

Attenuation of seismic migration smile artifacts with deep learning

Jewoo Yoo, Paul Zwartjes *

Aramco Overseas Company B.V., the Netherlands



ARTICLE INFO

Keywords:
Seismic data processing
Noise attenuation
Migration
Deep learning
U-net

ABSTRACT

Attenuation of migration artifacts on Kirchhoff migrated seismic data can be challenging due to the relatively low amplitude of migration artifacts compared to reflections as well as the overlap in the kinematics of reflection and migration smiles. Several ‘conventional’ filtering methods exist and recently deep learning based workflows have been proposed. A deep learning workflow can be a simple and fast alternative to existing methods. In case of supervised training of a deep neural network using training data made by physics-based modelling or actual migrations is expensive and lacks diversity in terms of noise, amplitude, frequency content and wavelet. This can result in poor generalization beyond the training data without re-training and transfer learning. In this paper we demonstrate successful applications of migration smile separation using a conventional U-net architecture. The novelty in our approach is that we do not use synthetic data created from physics-based modelling, but instead use only synthetic data build form basic geometric shapes. Our domain of application is the migrated common offset domain, or simply the stack of the pre-stack migrated data, where reflections resemble local geology and migration smiles are upward convex hyperbolic patterns. Both patterns were randomly perturbed in many ways while maintaining their intrinsic features. This approach is inspired by the common practice of data augmentation in deep learning for machine vision applications. Since many of the standard data augmentation techniques lack a geophysical motivation, we have instead perturbed our synthetic training data in ways to make more sense for a signal processing perspective or given our ‘domain knowledge’ of the problem at hand. We did not have to retrain the network to produce good results on the field dataset. The large variety and diversity in examples enabled the trained neural network to show encouraging results on synthetic and field datasets that were not used in training.

1. Introduction

Artifacts in migrated seismic data can be caused by acquisition geometry, noise or non-primary wavefields in the data, the migration algorithm itself or errors in the velocity model. Abma et al. (2007) discuss artifacts generated by the Kirchhoff migration algorithm and the effects of irregular azimuth distributions and holes in the acquisition. Aliasing and non-uniform sampling both cause artifacts during seismic data processing. Other sources of errors are listed in Ongkiehong and Huizer (1987), who state that the most important error in seismic acquisition is the error due to binning as a correction for non-uniform sampling. Although Kirchhoff migration will accept irregularly spaced traces, artifacts occur when the sampling is not regular in all dimensions. The obvious solution is to acquire more data, but this is not always practical or economical. Nemeth and Schuster (1999) discuss how least-squares migration can help reduce Kirchhoff migration artifacts due to irregular sampling. At the time, this was prohibitively expensive and

Duijndam et al. (2000) discuss how applying seismic data regularization (a.k.a. seismic data reconstruction) followed by conventional migration gives similar results, at least as far as reducing imaging artifacts. Correct handling of the positions is important as it can reduce artifacts in noise removal, multiple attenuation and prestack migration and, therefore, has been at the forefront of seismic data processing conferences for close to two decades. There are various seismic data reconstruction methods, of which the transformation-based reconstruction approach has proven to be the most efficient and robust. Several conventional transforms have been evaluated such as the Fourier transform (Duijndam et al., 1999), the parabolic Radon transform (Hugonet and Canadas, 1997) and the hyperbolic Radon transform (Trad et al., 2002). The damage caused by aliasing can be reduced by using appropriate prior information in the inversion, as has been shown by Herrmann et al. (2000) for the parabolic Radon and by Zwartjes and Sacchi (2007) for the Fourier transform. The Fourier domain methods were the first method to be extended in four spatial dimensions for marine (Zwartjes and Gisolf,

* Corresponding author.

E-mail address: paul.zwartjes@aramcooverseas.com (P. Zwartjes).

2006) and land seismic data (Trad, 2008) and then went on to become the de-facto standard for multi-dimensional interpolation under then name anti-leakage Fourier transform (Xu et al., 2010).

With the arrival of faster computers, the focus has returned to least-squares migration, not only to reduce acquisition artifacts but mainly to improve focusing (e.g. Witte et al., 2019).

There are other sources of migration artifacts next to sampling irregularities. Prior to migration an implicit assumption is made that all events are well-behaved primary events with clean amplitudes. Any sharp truncation, time shifts or abrupt amplitude change can produce migration artifacts. Tapering the edges of truncations and discontinuities generally reduces these artifacts. And even when the events are well-behaved and the sampling is regular, artifacts can occur through aliasing of events along the migration operator. This operator aliasing was originally dealt with by limiting the migration aperture, until Gray (1992) introduced frequency-dependent anti-alias operators, which were later improved to preserve steep dips as much as possible (Abma et al., 1999). Aliasing noise manifests itself as ghosts of events deeper in the section that vary in space and time, which complicates their identification. It is not a type of noise we will consider here. The effect of velocity errors in migration is, for instance, addressed in the tutorial by Zhu et al. (1998) who showed how smiles are caused by over-migration with a velocity that is too fast, and frowns are caused by under migration with a velocity that is too low. Even with perfectly optimal acquisition and perfect knowledge of the velocity model, there can be artifacts. For example, due to strong refraction of waves in the velocity model as discussed by Stolk and Symes (2004) or simply due to sub-optimal migration operator settings which leads to operator aliasing.

Two-way wave-equation methods like reverse time migration (RTM) have replaced Kirchhoff pre-stack depth migration in areas of complex geology. RTM also has its share of migration artifacts, especially low frequency artifacts caused by the unwanted cross-correlation of head waves, diving waves, and backscattered waves, which are not included in the imaging condition. The most popular approach for attenuating these artifacts is filtering after application of the imaging condition, see for example (Zhu and Cao, 2019 and Chen et al., 2020). This is also a topic we will not address here.

Apart from expensive solutions (improved sampling, least-squares migration), there are cheaper processing solutions to remove migration artifacts. Hale (2011) proposed a structure-oriented bilateral filter to handle migration noise. Masjukov and Shlyonkin (2015) propose a 6D filtering approach to suppress migration artifacts based on a dip decomposition in common offset volumes and a semblance-type measure computation via offset for all constant dip-gathers. Lou and Simpson (2019) present a dip filter consistent with the local velocity model in VSP migration, which preserves true dip structures and suppresses migration artifacts.

Recently, also deep learning-based methods for migration noise filtering have been proposed. In this paper we will investigate that approach further.

