Surface-related multiple elimination with deep learning

Ali Siahkoohi[1], Dirk J. Verschuur[2], and Felix J. Herrmann[1]

[1] School of Computational Science and Engineering, Georgia Institute of Technology
[2] Faculty of Applied Sciences, Delft University of Technology

SLIM
Georgia Institute of Technology

Berkhout, A. J., and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations: GEOPHYSICS, 62, 1586–1595.
Guitton, A., and D. J. Verschuur, 2004, Adaptive subtraction of multiples using the l1-norm: Geophysical Prospecting, 52, 27–38.
Groenestijn, G. J. V., and D. J. Verschuur, 2009, Estimating primaries by sparse inversion and application to near-offset data reconstruction: GEOPHYSICS, 74.
Lin, T. T., and F. J. Herrmann, 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, 78, R133–R150.

Surface-related multiple elimination

Prediction and subtraction problem

- SRME

Inverse problem w/ primary reflections as unknowns

- EPSI

Both are computationally expensive, especially in 3D

Questions to investigate

1. Can we approximate the action of EPSI and SMRE via CNNs?
2. Supervised learning: What should be input/output pairs for training CNNs?
3. Can CNNs handle the intricacies of field data?
4. Can the trained CNN be applied to data from another seismic survey?

Questions to investigate

1. Can we approximate the action of EPSI and SMRE via CNNs?
2. Supervised learning: What should be input/output pairs for training CNNs?
3. Can CNNs handle the intricacies of field data?
4. Can the trained CNN be applied to data from another seismic survey?

Hornik, K., M. Stinchcombe, and H. White, 1989, Multilayer feedforward networks are universal approximators: Neural networks, 2, 359–366.

Approximating EPSI w/ CNNs

EPSI and SRME algorithms can be considered as functions such that,

- they “map” data with multiples, to predicted primaries.

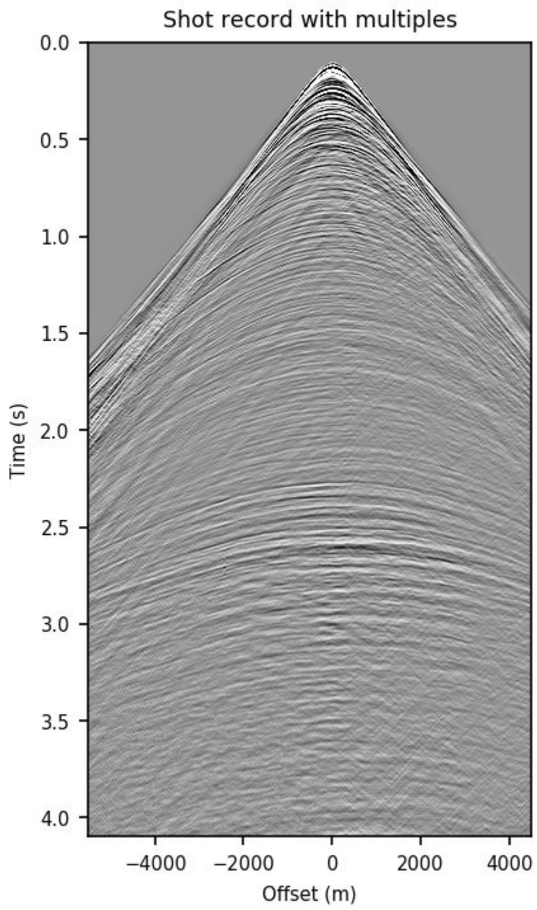
Answer:

- in theory, yes
- universal approximation theory: NNs can approximate any continuous function defined on a compact subspace, with arbitrary precision

