# Efficient swell noise removal using a global deep neural network model

Alejandro Valeciano[1*], Olga Brusova[1] and Cheng Cheng[1] describes implementing a global deep learning (DL) model for swell noise removal.


## Summary

This paper describes implementing a global deep learning (DL) model for swell noise removal. It uses new ways to generate training data from a worldwide data library that combines noise-free processed field shot gathers with swell noise recorded during acquisition. Using examples from around the world, we demonstrated that the DL model generalizes well, providing good denoising results in data not previously seen during training. Creating a global model allows for the disruption of the traditional processing sequence and improves efficiency.


## Introduction

Coherent noise interferes with processing algorithms' performance, contaminates seismic images, and degrades the quality of the final interpretative products. That is why it is targeted for early removal in the seismic data processing workflow (e.g., swell noise, seismic interference). In marine settings, rough weather conditions create low-frequency coherent noise in the data (swell noise). Swell noise usually appears in the 2-10 Hz band, showing high amplitude vertical stripes on the shot records (Figure 1). It affects several neighbouring traces, masking signals and reducing data quality (Elboth et al., 2009). Traditional swell noise attenuation methods need extensive parameter tuning and quality control (QC) to obtain clean data with minimal distortion on the signal.

Supervised deep learning (DL) methods work exceptionally well in image denoising tasks, with low computation and human resources (Zhang et al., 2017). They require using labelled datasets to train algorithms and accurately predict outcomes. It has been proven that DL networks can approximate complex functions or non-linear mappings. In contrast, traditional methods are limited to mathematical and physics-based strategies that could fail to model the complexity in the data.

DL models are often trained using large sets of labelled data. Ideally, the training set should contain samples from different geographical and survey acquisition configurations to create a DL model applicable without retraining (global). An obstacle to achieving this goal is that clean seismic or swell noise recordings are often unavailable from the field. Hence the quality and availability of the training set would depend on how and where domain experts process the data to create labels. Insufficient training data has prevented the creation of global DL models, delaying the adoption of existing DL denoising methods on seismic data.

## The traditional swell noise attenuation workflow

Traditional swell noise attenuation comprises alternative flavours of multidimensional transforms within overlapping windows (Fourier, tau-p, etc.) and thresholding (Schonewille et al., 2008, Elboth et al., 2009, Masoomzadeh et al., 2017). Generally, the spectral values that exceed some threshold at a given frequency are considered noise. Practitioners must define multiple parameters like threshold values and multidimensional window sizes often tuned in a few representative sail lines (Figure 2). The goal is to ensure the denoise process does not damage the signal. After a few weeks of parameter optimization, the workflow is applied to the whole survey and later undergoes careful QC. Finding the parameters to apply to an entire seismic survey can be tedious and time-consuming, often resulting in complex batch variant parameters and suboptimal noise subtraction.