2. Deep learning and migration smile attenuation

Some of the proposed methods for migration noise removal are very computationally expensive and/or require careful parameterization or rely on manually designed features. Regardless of the underlying problem, replacing strenuous parameterization and hand-crafted features is exactly what has made deep neural networks successful in many different fields. Deep learning, as a new data-driven technique compared to conventional approaches, has been enthusiastically embraced by researchers in the geophysical community (e.g. Yu and Ma, 2021). One of the most popular convolutional neural networks (CNN) in geophysical applications is the U-net (Ronneberger et al., 2015) with its long-range connections between the encoding and decoding levels that improve reconstruction accuracy. Although most geophysicists will be familiar with the concept of convolutions (which in the neural networks are

actually correlations), the overview given by Dumoulin and Visin (2018) will benefit the reader that is interested in the practical details. The original U-net was used for segmentation of biomedical images and has been used in geophysical segmentation applications such as salt segmentation (Shi et al., 2019), first break picking (Yuan et al., 2020), diffraction recognition (Markovic et al., 2022) and reflection-diffraction separation (Zwartjes and Yoo, 2022).

The U-net design is basically that of an enhanced auto-encoder, which is why it is also used frequently in regression problems, such as deghosting (De Jonge et al., 2022), interpolation (Fang et al. 2021), noise attenuation (Sun et al., 2019) and velocity model building (Yang and Ma, 2019). Klochikhina et al. (2020) and Bugge et al. (2021) have both proposed a U-net style network trained on synthetic and real data to attenuate migration artifacts.

At first glance, it appears that the computational complexity and human interaction has merely been shifted to the neural network training phase. This is because the training requires extensive neural network design choices (number of layers, filters, filters size, objective function, etc.) and hyperparameter choices (learning rate, optimization algorithm, etc.). Unfortunately, there is no theoretically optimal choice for neural networks for a particular problem and therefore this design process is empirical in nature; it mostly comes down to trial and error. The neural network architecture can be tweaked with various elements that can each have a positive impact on the final outcome, as for example shown in Zwartjes and Yoo (2022). However, in practice the differences will be small and incremental. Most likely for this reason the geophysical community has accepted the U-net architecture as sufficient for most practical problems. For this reason, we will not experiment with neural network design and rely on the U-net architecture and acknowledge that incremental improvements will be possible with different designs.

It is well known that neural networks are data hungry and, because they are completely ignorant of the problem at hand, require a very rich a diverse collection of examples in order to learn to distinguish between desired and undesired features. In our view, the key factor to achieve successful separation of migration smiles and proper seismic reflections relies on the dataset used for the training rather than the chosen deep-learning network architecture. We discuss in details how the synthetic training datasets are generated.

In the approach by Klochikhina et al. (2020) a training dataset was constructed by running well-sampled and coarsely sampled migrations. We are not in favor of this approach as it is difficult to make a generalized network for migration smile attenuation in this way as the training dataset has to be similar to the target dataset. This usually means that the network requires re-training before it is applied to the new dataset. Bugge et al. (2021) instead create realistically looking synthetic data and add migration smile noise. We have applied a similar approach based on synthetic data only training for reflection-diffraction separation (Zwartjes and Yoo, 2022) which we will now apply to the attenuation of migration smile noise.

3. Training data

We apply our workflow to migrated common offset sections or simply to the pre-or post-stack migrated stack, where reflections resemble local geology and migration artifacts have an upward convex hyperbolic shape. The aim is to train a neural network that finds and extracts the hyperbolic migration artifacts and suppresses the smoothly varying background reflectivity. In essence, this means that we are following the same philosophy as Fomel et al. (2007), who aimed to remove the smooth reflection data with plane-wave destruction filters for separation (downward convex) diffractions from the migrated common offset sections. The advantage of the deep neural networks is that they are much more powerful than plane-wave destruction filters in performing this task, provided we are able to create a diverse enough training dataset.

The novelty of our approach is that all synthetic events are

constructed from basic geometric shapes. One of the motivations to take this approach is that it simply is more computationally efficient as it does not require any actual seismic data migrations to simulate migration smiles from coarsely sampled data. Another reason is that it allows for more diversity in the training data, beyond what is possible or physically realistic with physics-based modelling. This in turn is motivated by data augmentation results in other fields (Shorten and Khoshgoftaar, 2019; Perez and Wang, 2017) where it is observed that non-physical distortions applied to the input data, such mixing of different images, geometric transformations and random erasing of samples, actually help make the trained neural network better in predicting in data that was not used during training. We expect this approach will also help in this geophysical problem. But instead of applying the common data augmentation techniques applied in the computer vision field, we have applied data augmentation methods that are more meaningful in the field of seismic processing.

The main distinguishing characteristics of migration smiles are their upward convex hyperbolic shape with gradual or steep flanks. However, since reflections can have all or parts of these properties also, the hyperbolic shape cannot be a ‘necessary and sufficient condition’ to distinguish migration smiles. Relying on shape alone yields a trained network that also tends to predict steep slopes as migration smiles signals, even where they actually originate from reflections. To overcome this, we added additional characteristics of migration smiles to the synthetic data. The additional characteristics for the smile are (1) a weak amplitude relative to the reflection, (2) and a location of the apex tangential to the reflection. We hypothesized that a signal can be considered a migration smile when it has these two features plus the upward convex hyperbolic shape. In our training dataset the migration smile aperture and amplitude are varied randomly but consistent with these assumptions.

Table 1 lists the process of creating a single pair of input and label gathers. The input gather is a sum of a reflection-only section and a migration smile section. The label section is just the smile-only section. The process starts with portioning the gather into a number of segments, as Fig. 1 shows. For each of these segments a curvy trend with random statics is defined, as in Fig. 2. Each segment is then ‘filled’ with a reflectivity sequence, as in Fig. 3a. The number and location of events in each segment is random. Each event within a segment is then convolved with a particular wavelet of which the dominant frequency and amplitude are also chosen randomly, but held constant along an event. Random noise is added in 70% of the cases. The time sampling of the training data is 4 ms for all gathers.

The smile-only segment of a section, in Fig. 3b is constructed based on the geometry of the reflection-only segment within that section meaning the artifact apex is positioned at the location of a reflector. At every location along a reflector it is randomly decided whether to place a

Table 1
Recipe for synthetic data creation.

