

Deep-learning based ocean bottom seismic wavefield recovery

Ali Siahkoohi¹, Rajiv Kumar², and Felix J. Herrmann¹

¹*School of Computational Science and Engineering, Georgia Institute of Technology,*

²*Formerly Georgia Institute of Technology; Currently DownUnder GeoSolutions, Perth, Australia*

SUMMARY

Ocean bottom surveys usually suffer from having very sparse receivers. Assuming a desirable source sampling, achievable by existing methods such as (simultaneous-source) randomized marine acquisition, we propose a deep-learning based scheme to bring the receivers to the same spatial grid as sources using a convolutional neural network. By exploiting source-receiver reciprocity, we construct training pairs by artificially subsampling the fully-sampled single-receiver frequency slices using a random training mask and later, we deploy the trained neural network to fill-in the gaps in single-source frequency slices. Our experiments show that a random training mask is essential for successful wavefield recovery, even when receivers are on a periodic grid. No external training data is required and experiments on a 3D synthetic data set demonstrate that we are able to recover receivers for up to 90% missing receivers, missing either randomly or periodically, with a better recovery for random case, at low to midrange frequencies.

INTRODUCTION

Densely sampled wide-azimuth seismic data yields higher resolution images using current imaging techniques. Consequently, seismic data acquisition plays a major role in successful seismic imaging and inversion. In marine seismic surveys, data can be acquired either using a dynamic or static acquisition geometry. In dynamic geometries, sources and receivers are towed behind vessels, as opposed to static geometries, where ocean bottom nodes (OBN) or cables (OBC) are placed on the ocean bottom, and where seismic sources are towed behind one or more vessels. Compared to the other acquisition modalities, OBN geometries have several advantages. For instance, OBN geometries can cover a wider azimuth compared to dynamic geometries. Moreover, static geometry is less influenced by delays in surveying due to bad weather, physical obstacles such as exploration platforms and wells, and fluctuations in receiver depth caused by waves in the sea.

While ocean bottom seismic acquisition has many advantages, because of economic considerations, the nodes live on a very coarse grid. On the other hand, relying on existing work on randomized marine acquisition (Kumar et al., 2015; Cheng and Sacchi, 2015), we can safely assume that sources are finely sampled. Therefore, successful ocean bottom seismic acquisition relies on recovering “missing” receivers (nodes) such that the sources and receivers live on the same (spatial) grid. Recovering missing nodes is a very challenging process since the spatial sampling is large in both the inline and crossline directions. To address this problem, we present a deep-learning based methodology built on source-receiver reciprocity to recover missing receivers. This is accomplished by training a neural network to recover artificially subsampled single-receiver frequency slices, subsampled by a random subsampling mask. In other words, we rely on randomized sampling during training even if the

existing receivers are sampled on a periodic grid. Although this technique is applicable to any type of receiver sampling, e.g., periodic or random, for sampling rates as low as 10%, we observed better recovery when applied to randomly missing case.

Recent successes in the application of machine learning algorithms have been reported in the areas of seismic interpretation (Huang et al., 2017; Di et al., 2018; Veillard et al., 2018), seismic data processing (Siahkoohi et al., 2018b; Sun and Demanet, 2018; Ovcharenko et al., 2018), seismic modeling and imaging (Siahkoohi et al., 2019; Rizzuti et al., 2019; Moseley et al., 2018), and inversion (Mosser et al., 2018; Kothari et al., 2019). Several authors have addressed the seismic wavefield recovery (Siahkoohi et al., 2018a; Mandelli et al., 2018) as well. Mandelli et al. (2018) trained a neural network to jointly denoise and interpolate 2D shot records by operating on extracted patches from shot records. Siahkoohi et al. (2018a) recovered seismic wavefields by reconstructing subsampled frequency slices, by assuming that 5% of shot records are fully-sampled. We extend ideas in Siahkoohi et al. (2018a) by relaxing the assumption of having few fully-sampled shot records by exploiting reciprocity.