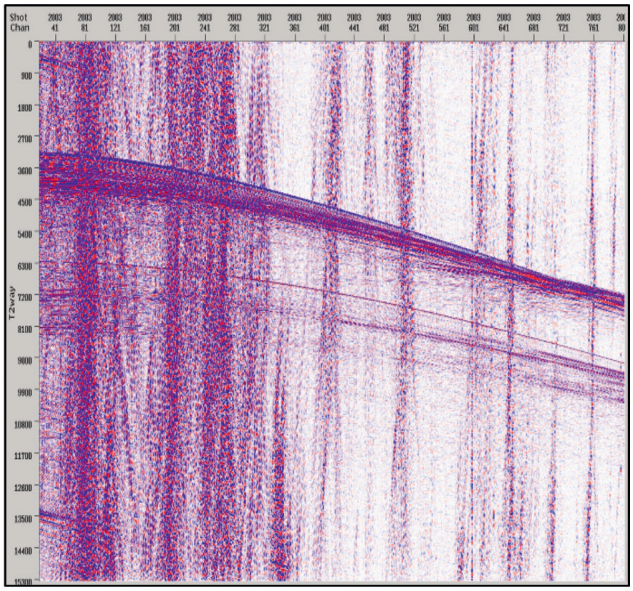


**Figure 1** Shot records contaminated with swell noise.

[1] TGS
[*] Corresponding author, E-mail: Alejandro.Valenciano@tgs.com
DOI: xxx

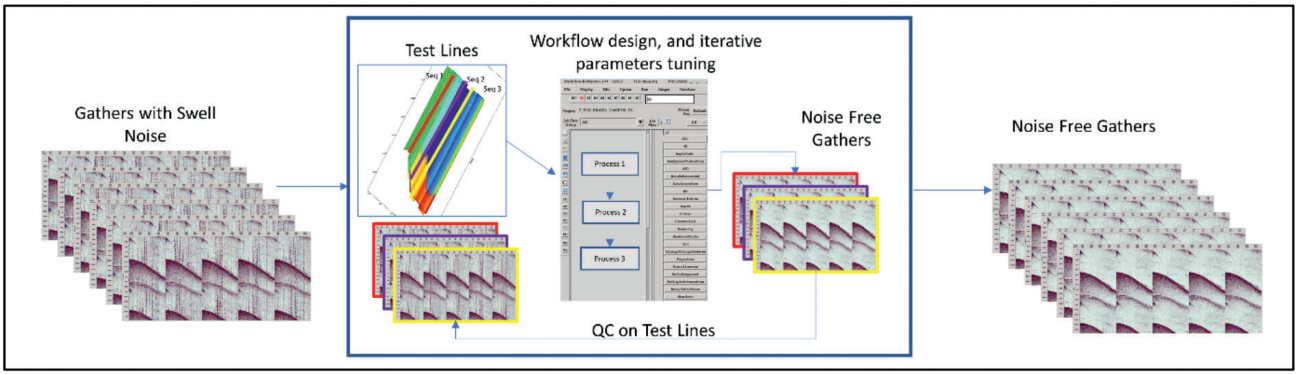


**Figure 2** Traditional swell noise attenuation workflow. The clean data is obtained after noise subtraction.

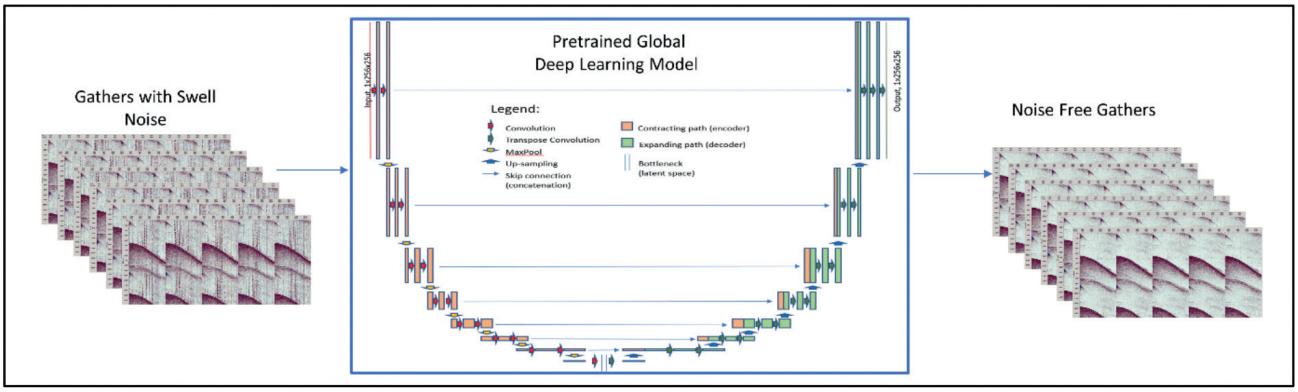


**Figure 3** ML swell noise attenuation workflow. The clean data is obtained after noise subtraction.

## Supervised deep learning for swell noise attenuation

Developing a global DL model impacts swell noise attenuation efficiency (Figure 3). The global model can be applied to the data as soon as it is acquired, saving processing time for a typical project without the need to retrain with input/output pairs from the specific dataset to be processed (local training). The two main components to achieve a global model are model architecture and training dataset.

We use a deep convolutional U-Net architecture to model the swell noise (Figure 4). The network consists of an encoder-decoder framework with symmetric convolutional-deconvolutional layers and skip connections. The convolutional layers capture the input image contents and generated features, while the deconvolutional layers upscale the latent space feature maps and recover desired details of the target. Skip connections make the U-Net use fine-grained information learnt in the encoder part to construct an image in the decoder part. Each encoder block has two convolution layers followed by batch normalization and non-linear leaky ReLU activation function. Each convolutional layer is a 2d filter (kernel) applied to a part of the image defined by a filter size. MaxPooling is used to reduce the size of the features. The input image is repeatedly convolved with filters and subsampled, creating smaller images with more complex activations. The decoder block is almost identical to the encoder block, except it uses deconvolution operation on the feature maps and upscaling instead of the pooling operation. The bottleneck layer is between the encoder and decoder parts of the network. It attempts to model a compressed representation of the data, given the prior assumption of Gaussian distribution. This design for the bottleneck layer works better for an under-determined problem like swell noise removal.