Per gather
• Define number segments [1,4]
• Define reflection-only gather
• Define segmentation partitioning ('faults' and 'unconformities')
• Define structural reflectivity with each segment using linear and sinusoidal trends with small static shifts
• Define location and number of reflections [50–200]
• Convolve with wavelet
• Ricker, Gaussian, 1st der. of Gaussian
• Frequency content per event (dominant $f_M = [5–90]$)
• Amplitude variation per event [0.1–1]
• Add noise to 70% of gathers
• Define migration smile
• Number varies per event
• Amplitude a fraction of reflection amplitude
• Hyperbola opening angle varies per migration smile
• Sum reflection and smiles gathers to create an input gather

migration smile or not. For each smile event that is placed, the wavelet, dominant frequency, opening angle and amplitude are all randomly chosen. That means that each migration smile along a reflector is different. This is not possible with physics-based modelling, and simply not realistic to expect to occur in real data. However, it does meet our objective of geophysically motivated data augmentation. Fig. 4c shows the input to the neural network which consist of the sum of each reflection-only gather (Fig. 4a) and corresponding smile-only gather (Fig. 4b). Note that random noise is only added to the reflection-only sections.

The size of the gathers in Fig. 4 is 256 traces \times 256 time samples. In total, 9000 training and 500 validation gathers were generated in with this randomized workflow.

3.1. Overlap in reflections and migration smiles near the apex

Near the apex of the smile, the reflection and migration smile energy will invariably overlap. Only with true amplitude modelling will it be possible to attenuate the correct fraction of migration smile energy and leave the reflection energy untouched. In practice, this is next to impossible. Our objective is to preserve the apex of reflection events and not remove migration smiles at the expense of migrated reflections. Our approach is to remove the samples of the smile event that are within one half of the wavelet main lobe width between zero-crossings of the migration smile apex time, as shown in Fig. 5.

4. Neural network

The U-net network in our study follows the original U-net design and consists of nine stages, with four stages in an encoding branch, a middle stage and four stages in a decoding branch.¹ We stick to 3×3 filter size and have 1 convolutional layer per stage. Our loss function is the Huber loss function, which is defined as

$$L(x) = \sum_i \begin{cases} 0.5 (y_i - x_i)^2 & \text{if } |y_i - x_i| \leq \delta \\ 0.5\delta^2 + \delta(|y_i - x_i| - \delta) & \text{if } |y_i - x_i| > \delta \end{cases}, \quad (1)$$

where y_i is the network prediction for input x_i , (x_i has been reshaped to a 1D vector in this formulation). We do not use a mean-squared error loss function since that suppresses outliers by design and, hence, tends to produce smoothed reconstruction results. We use the Adam optimization algorithm (Kingma and Ba, 2015). Some minor modifications w.r.t. the original U-net is that we have included Batch Normalization and replaced the ReLU activation function with the LeakyReLU. Also, we do not use bi-linear interpolation in the upsampling, but a transposed convolution. The neural network is graphically shown in Fig. 6.

5. Results

5.1. Waka-3D survey

The trained neural network was applied to a 2D section of the pre-stacked time-migrated stack of the Waka 3D seismic data. The network was not retrained on this dataset. There are roughly 5800 traces in a crossline section and the section is 6 s long, sampled at 4 ms. The spacing along the section is 12.5 m. The input data, consisting of 5800 \times 1500 samples is processed by the neural network in 50% overlapping windows of size 256 \times 256, the size at which the network was trained. Fig. 7 shows a close-up of a crossline section with 1200 traces for times 0.5–2.5 s. The parameter for the pre-stack migration of this open source dataset are not known, but it can be clearly seen in the top figure of Fig. 7 that migration smiles are present. The bottom figure shows the predicted

¹ With 'stage' we refer to all the operations in between two pooling and/or upsampling operations.

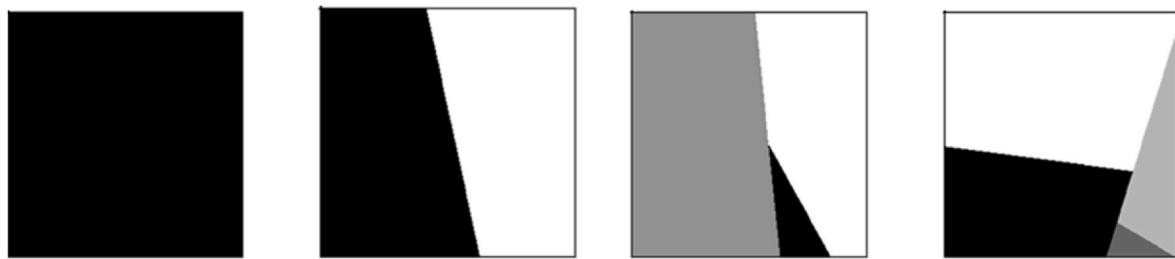


Fig. 1. Example of the creation of segment within a gather using simple geometric shapes. Each of these segments is ‘filled’ with a different reflectivity sequence to which migration smiles are added.

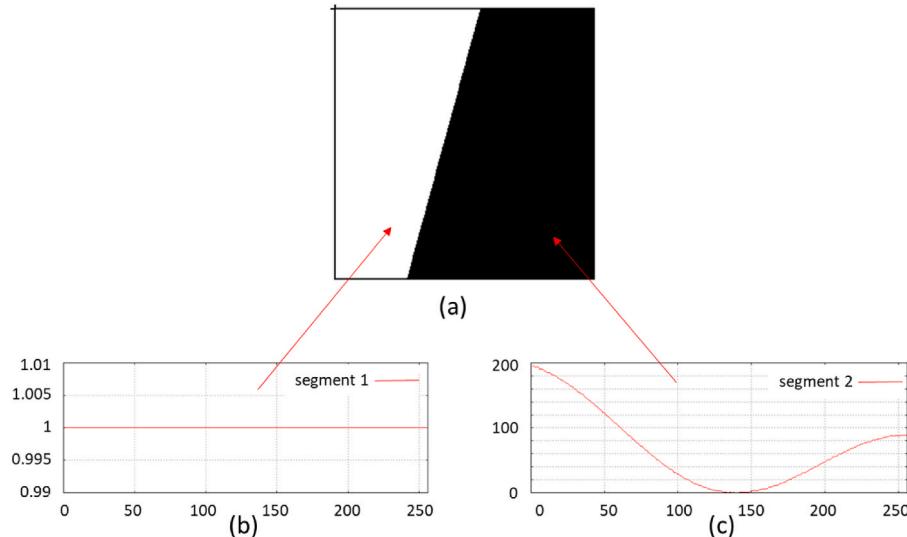


Fig. 2. (a) Example of a gather with two segments. (b,c) show a linear trend with sinusoidal variation and random time shifts. This defines the ‘background’ reflectivity structure on which the migration smile events are later placed.