Our paper is organized as follows. First, we mention two organizations for monochromatic seismic data. Next, we discuss how we construct training data by using reciprocity. After a brief introduction to Generative Adversarial Networks (GANs, Goodfellow et al., 2014), we state the utilized training objective function and the Convolutional Neural Networks (CNN) architecture. Finally, we demonstrate our results on a synthetic 3D synthetic data set.

THEORY

We describe how ocean bottom acquisition with fine receiver sampling can be achieved, where the receivers originally live on a coarse grid. In a typical ocean bottom acquisition geometry, sources are often subsampled mainly because of economic considerations. As mentioned earlier, we will assume sources are finely sampled on a regular periodic grid but the scheme can be applied to geometries with irregular source sampling with simultaneous sources. Our framework operates on monochromatic frequency slices. Note that, in 3D seismic data acquisition, by working with one frequency only, we reduce the 5D seismic data (two source dimension, two receiver dimension, and one temporal dimension) into 4D. We first introduce two domains in which seismic data can be represented along the four remaining dimensions.

Matricization of monochromatic seismic data

We consider single-source frequency slices as 2D complex-valued arrays of size $N_{R_x} \times N_{R_y}$, where N_{R_x} and N_{R_y} are number of receivers in x and y coordinates, respectively. Similarly, single-receiver frequency slices are a function of two source dimensions and they can be represented as 2D complex-valued

arrays of size $N_{S_x} \times N_{S_y}$, where N_{S_x} and N_{S_y} are number of sources in x and y coordinates, respectively. There are two choices for matricization of monochromatic seismic data over the two source and receiver coordinates (Demanet, 2006; Silva and Herrmann, 2013). One matricization is obtained by putting all the vectorized single-source frequency slices in the columns of a $N_{R_x} N_{R_y} \times N_{S_x} N_{S_y}$ array. This matricization corresponds to placing the (Rec x , Rec y) dimensions in the rows and (Src x , Src y) dimensions in the columns of a matrix. Evidently, the rows in this matrix correspond to vectorized single-receiver frequency slices. Another organization of seismic data in a matrix form can be obtained by placing (Rec y , Src y) dimensions in the rows and (Rec x , Src x) dimensions in the columns. This matrix can be considered as a matrix that includes single-receiver frequency slices as sub-matrix blocks. Suppose now that the receivers are periodically coarsely subsampled by a factor of 10. Figures 1a and 1b illustrate the (Rec y , Src y) \times (Rec x , Src x) organization of fully-sampled and subsampled monochromatic seismic data, respectively. Similarly, Figures 1c and 1d show the fully-sampled and subsampled monochromatic seismic data in (Rec x , Rec y) \times (Src x , Src y) organizations, respectively.

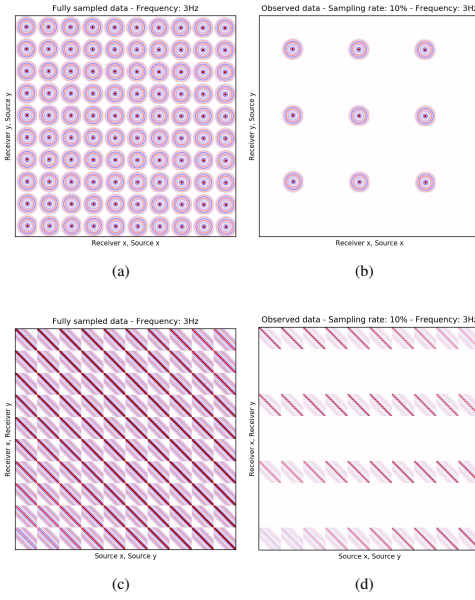


Figure 1: Matricization of monochromatic seismic data in (Rec y , Src y) \times (Rec x , Src x) domain, with a) fine, and b) coarse receiver sampling, and in (Rec x , Rec y) \times (Src x , Src y) domain, with c) fine, and d) coarse receiver sampling.