How effective a deep learning method solves a problem depends heavily on the dataset used for model training. DL uses error backpropagation and gradient descent methods to tune model parameters (weights and biases). Loss functions and gradient solvers affect the model's ability to learn specific objectives, while training data affect generalization and exceptional cases.

A pair of noisy input shot gathers, and the desired output label (cleaned gather or noise model) are provided in a typical supervised model training. Although it is easy to obtain input shot gathers, finding the associate noise-free or noise gathers is not

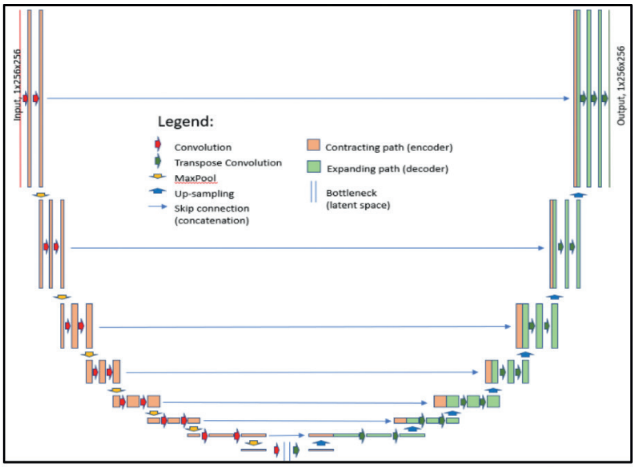


**Figure 4** Deep convolutional neural network U-Net architecture.

trivial. Traditional processing workflows do not provide complete noise removal and would sacrifice leaving some noise in the data to preserve the signal integrity. This data is not ideal for deep learning training. Deep learning models trained on synthetic data can work well when applied to field seismic (Klochikhina et al., 2020). Unfortunately, it is not easy to model swell noise; thus, we cannot use a synthetic training dataset.

We propose a method to create a semi-synthetic training dataset. It uses seismic shot records, after harsh denoising, from multiple projects worldwide to make the clean input data. Any residual swell noise is attenuated in disregard of the signal fidelity. A combination of the clean signal and swell noise records from the field is used as the noisy input to our model (semi-synthetic data). Then we set the network to predict swell noise from the

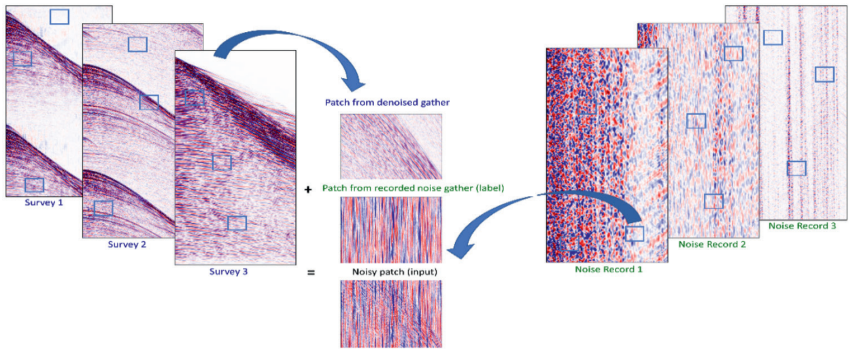


**Figure 5** Schematics of how the training set is created by adding patches of clean data and swell noise field records.