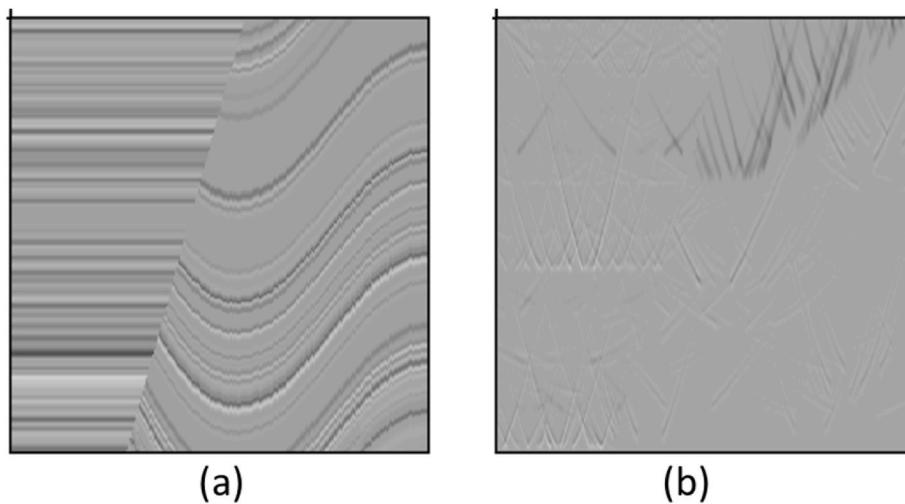


Fig. 3(a). Continuation of the example in Fig. 2. A number of events has been defined within a reflection-only gather. (b) Along the events in (a) migration smile hyperbolae are placed. These hyperbolae have the following random properties: wavelet, amplitude, opening angle, and number per event.

migration smiles. The middle figure the cleaned output, obtained by subtracting the migration smiles from the input data. All figures are shown with the same amplitude scale. Note that the predicted migration smiles do not show a significant imprint of the underlying geology.

5.2. Penobscot 3D survey

The second example shows the neural network applied a common offset section at 475m of the pre-stack migrated 3D Penobscot survey. There are roughly 480 traces in these 2D common offset section and the sections are 6s long, sampled at 4 ms. The spacing along the section is 25

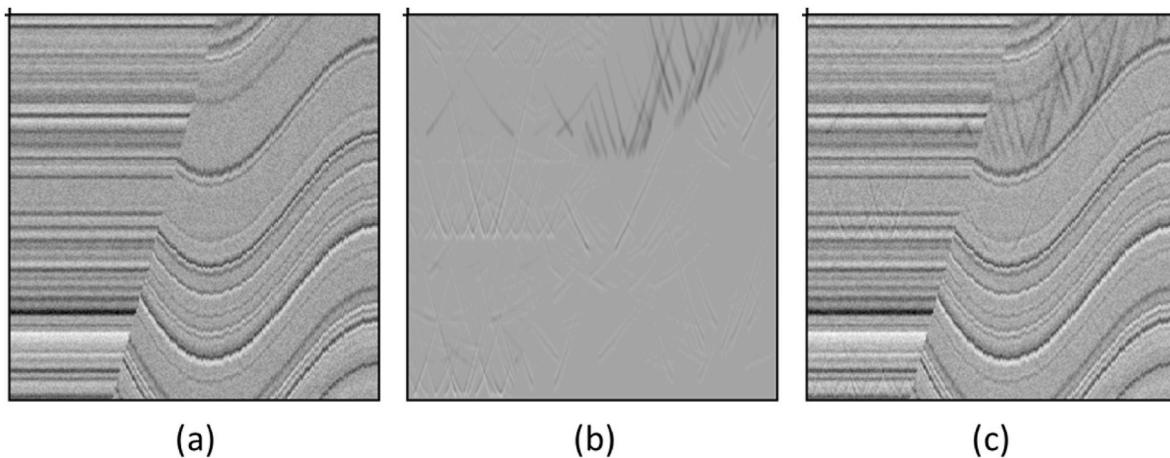


Fig. 4. Continuation of Fig. 3. (a) reflection-only gather, (b) migration-smile-only gather, (c) combination of (a) and (b). Note that (c) is input into the neural network, whereas (b) is the desired output.

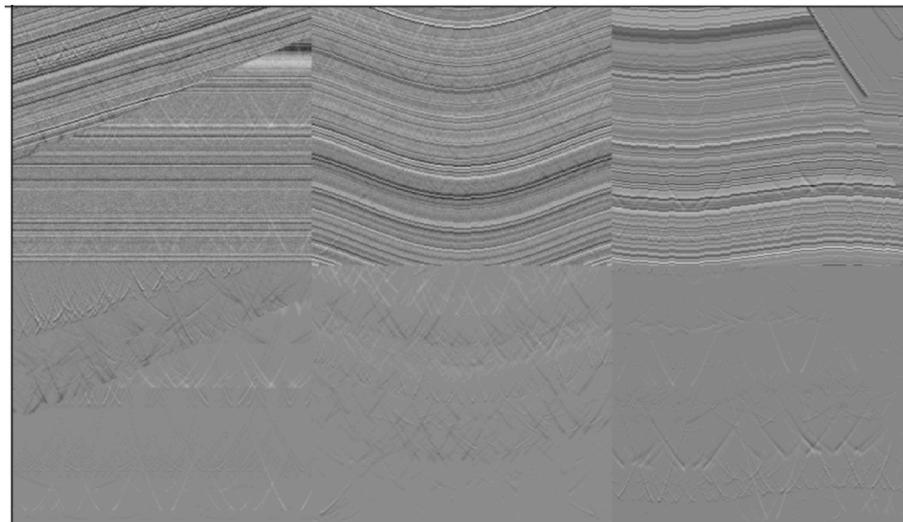


Fig. 5. Examples of gathers in the training dataset. The top row shows three gathers of training data input and the bottom row shows the corresponding migration-smile-only data. Note that the apex of the migration smile has been muted. This will ensure that the neural network will learn not to remove the apex of the diffractions and not to distort the reflection at the apex of the migration smiles.

m. The data were migrated with a 65° maximum angle and aperture of 4100m. Again, we process the data in 50% overlapping windows of size 256×256 . Fig. 8 shows a close-up for times 0–2 s. It can be clearly seen in the top figure of Fig. 7 that migration smiles are present. The top figure shows the input data. The middle figure the cleaned output, obtained by subtracting the migration smiles from the input data. All figures are shown with the same amplitude scale. Note that the predicted migration smiles do not show a significant imprint of the underlying geology.