Learning to reconstruct by exploiting reciprocity

We are after bringing receivers to the same grid as sources—i.e., recovering the empty rows in Figure 1d. In order to reach this goal, we train a CNN \mathcal{G}_θ , parameterized by θ , consisting of all the convolutional kernels and biases. After training, the CNN takes the columns of (Rec x , Rec y) \times (Src x , Src y) organization—i.e., single-source frequency slices, as input and recovers the values in the gaps. Obviously, if we have such a network, we would be able to recover Figure 1c. The choice of CNN for this task is motivated by the universal approxima-

tion theorem, by Hornik et al. (1989), which states that given any continuous function defined on a compact subspace of \mathbb{R}^n , there exists a standard feed-forward single layer neural network with enough neurons that is capable of approximating the function with arbitrary precision. Although this theorem proves the existence of such a neural network, in practice, CNNs are utilized since they are known to generalize well—i.e., maintain the quality of performance when applied to previously unseen data.

Source-receiver reciprocity holds true under certain conditions such as equal source/receiver directivity. As mentioned earlier, we want to train the CNN, \mathcal{G}_θ , to recover subsampled single-source frequency slices. Since receivers are severely subsampled, we do not have access to desired output for training the CNN. To address this problem, we assume source-receiver reciprocity holds—i.e., single-receiver frequency slices are reciprocal to single-source frequency slices. Therefore, to construct training pairs for the CNN, \mathcal{G}_θ , we artificially subsample the fully-sampled single-receiver frequency slices (non-empty rows in Figure 1d) using an arbitrary *training mask*. Note that we are free in choosing the training mask to subsample the fully-sampled single-receiver frequency slices. Based on our experiments, random training mask, with the same subsampling ratio as the receivers, is essential for successful wavefield recovery, even if the existing receivers are sampled on a periodic grid. Ultimately, based on source-receiver reciprocity, we hope that a well-trained CNN will perform equally well when applied to the single-source frequency slices—i.e., all columns of Figure 1d.

Generative Adversarial Networks

Let $S_R = \{\mathbf{r}_i\}_{i=1}^{N_R}$ be the set of fully-sampled single-receiver frequency slices, where N_R is the number of receivers actually sampled along the two receiver coordinates in the field. Denote $S_{\hat{R}} = \{\hat{\mathbf{r}}_i\}_{i=1}^{N_R}$ as the set of single-receiver frequency slices that are subsampled by applying a training mask to slices in S_R . We use a random training mask, with the same subsampling ratio as receivers, since based on our experiments it leads to higher signal-to-noise ratio (SNR), regardless of type of receiver sampling, e.g., periodic or random. GANs introduce a framework in which the CNN, \mathcal{G}_θ , is coupled with an additional cnn, the discriminator, \mathcal{D}_ϕ , that learns to distinguish between fully-sampled frequency slices and the ones that have been recovered by \mathcal{G}_θ . Once a GAN is trained, the range of \mathcal{G}_θ function will be indistinguishable from samples drawn from the probability distribution of fully-sampled frequency slices. To enforce the relationship between each specific pair of subsampled and fully-sampled frequency slices, we include an additional ℓ_1 -norm misfit term weighted by λ (Isola et al., 2017). We use the following objective function for training GANs with input-output pairs:

$$\begin{aligned} \min_{\theta} \mathbb{E}_{\hat{\mathbf{r}} \sim p_{\hat{R}}(\hat{\mathbf{r}}), \mathbf{r} \sim p_R(\mathbf{r})} & \left[(1 - \mathcal{D}_\phi(\mathcal{G}_\theta(\hat{\mathbf{r}})))^2 + \lambda \|\mathcal{G}_\theta(\hat{\mathbf{r}}) - \mathbf{r}\|_1 \right], \\ \min_{\phi} \mathbb{E}_{\hat{\mathbf{r}} \sim p_{\hat{R}}(\hat{\mathbf{r}}), \mathbf{r} \sim p_R(\mathbf{r})} & \left[(\mathcal{D}_\phi(\mathcal{G}_\theta(\hat{\mathbf{r}})))^2 + (1 - \mathcal{D}_\phi(\mathbf{r}))^2 \right]. \end{aligned} \quad (1)$$