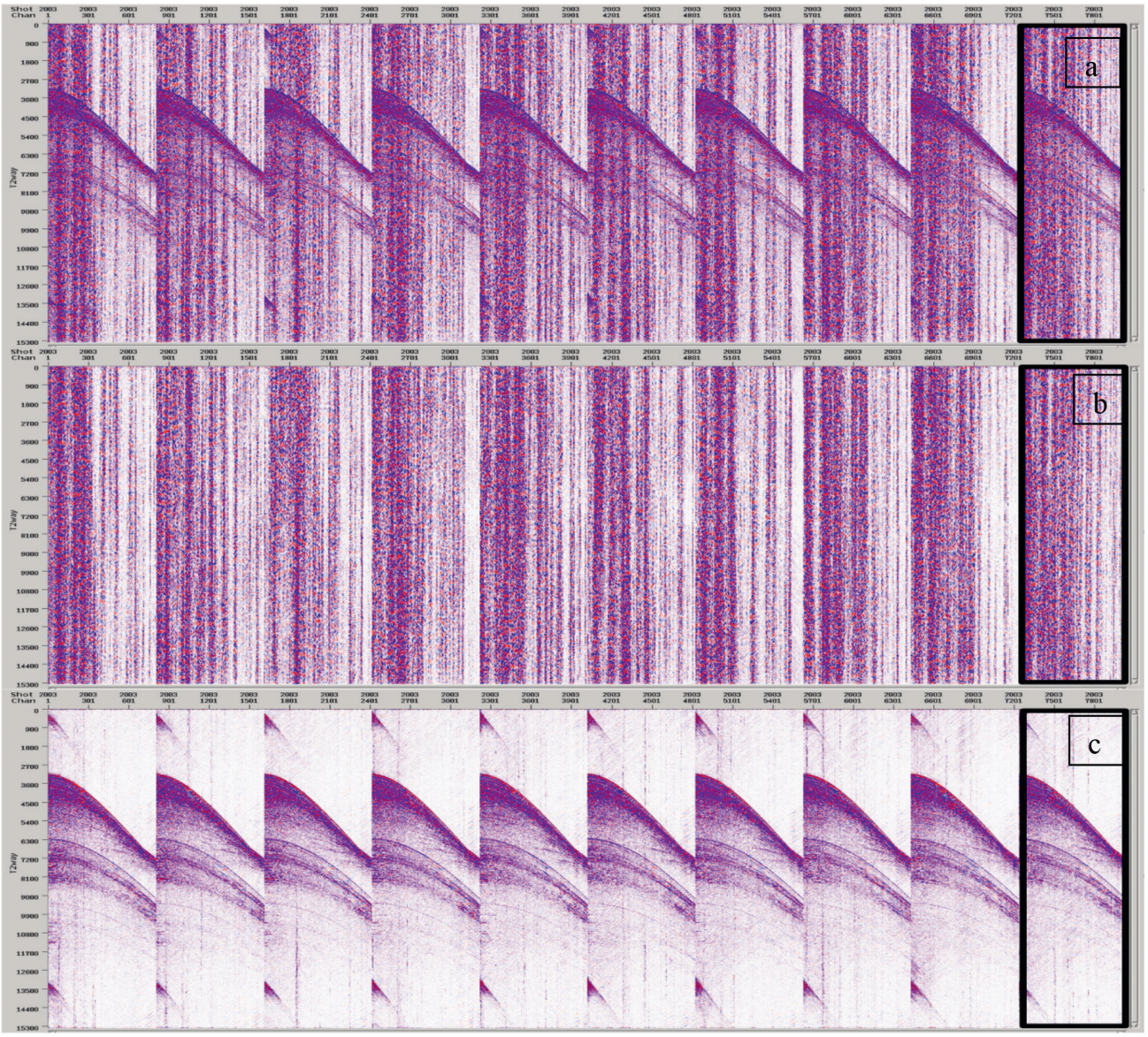


**Figure 6** Example of deep learning swell noise attenuation on a validation sequence: noisy input shot gathers (a), noise model (b), and after DL noise attenuation (c).

