reduces the preparation time compared to more traditional denoise methods. The user is considerably less encumbered with testing parameters and tuning workflows using the neural network, freeing time which can be spent giving greater attention to more advanced geophysical processes and analysis of QI attributes extracted from the data.

Use case 3 — Diffraction event detection

In the pre-migration domain, diffraction events, sometimes referred to as ‘tails’, exhibit similar characteristics to the ‘migration smile’ artifacts discussed earlier, albeit in a downward dipping fashion. Building on the similarity of these two artifact types, the same CNN U-net architecture as featured in use case 2 above was used but with an inverted time axis (i.e., free surface pointing downward on seismic traces). Contrary to the earlier example, here the diffraction ‘noise’ is actually desired signal, which can be re-inserted into the migration process to produce a sharper image of small-scale heterogeneities.

In Figure 8, a near-offset volume is taken through migration before and after the diffraction identification and separation. The same migration algorithm and velocity models were used, so the images are fully consistent, although one may argue that focusing diffraction tails would require further attention. As expected, the separation is not perfect, as reflection leakage can be observed

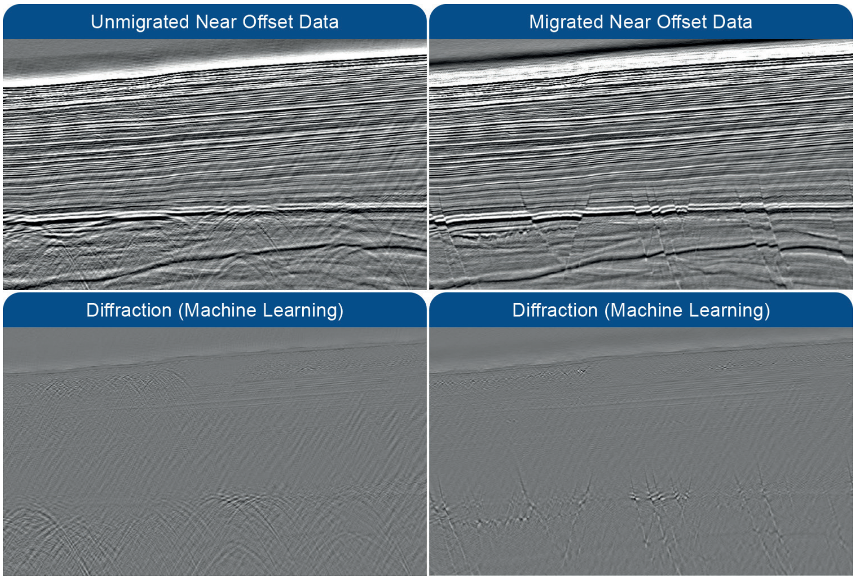

Figure 8 Near offset volume from shallow water offshore Newfoundland and Labrador, Canada: conventional imaging (top) versus diffraction imaging (bottom). The workflow successfully identified both large- and small-scale diffraction patterns, though some weak reflection leakage can be seen. The separation is most impressive towards the shallow (top of the images) where strong reflections were masking weaker diffraction tails.

in the diffraction products. However, the features highlighted by the new image now appear much stronger than the background geology, especially towards the top of the section where highly reflective horizon events were masking smaller details.

While the section views are very useful for understanding large fault patterns, we find the diffraction images to be rather chaotic and still noisy for small-scale features. Looking at time slices of the same migration volume (Figure 9) reveals patterns and occasionally isolated objects. The benefits of such an approach are significant when the diffraction generators are located close to reflections of similar or higher amplitudes. Furthermore, the separation, although not perfect, does not seem to suffer from amplitude footprint effects and can resolve details down to the scale of the imaging bin size.

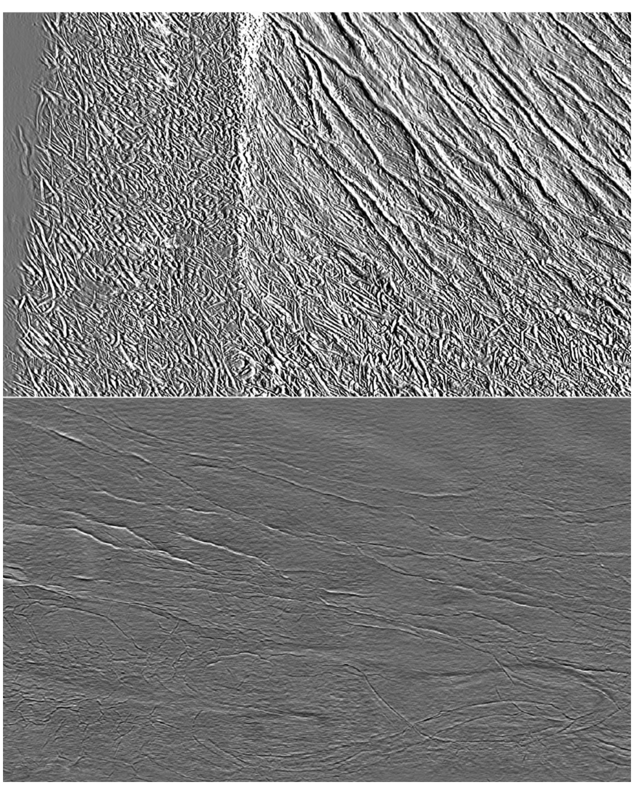

Figure 9 Slices through the diffraction image volume: seabed (top) and 1368 ms under seabed (bottom). Glacial features are dominating the seabed image except towards the left side of the image: the water bottom goes rapidly deeper and is therefore practically free of iceberg marks. The deeper image shows both a large network of sub-parallel faults as well as small scale polygonal faulting in the lower left side of the image.