where the expectations are approximated with the empirical mean computed with respect to pairs $(\hat{\mathbf{r}}_i, \mathbf{r}_i)$, $i = 1, 2, \dots, N_R$ of subsampled and fully-sampled single-receiver frequency slices drawn from the probability distributions $p_{\hat{\mathbf{r}}}(\hat{\mathbf{r}})$ and $p_{\mathbf{r}}(\mathbf{r})$. As proposed by Johnson et al. (2016), we use a ResNet (He et al., 2016) for the generator \mathcal{G}_θ and we follow Isola et al. (2017) for the discriminator \mathcal{D}_ϕ architecture. The hyper-parameter λ ensures that the generator maps specific pairs $(\hat{\mathbf{r}}_i, \mathbf{r}_i)$ to each other (Isola et al., 2017). Solving the optimization objective 1 is typically based on Stochastic Gradient Descent (SGD) or one of its variants (Bottou et al., 2018; Goodfellow et al., 2016).

EXPERIMENTS

We apply our framework to a synthetic dataset simulated on a portion of BG Compass model. The data is generated on a 172×172 2D periodic grid of sources and a 172×172 2D periodic grid of receivers, both with 25 m spatial sampling in both x and y directions. We will apply our algorithm to 3, 5, 10, and 15 Hz monochromatic seismic data extracted from the mentioned data set. We will evaluate the performance of the algorithm in cases where the we miss 90% of receivers, either randomly or periodic.

To perform our numerical experiment, we remove receivers by a factor of 10 for both random and periodic case. For the random case, Figures 2a and 2e in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ domain and Figures 2b and 2f in $(\text{Rec } x, \text{Rec } y) \times (\text{Src } x, \text{Src } y)$ domain indicate the randomly picked receivers for 10 and 15 Hz, respectively. Similarly, periodically subsampled receivers for 10 and 15 Hz monochromatic seismic data in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ and $(\text{Rec } x, \text{Rec } y) \times (\text{Src } x, \text{Src } y)$ domain can be seen in Figures 2i, 2m, 2j, and 2n. For the sake of clarity, we are only showing a small portion of the data in Figure 2 since the dimensions of the monochromatic seismic data in both matricizations is 29584×29584 . Since only 10% of the receivers are recording, we only have 2958 training pairs of fully-sampled and artificially subsampled single-receiver frequency slices, obtained by using a random training mask. For each frequency, we trained a CNN by minimizing Equation 1 for 100 passes over 2958 training pairs using $\lambda = 1000$, which balances generator’s tasks for fooling the discriminator and mapping specific pairs $(\hat{\mathbf{r}}_i, \mathbf{r}_i)$ to each other.

After training the CNN, for each frequency, we reconstruct the missing receivers by first, extracting single-source frequency slices from all the columns in the $(\text{Rec } x, \text{Rec } y) \times (\text{Src } x, \text{Src } y)$ organization (see Figures 2b, 2f, 2j, and 2n). Next, we feed the extracted single-source frequency-slices to the trained CNN. Based on source-receiver reciprocity, we can be optimistic that the CNN will perform well when applied to the single-source frequency slices. Figures 2c, 2g, 2k, and 2o illustrate the reconstructed single-receiver frequency slices in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ domain for 10 and 15 Hz monochromatic seismic data, for the random and periodic case, respectively. The reconstruction error for 10 and 15 Hz are included in Figures 2d, 2h, 2l, and 2p, respectively.