6. Discussion and conclusion

Conventional techniques for separation of migration smiles can be computationally expensive and difficult to parameterize. A workflow based on deep learning can be a simple and fast alternative, but a common challenge with deep neural networks is generalization to data different from the training dataset. A conventional approach to create training data is to use physics-based modelling, followed by an actual migration, in order to create physically realistic migration smiles. We feel that this is not helpful for generalizing beyond the training data, because this approach provides, by design, a dataset that is limited in

variation. For example, wavefield modelling is performed with a single, fixed wavelet which explicitly limits the variation in the data. To mitigate this lack of diversity one would have to repeat the modelling with many different wavelets. However, this conventional modelling workflow is computationally demanding due to the finite difference based wavefield modelling and seismic migration algorithms used. In practice this means that it becomes prohibitively expensive to create a diverse training dataset, which runs counter the argument that deep learning can provide a simple and fast alternative.

Our position is that it is far more important to create a proper training dataset and, hence, the main focus and novelty of this work is the workflow of generating the synthetic training dataset. The training data was created not by wave equation modelling or by running actual migrations, but simply by convolving a reflectivity model with a wavelet. A large variety of reflectivity of geologically inspired reflectivity patterns were generated, to which upward convex hyperbolic shapes (migration smile) were added. Both patterns were randomly perturbed in many ways. Since migration smiles originate from reflectors and have weaker amplitudes than those reflectors, we used that relationship in the data creation, but each migration smile is independently created from the next. By providing both narrow and wider

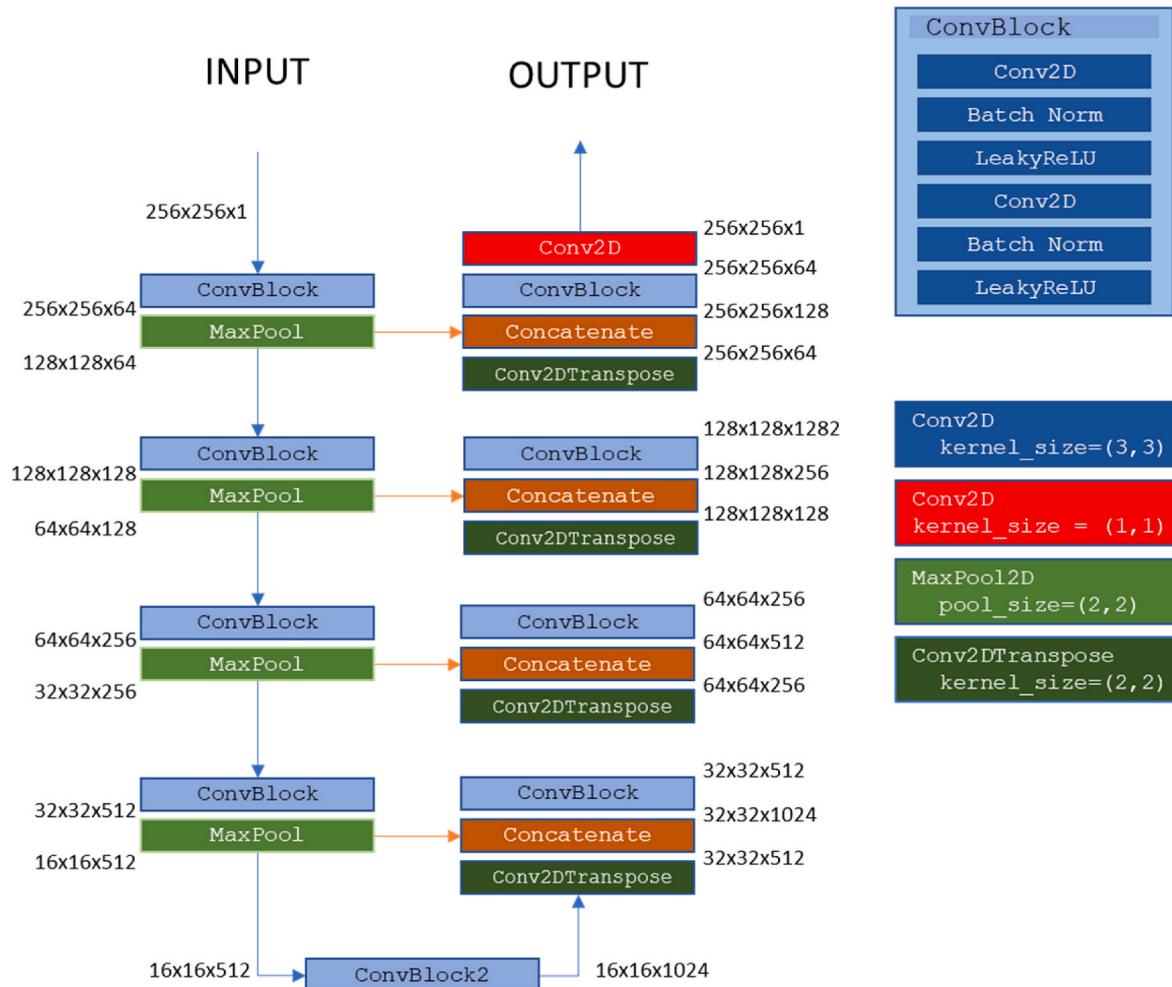


Fig. 6. The U-net network used for migration smile attenuation.

hyperbolae generated with various wavelets, amplitudes and frequencies, we have trained the network to deal with migration smiles at near and far offsets. By mixing various narrow and wide hyperbolae within the same training example we have in fact deviated from what is physically realistic, but this appears to have a positive effect on the generalization of the network. This approach is also taken in other machine vision problems, where it is observed that non-physical distortions applied to the input data, such mixing of different images, geometric transformations and random erasing of samples, actually help make the trained neural network better in predicting in data that was not used during training.