The method presented here is now being tested on various datasets to aid detailed interpretation for both deep and shallow targets, and using conventional as well as high-resolution and ultra high-resolution 3D seismic data (with approximately 1 m x 1m bin size and 0.25 ms temporal sampling).

It is still early days, but so far the machine learning CNN U-net-based workflow has robustly worked for any type of 3D seismic data and any type of frequency bandwidth that it was applied to.

Conclusions

We have discussed several use cases of ML methods in data processing steps which are always required on any type of seismic data nowadays. The examples shown have proven our implementations to be reliable for a large variety of datasets, nearly irrespective of the region of origin and of the geophysical and geological settings. The ML workflows have either replaced complete steps or been integrated with other tools to achieve at the very least the same quality as before, but most often better.

Running machine learning workflows can be compute intensive. However, access to cloud computing has all but removed any computational limits and enabled large-scale deployment of ML technology. The main benefit of employing machine learning technology is that it has allowed us to free up valuable time of the project geophysicists by shifting more of the denoising effort to the computer. This has more broadly allowed the processors to spend time on QC and quality improvements, leading to better project outcomes.

The methods described in this paper incorporate a significant amount of automation. However, the role of geophysicist remains crucial, in selecting the appropriate flows, defining the application range, adding adequate complementary processes, or simply discarding the results in parts of a dataset. It is important that using ML technology does not amount to black box processing and that the application user remains in full control. A natural question to ask is how far we can progress towards a fully automated data processing workflow and how quickly and whether experienced geophysicists will be replaced by machines (Brittan et al., 2021).

We believe that we are still a long way away from fully automated seismic data processing and that with the continued changes to the type and characteristics of the data we record and the constant evolution of high-fidelity final imaging products, geophysicists will continue to play a critical role in executing successful imaging projects. We have demonstrated in this paper that in specific areas of data processing, fully industrial and dependable implementations of ML technology can ultimately benefit both data providers and data recipients.

Acknowledgement

The authors would like to thank PGS for permission to publish this paper.

References

Bekara, M., and Day, A. [2019]. Automatic QC of denoise processing using a machine learning classification. *First Break*, **37** (9), 51-58.

Brittan, J., O'Driscoll, R., Walpole, J. and Cobo Y. [2021]. Root Butlers, Flying Cars, Automated Seismic Processing and Imaging – How near are we? *First EAGE Workshop on Optimizing Project Turnaround Performance*, February 2021, Volume 2021, p.1-5.

Carlson, D., Long, A., Söllner, W., Tabti, H., Tenghamn, R. and Lunde N. [2007]. Increased resolution and penetration from a towed dual-sensor streamer, *First Break*, **25**, 71-77.

Chemingui, N., and Biondi, B. [2002]. Seismic data reconstruction by inversion to common offset. *Geophysics*, **67**, 1575-1585, doi: https://doi.org/10.1190/1.1512803

Farmani, B. and Pedersen, M.W. [2022]. Stepping Towards Automated Multisensor Noise Attenuation Guided by Deep Learning. *83rd EAGE Annual Conference & Exhibition*, Jun 2022, Volume 2022, p.1-5.

Farmani, B., Pal, Y., Pedersen, M.W. and Hodges, E. [2023]. Motion sensor noise attenuation using deep learning. *First Break*, **41**(2), 45-51.

Farmani, B., Lenses, M. and Pal, Y. [2023]. *Multisensor Noise Attenuation with RIDNet*. 84th EAGE Annual Conference & Exhibition, Jun 2023, Volume 2023, p.1-5

Klochikhina, E., Crawley, S., Frolov, S., Chemingui, N. and Martin T. [2020]. Leveraging Deep Learning For Seismic Image Denoising. *First Break*, **38**(7), 41-48.

Klochikhina, E., Crawley, S., and Chemingui, N. [2021]. *Seismic image denoising with convolution neural network.*, First International Meeting for Applied Geoscience & Energy

Long, A., Fromyr, E., Page, C., Pramik, W., Laurain, R., and Buchan, I. [2006]. *Multi-Azimuth and Wide-Azimuth lessons for better seismic imaging in complex settings.* SEG 76th Annual Meeting, New Orleans.

Schonewille, M.A. [2000]. *Fourier reconstruction of irregularly sampled seismic data*: Ph.D. dissertation, Delft University of Technology