The obtained results indicate that the CNN trained on pairs of fully-sampled and subsampled single-source frequency slices is able to recover the single-source frequency slices with 90% missing receivers, either missing randomly or periodically, with

higher accuracy in randomly missing case. Our experiments showed that a CNN trained on training pairs obtained by subsampling the single-receiver frequency slices with a random training mask equal to the randomly missing receiver sampling mask, performs better than using a periodic training mask on both randomly and periodically missing receivers cases. While we can not fully justify this observation, but we know from Compressive Sensing that by utilizing a randomized sampling, the underlying signal can be recovered with a higher probability. We make a similar observation in our experiment, namely as the frequency increases, the performance of the method degrades faster for the periodically missing receiver case. In general, the decrease in SNR as frequency increases is understandable since in higher frequencies, the frequency slices are more intricate. Table 1 includes the average recovery SNRs for 3, 5, 10, and 15 Hz monochromatic seismic data for randomly or periodically missing receivers. Note that for the periodic case, we have used the same network as in the random case.

Sampling mask	Frequency	Average recovery SNR
random	3 Hz	32.66 dB
random	5 Hz	29.07 dB
random	10 Hz	23.46 dB
random	15 Hz	17.31 dB
periodic	3 Hz	32.17 dB
periodic	5 Hz	28.32 dB
periodic	10 Hz	20.82 dB
periodic	15 Hz	9.12 dB

Table 1: Average reconstruction SNR for 90% random/periodic missing receivers.

DISCUSSION & CONCLUSIONS

We introduced a framework in which data recorded via an ocean bottom acquisition survey, can be mapped to a finer receiver sampling grid, regardless of type of receiver sampling, e.g., periodic or random. An important distinction of this contribution compared to similar works is that it does not need any external training data. This is accomplished by assuming that source-receiver reciprocity holds and the sources are densely sampled, which is possible using existing methods such as (simultaneous-source) randomized marine acquisition. Based on later assumption, single-receiver frequency slices, corresponding to existing receivers, are fully-sampled. We trained a CNN with training pairs consisting of the fully-sampled and artificially subsampled single-receiver frequency slices, where the subsampling is performed using a random training mask. Our numerical experiment on a 3D synthetic data set, with 90% missing receivers, missing either randomly or periodically, over low to midrange frequencies indicated that a random training mask is able to recover the wavefields with high SNR and in general, our method is more accurate when applied to randomly missing receivers case. In the future, more work is needed to improve the accuracy of the method on higher frequencies.