Another difficulty in migration smile attenuation is the separation of migration smiles without removing proper reflection energy due to the relatively low amplitude of smiles compared to reflections as well as the overlap in the kinematics. In order to deal with the overlap in the kinematics of reflection and migration smile events, we have applied a simple, but effective trick of muting the apex of the synthetic migration smile hyperbolae. Where we have therefore not solved this problem, but simply avoided it, accepting residual diffraction energy on the reflection. In most cases where the objective is structural interpretation of the subsurface this is an acceptable compromise. The effect of this apex muting approach is that the network learns to recognize and predict only the non-apex part of the hyperbola, which has positive effects on both synthetic and field data in preserving reflections and avoid leakage of reflection energy into the migration-smile-only data and stacks.

The large variety and diversity in examples enabled the trained neural network to show encouraging results on synthetic and field

datasets that were not used in training. We did not have to retrain the network to demonstrate good results on the field dataset.

7. Conclusion

In this study we have demonstrated a successful application of migration smile separation using a conventional U-net architecture trained with synthetic data. A compromise was made to not predict and subtract the migration smile apex in order to avoid attenuating the primary reflection at the intersection between migration smile and reflection. The key novelty of our approach that we exclusively use data generated with a convolutional model, perturbed in many random ways not achievable with physics-based modelling. The result is a network that is more robust to variations in the input data, which is supported by the fact that we did not need to retrain our model prior to application to unseen field data. This approach carries over to similar problems in which convolutional neural networks are applied to signal or image processing operations. In fact, since deep neural networks learn patterns and relationships between features in the data it is in the majority of the cases not required to use physics-based modelling to create training data, which eliminates a huge computational effort. Future efforts along these lines will be to extend these efforts to 3D seismic data and a wider variety of seismic data processing problems.

Data and materials availability

The Waka 3D seismic data can be found at: <https://wiki.seg.org/wiki>

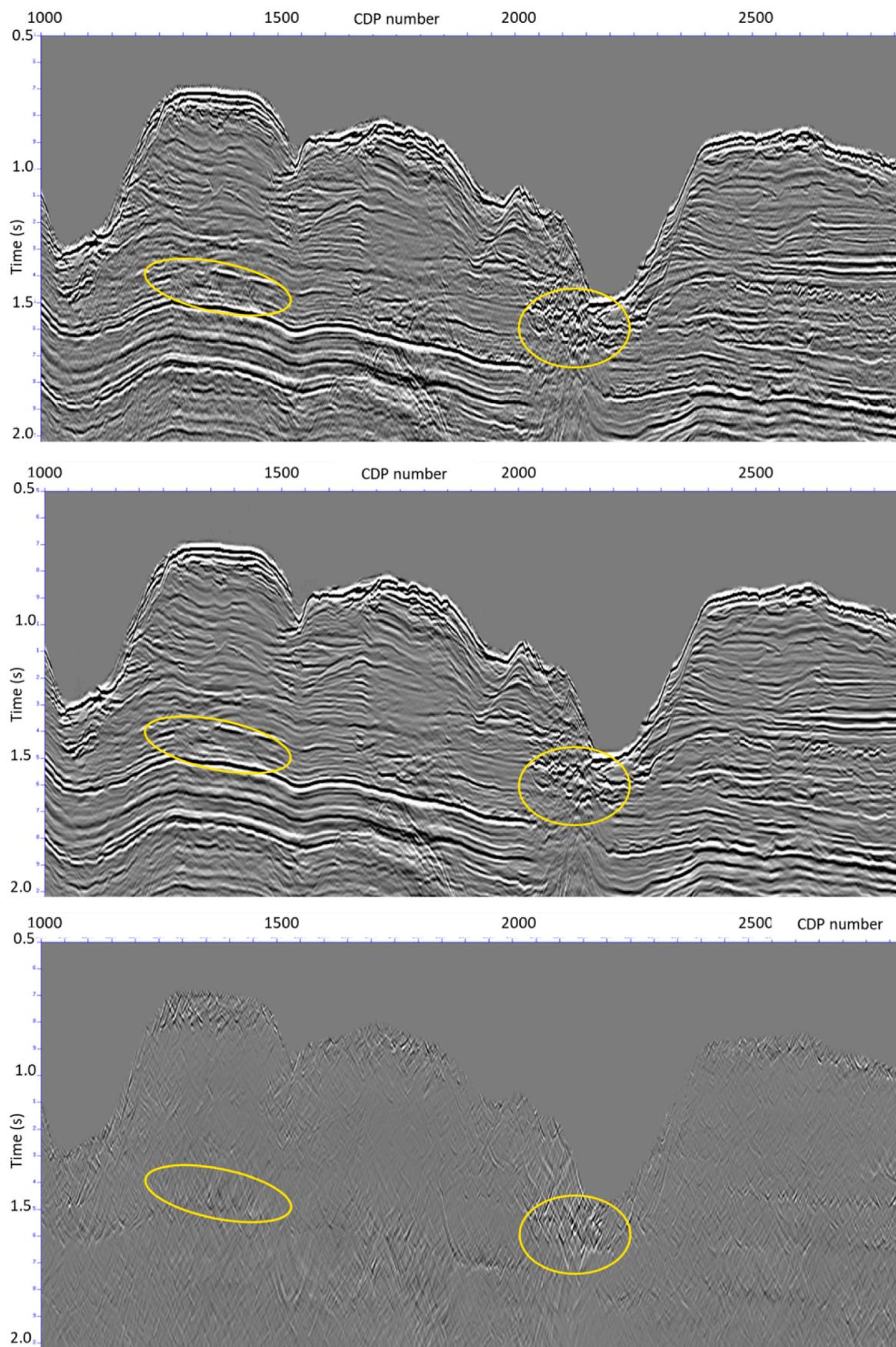


Fig. 7. Result on the pre-stack migrated Waka 3D survey. Top row: input. Middle row: predictions subtracted from input. Bottom row: predicted migration smiles. The yellow ellipses are inserted to help identify a number of migration smiles on the stacks.