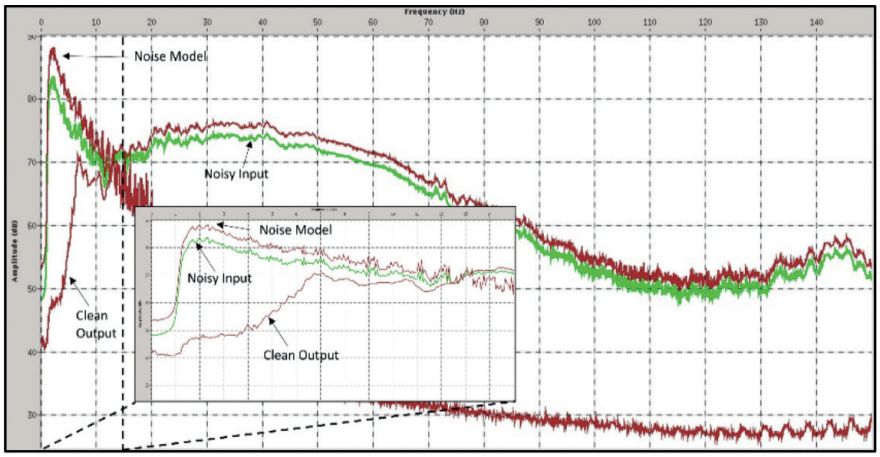


**Figure 7** Frequency spectra extracted from the area in the black rectangle in Figure 6. Noisy input is shown in green. The clean data and noise are shown in brown. The zoomed-in rectangle in the low-frequency highlights the differences.

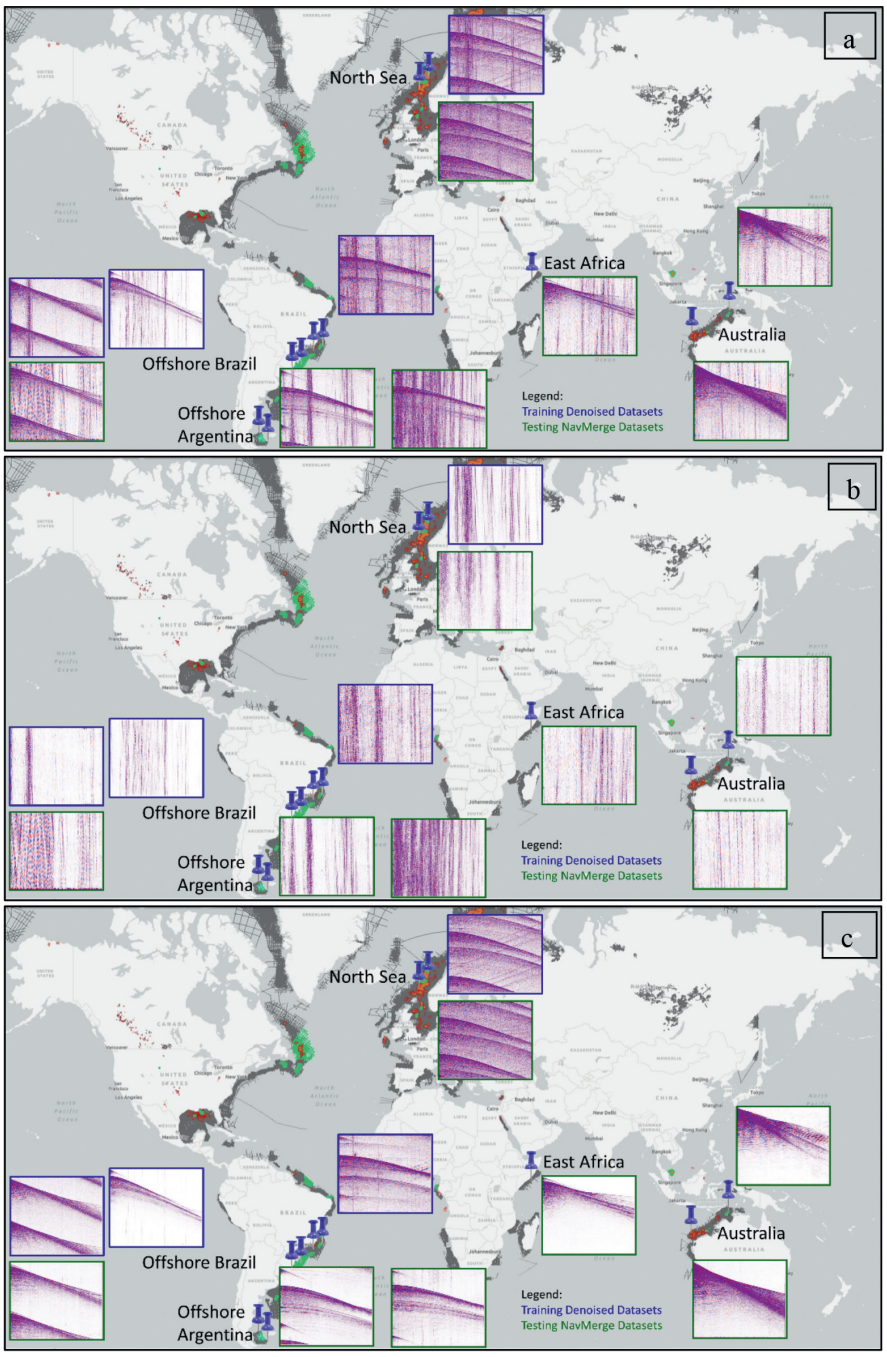


**Figure 8** DL swell noise removal on data not used during the training. Input noisy shot gathers (a), DL predicted swell noise (b), and denoised data (c).