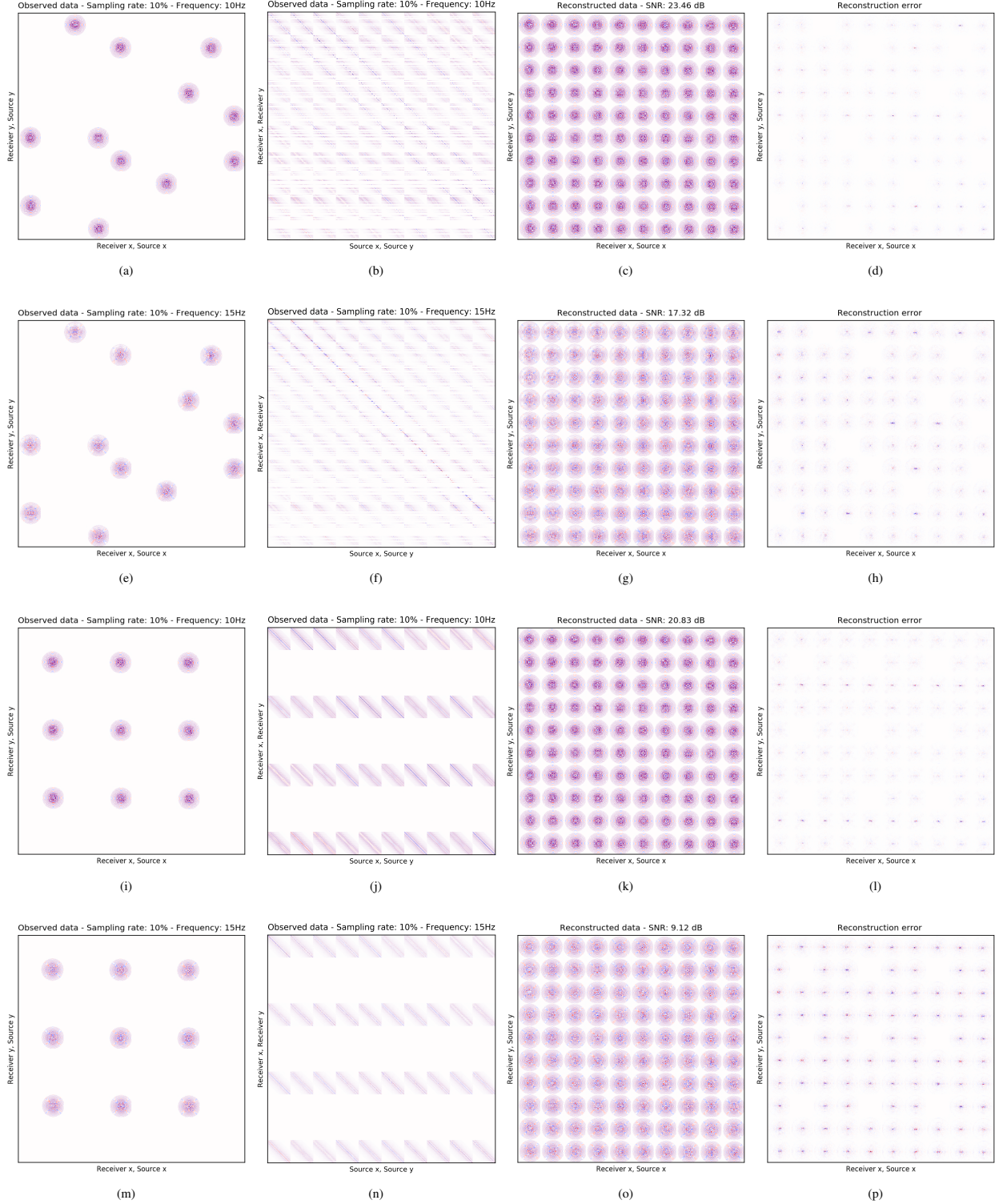


Figure 2: Monochromatic seismic data recovery with 90% missing receivers, either random or periodic. Monochromatic seismic data in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ domain when missing 90% of receivers a) randomly, at 10 Hz, e) randomly, at 15 Hz, i) periodically, at 10 Hz, and m) periodically, at 15 Hz. Monochromatic seismic data in $(\text{Rec } x, \text{Rec } y) \times (\text{Src } x, \text{Src } y)$ domain when missing 90% of receivers b) randomly, at 10 Hz, f) randomly, at 15 Hz, j) periodically, at 10 Hz, and n) periodically, at 15 Hz. Recovered monochromatic seismic data in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ domain when missing 90% of receivers c) randomly, at 10 Hz, g) randomly, at 15 Hz, k) periodically, at 10 Hz, and o) periodically, at 15 Hz. Reconstruction error of recovered monochromatic seismic data in $(\text{Rec } y, \text{Src } y) \times (\text{Rec } x, \text{Src } x)$ domain when missing 90% of receivers d) randomly, at 10 Hz, h) randomly, at 15 Hz, l) periodically, at 10 Hz, and p) periodically, at 15 Hz.