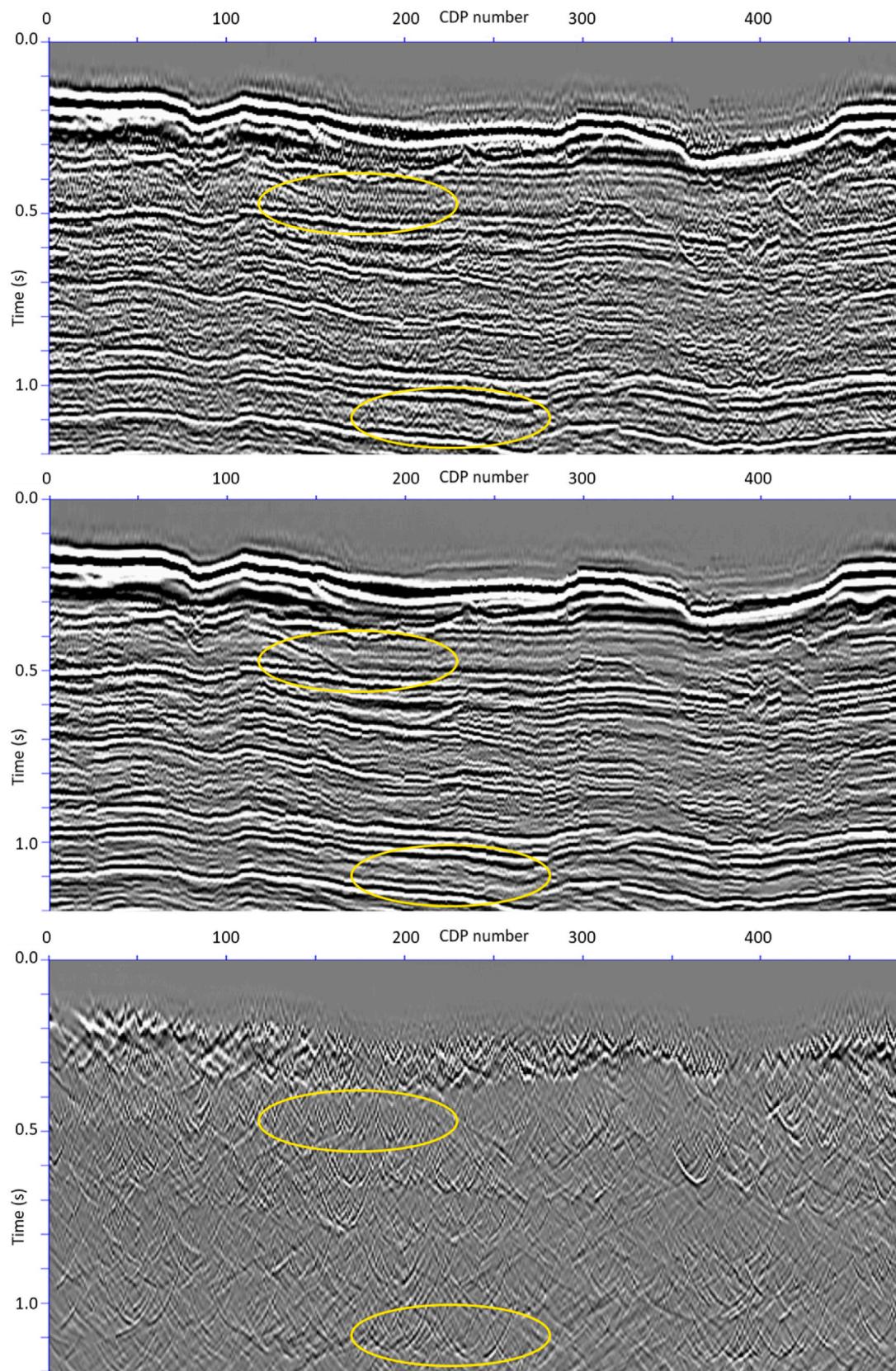


Fig. 8. Common offset section at 475m of the 3D prestack migrated Penobscot survey. Top row: input. Middle row: predictions subtracted from input. Bottom row: predicted migration smiles. Although there are numerous migration smiles in this stack, the yellow ellipses highlight some that are more easily identified in a weak reflectivity zone above a strong seismic reflection.

/Waka-3D/

The Penobscot 3D seismic data can be found at: https://wiki.seg.org/w/index.php?title=Penobscot_3D&oldid=1039972214

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We want to acknowledge the New Zealand Petroleum and Minerals for providing the Waka-3D seismic data, and the Nova Scotia Department of Energy for providing the Penobscot 3D seismic data.