input data. We work with the data patches randomly extracted from the shot gathers in the training dataset. Augmentation techniques such as scaling, polarity changing, and horizontal flipping enhance the dataset statistics. Figure 5 shows how we form the training patches.

## Model training and validation

We generate a large number of training patches using the previous section's approach. A set was used in model training, and the remaining were utilized as the validation set. Separating the data into training/validation/test groups followed machine learning best practices. An AdamW optimizer (Loshchilov and Hutter, 2017), a dropout regularization, and a cosine-annealing learning-rate schedule worked best to avoid overfitting. Hyper-parameters (patch sizes, learning rate) that define how well the neural network performs in a general dataset were also tuned in the multistage training.

The results of the ML denoise are shown in Figure 6 in data not used during training, with a display of 10db gain on the amplitude limits. A group of input shot gathers is shown in Figure 6a. Swell noise is visible in the seismic record as vertical stripes. The swell noise is more substantial in the last two shot gathers. Figure 6c displays clean shot gathers with swell noise effectively removed. The signal does not leak into the predicted noise model shown in Figure 6b.

Figure 7 shows the spectra of the area indicated by the black rectangle in Figure 6. The noisy input and denoised data spectra are almost identical except for the first 7 Hz, where swell noise dominates. The ML denoising approach effectively suppressed the lowest frequencies corresponding to swell noise.

## Generalization on data from a worldwide data library

We test our ML denoise model on different datasets to ensure our global model generalizes well beyond the training data. Figure 8 summarizes the swell noise removal results using DL from multiple datasets distributed worldwide. The input noise gathers are shown in Figure 8a, and the corresponding swell noise model predicted by DL is shown in Figure 8b. Figure 8c shows the cleaned data after direct swell noise model subtraction. As in the validation case, the deep learning noise removal workflow was effective.

We expect the current model to work optimally for most datasets/areas. Although, a sub-optimal performance might be observed in some datasets at the present stage of the training library. This can be addressed by adding more data and labels to the training set. After retraining, the model should perform better on similar data. The DL models will improve over time to overcome limited data exposure during training and become more general.

## Conclusions

Supervised deep neural networks can efficiently remove swell noise from seismic shot records. The potential improvements are fully realized by creating a global deep-learning (DL) model that does not need network retraining for new projects. To that end, we adapted a deep convolutional U-Net architecture and developed a semi-synthetic training data set using a novel strategy. The training data combined recorded swell noise from the field with clean processed data to make optimal model training labels. The DL model proved to generalize well as it produced clean shot gathers on data from different parts of the world that were not part of the training set. The application of the DL model allowed the disruption of the traditional processing sequence and improved efficiency.

## Acknowledgments

We thank TGS for allowing the publication of this paper. We want to express special gratitude to the TGS advisors and management (Scott Michell, Mark Roberts, Cable Warren, CenYen Ong, Graeme Stock, Raheel Malik, and Allan Willis) for their support and thoughtful discussions.

## References

Michel Schonewille, Alan Vigner and Alan Ryder [2008]. 'Swell-noise attenuation using an iterative FX prediction filtering approach,' SEG Technical Program Expanded Abstracts: 2647-2651.

Elboth, T. and Andreassen, O. [2009]. Flow and swell noise in marine seismic: Geophysics,**74**(2), Q17-Q25

Klochikhina, E., S. Crawley, S. Frolov, N. Chemingui, and T. Martin, 2020, Leveraging deep learning for seismic image denoising: *First Break*, **38**, 41-48.

Loshchilov, I. and Hutter, F. [2019]. Decoupled Weight Decay Regularization: arXiv:1711.05101v3

Masoomzadeh, H., Ratnett, N., Travis, T. and Salem, A. [2017]. Multidomain denoise: A robust and efficient method of suppressing incoherent noise:79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, TUP414.

Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. [2017]. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising: *IEEE Transactions on Image Processing*, 26, 3142-315