REFERENCES

- Bottou, L., F. E. Curtis, and J. Nocedal, 2018, Optimization methods for large-scale machine learning: *SIAM Review*, **60**, 223–311.
- Cheng, J., and M. D. Sacchi, 2015, Separation and reconstruction of simultaneous source data via iterative rank reduction: *GEOPHYSICS*, **80**, V57–V66.
- Demagnet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
- Di, H., Z. Wang, and G. AlRegib, 2018, Deep convolutional neural networks for seismic salt-body delineation: Presented at the AAPG Annual Convention and Exhibition.
- Goodfellow, I., Y. Bengio, and A. Courville, 2016, *Deep learning*: MIT Press. (<http://www.deeplearningbook.org>).
- Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, 2014, Generative Adversarial Nets: Proceedings of the 27th International Conference on Neural Information Processing Systems, 2672–2680.
- He, K., X. Zhang, S. Ren, and J. Sun, 2016, Deep Residual Learning for Image Recognition: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.
- Hornik, K., M. Stinchcombe, and H. White, 1989, Multilayer feedforward networks are universal approximators: *Neural networks*, **2**, 359–366.
- Huang, L., X. Dong, and T. E. Clee, 2017, A scalable deep learning platform for identifying geologic features from seismic attributes: *The Leading Edge*, **36**, 249–256.
- Isola, P., J.-Y. Zhu, T. Zhou, and A. A. Efros, 2017, Image-to-Image Translation with Conditional Adversarial Networks: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5967–5976.
- Johnson, J., A. Alahi, and L. Fei-Fei, 2016, Perceptual Losses for Real-Time Style Transfer and Super-Resolution: Computer Vision – European Conference on Computer Vision (ECCV) 2016, Springer International Publishing, 694–711.
- Kothari, K., S. Gupta, M. v. de Hoop, and I. Dokmanic, 2019, Random mesh projectors for inverse problems: Presented at the International Conference on Learning Representations.
- Kumar, R., H. Wason, and F. J. Herrmann, 2015, Source separation for simultaneous towed-streamer marine acquisition — a compressed sensing approach: *GEOPHYSICS*, **80**, WD73–WD88.
- Mandelli, S., F. Borra, V. Lipari, P. Bestagini, A. Sarti, and S. Tubaro, 2018, Seismic data interpolation through convolutional autoencoder: SEG Technical Program Expanded Abstracts 2018, 4101–4105.
- Moseley, B., A. Markham, and T. Nissen-Meyer, 2018, Fast approximate simulation of seismic waves with deep learning: arXiv preprint arXiv:1807.06873.
- Mosser, L., O. Dubrule, and M. Blunt, 2018, Stochastic Seismic Waveform Inversion Using Generative Adversarial Networks As A Geological Prior: First EAGE/PESGB Workshop Machine Learning.
- Ovcharenko, O., V. Kazei, D. Peter, X. Zhang, and T. Alkhalifah, 2018, Low-Frequency Data Extrapolation Using a Feed-Forward ANN: 80th EAGE Conference and Exhibition 2018.
- Rizzuti, G., A. Siahkoohi, and F. J. Herrmann, 2019, Learned iterative solvers for the helmholtz equation: 81st EAGE Conference and Exhibition 2019.
- Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2018a, Seismic Data Reconstruction with Generative Adversarial Networks: 80th EAGE Conference and Exhibition 2018.
- Siahkoohi, A., M. Louboutin, and F. J. Herrmann, 2019, The importance of transfer learning in seismic modeling and imaging. (Submitted to *GEOPHYSICS* in February 2019).
- Siahkoohi, A., M. Louboutin, R. Kumar, and F. J. Herrmann, 2018b, Deep-convolutional neural networks in prestack seismic: Two exploratory examples: SEG Technical Program Expanded Abstracts 2018, 2196–2200.
- Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the , Sampling Theory and Applications conference.
- Sun, H., and L. Demagnet, 2018, Low-frequency extrapolation with deep learning: SEG Technical Program Expanded Abstracts 2018, 2011–2015.
- Veillard, A., O. Morère, M. Grout, and J. Gruffeille, 2018, Fast 3D Seismic Interpretation with Unsupervised Deep Learning: Application to a Potash Network in the North Sea: 80th EAGE Conference and Exhibition 2018.