References

- Abma, R., Sun, J., Bernitsas, N., 1999. Antialiasing methods in Kirchhoff migration. *Geophysics* 64, 1783–1792.
- Abma, R., Kelley, C., Kaldy, J., 2007. Sources and treatments of migration-introduced artifacts and noise. *SEG Tech. Progr. Expand. Abstr.* 2349–2353. <https://doi.org/10.1190/1.2792955>.
- Bugge, A.J., Lie, J.E., Evensen, A.K., Nilsen, E.H., Danielsen, F., Gratacos, B., 2021. One-click processing with synthetically pretrained neural networks. *SEG Tech. Progr. Expand. Abstr.* 1851–1855. <https://doi.org/10.1190/segam2021-3583313.1>.
- Chen, Y., Huang, Y., Huang, L., 2020. Suppressing migration image artifacts using a support vector machine method. *Geophysics* 85, S255–S268. <https://doi.org/10.1190/geo2019-0157.1>.
- De Jonge, Th., Vinje, V., Zhao, P., Poole, G., Iversen, E., 2022. Source and receiver deghosting by demigration-based supervised learning. *Geophys. Prospect.* <https://doi.org/10.1111/1365-2478.13253>.
- Duijndam, A.J.W., Schoneville, M.A., Hindriks, C.O., 1999. Reconstruction of band-limited signals, irregularly sampled along one spatial direction. *Geophysics* 64, 524–538. <https://doi.org/10.1190/1.1444559>.
- Duijndam, A.J.W., Volker, A.W.F., Zwartjes, P.M., 2000. Reconstruction as efficient alternative for least squares migration. *SEG Tech. Progr. Expand. Abstr.* 1012–1015. <https://doi.org/10.1190/1.1815554>.
- Dumoulin, V., Visin, F., 2018. A guide to convolution arithmetic for deep learning. <https://arxiv.org/abs/1603.07285>.
- Fang, W., Fu, L., Zhang, M., Li, Z., 2021. Seismic data interpolation based on U-net with texture loss. *Geophysics* 86, V41–V54. <https://doi.org/10.1190/geo2019-0615.1>.
- Fomel, S., Landa, E., Taner, M.T., 2007. Poststack velocity analysis by separation and imaging of seismic diffractions. *Geophysics* 72, U89–U94. <https://doi.org/10.1190/1.2781533>.
- Gray, S., 1992. Frequency-selective design of the Kirchhoff migration operator. *Geophys. Prospect.* 40 (5), 565–571. <https://doi.org/10.1111/j.1365-2478.1992.tb00541.x>.
- Hale, D., 2011. Structure-oriented bilateral filtering of seismic images. *SEG Tech. Progr. Expand. Abstr.* 3596–3600. <https://doi.org/10.1190/1.3627947>.
- Herrmann, P., Mojesky, T., Magesan, M., Hugonnet, P., 2000. De-aliased, high-resolution Radon transforms. *SEG Tech. Progr. Expand. Abstr.* 1953–1956. <https://doi.org/10.1190/1.1815818>.
- Hugonnet, P., Canadas, G., 1997. Regridding of irregular data using 3D Radon decompositions. *SEG Tech. Progr. Expand. Abstr.* 1111–1114. <https://doi.org/10.1190/1.1885585>.
- Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. In: International Conference on Learning Representations (ICLR), pp. 1–15. <https://arxiv.org/abs/1412.6980>.
- Klochikhina, E., Frolov, S., Chemingui, N., 2020. Deep learning for migration artifacts attenuation. EAGE 2020 Annual Conference & Exhibition Online. <https://doi.org/10.3997/2214-4609.202011932>.
- Lou, M., Simpson, H., 2019. An efficient dip filtering approach to improve VSP imaging quality. Fifth EAGE Workshop on Borehole Geophysics 1–3. <https://doi.org/10.3997/2214-4609.2019X604016>.
- Markovic, M., Malehmir, R., Malehmir, A., 2022. Diffraction pattern recognition using deep semantic segmentation. *Near Surf. Geophys.* 20 (5), 507–518. <https://doi.org/10.1002/ng.12227>.
- Masjukov, A., Shlyonkin, V., 2015. Eliminating 3D pre-stack migration artefacts by 6D filtering. *Geophys. Prospect.* 63 (3), 626–636. <https://doi.org/10.1111/1365-2478.12235>.
- Nemeth, T., Wu, C., Schuster, G.T., 1999. Least-squares migration of incomplete reflection data. *Geophysics* 64, 208–221. <https://doi.org/10.1190/1.1444517>.
- Ongkiehong, L., Huizer, W., 1987. Dynamic range of the seismic system. *First Break* 5, 435–439.
- Perez, L., Wang, J., 2017. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning, 04621. <https://doi.org/10.48550/arXiv.1712.04621> arXiv: 1712.04621.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: medical image computing and computer-assisted intervention – MICCAI 2015. MICCAI 2015. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (Eds.), Lecture Notes in Computer Science, vol. 9351. Springer, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
- Shi, Y., Wu, X., Fomel, S., 2019. SaltSeg: automatic 3D salt segmentation using a deep convolutional neural network. *Interpretation* 7 (3), SE113–SE122. <https://doi.org/10.1190/INT-2018-0235.1>.
- Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *Journal of Big Data* 6. <https://doi.org/10.1186/s40537-019-0197-0> article 60.
- Stolk, C., Symes, W., 2004. Kinematic artifacts in prestack depth migration. *Geophysics* 69, 562–575. <https://doi.org/10.1190/1.1707076>.
- Sun, J., Slang, S., Elboth, T., Greiner, T.L., McDonald, S., Gelius, L.-J., 2019. Attenuation of marine seismic interference noise employing a customized U-Net. *Geophys. Prospect.* 68, 845–871. <https://doi.org/10.1111/1365-2478.12893>.
- Trad, D., 2008. Five dimensional seismic data interpolation. *SEG Tech. Progr. Expand. Abstr.* 978–982. <https://doi.org/10.1190/1.3063801>.
- Trad, D.O., Ulrych, T.J., Sacchi, M.D., 2002. Accurate interpolation with high-resolution time-variant Radon transforms. *Geophysics* 67, 644–656. <https://doi.org/10.1190/1.1468626>.
- Witte, P.A., Louboutin, M., Luporini, F., Gorman, G.J., Herrmann, F.J., 2019. Compressive least-squares migration with on-the-fly Fourier transforms. *Geophysics* 84, R655–R672. <https://doi.org/10.1190/geo2018-0490.1>.
- Xu, S., Zhang, Y., Lambaré, G., 2010. Antileakage Fourier transform for seismic data regularization in higher dimensions. *Geophysics* 75, WB113–WB120. <https://doi.org/10.1190/1.3507248>.
- Yang, F., Ma, J., 2019. Deep-learning inversion: a next-generation seismic velocity model building method. *Geophysics* 84, R583–R599. <https://doi.org/10.1190/geo2018-0249.1>.
- Yu, S., Ma, J., 2021. Deep learning for geophysics: current and future trends. *Rev. Geophys.* 59 (3). <https://doi.org/10.1029/2021RG000742>.
- Yuan, P., Wang, S., Hu, W., Wu, X., Chen, J., Nguyen, H.V., 2020. A robust first-arrival picking workflow using convolutional and recurrent neural networks. *Geophysics* 85 (5), U109–U119. <https://doi.org/10.1190/geo2019-0437.1>.
- Zhu, Z., Cao, D., 2019. Low frequency artifact attenuation of reverse-time-migration through anisotropic tensor. *SEG Tech. Progr. Expand. Abstr.* 4361–4364. <https://doi.org/10.1190/segam2019-3203816.1>.
- Zhu, J., Lines, L., Gray, S., 1998. Smiles and frowns in migration/velocity analysis. *Geophysics* 63, 1200–1209. <https://doi.org/10.1190/1.1444420>.
- Zwartjes, P., Gisolf, A., 2006. Fourier reconstruction of marine-streamer data in four spatial coordinates. *Geophysics* 71, V171–V186. <https://doi.org/10.1190/1.2348633>.
- Zwartjes, P.M., Sacchi, M.D., 2007. Fourier reconstruction of nonuniformly sampled, aliased seismic data. *Geophysics* 72, V21–V32. <https://doi.org/10.1190/1.2399442>.
- Zwartjes, P., Yoo, J., 2022. Common-offset domain reflection-diffraction separation with deep learning. *Geophys. Prospect.* <https://doi.org/10.1111/1365-2478.13278>.