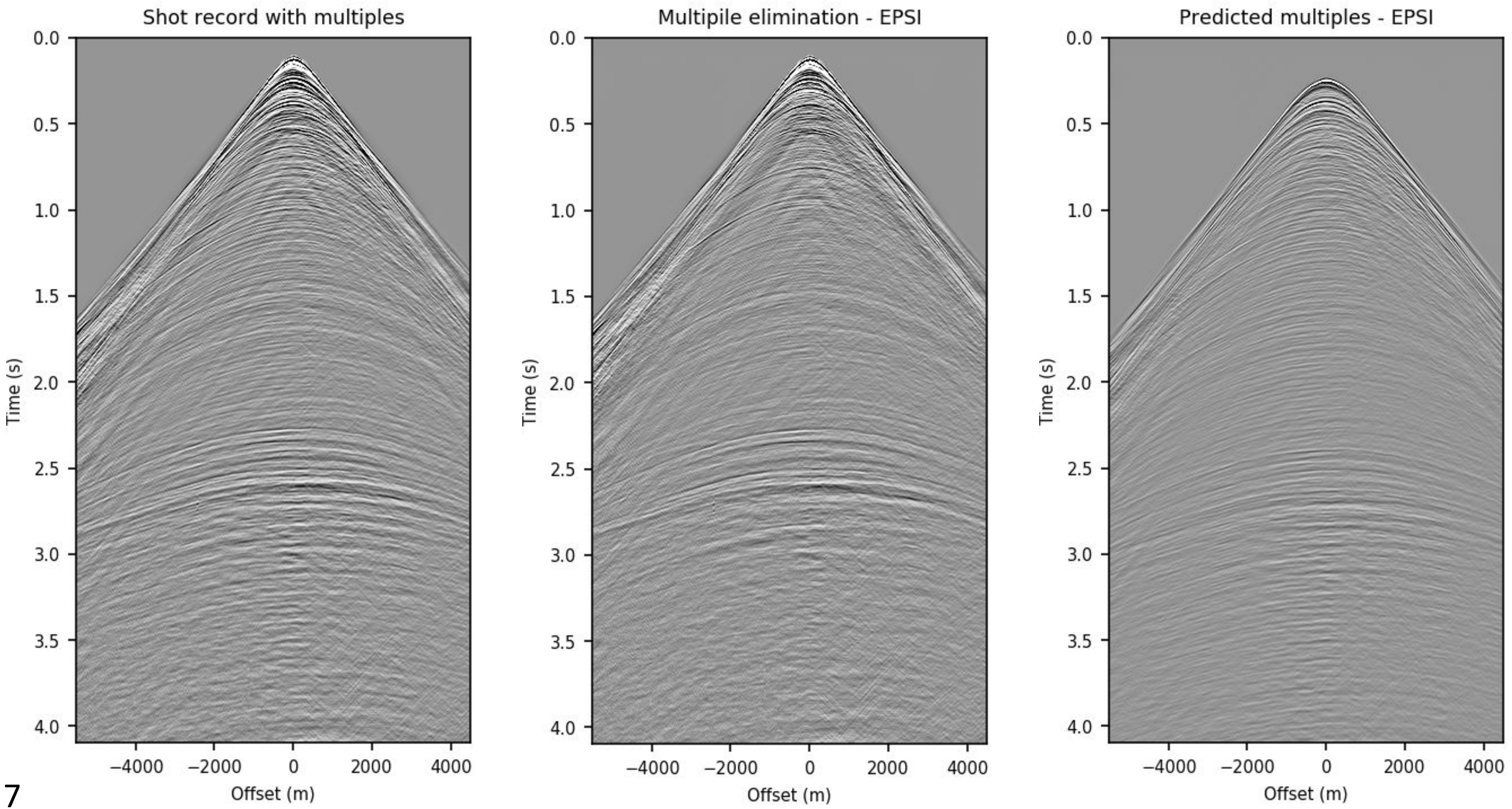
Nelson field data set

After exploiting reciprocity and applying near-offset interpolation:

- 401 shot records
- Each shot, 401 traces
- 1024 time samples per trace
- Time sampling interval: 4 ms
- Source/receivers spacing: 12.5 m



Multiple elimination w/ EPSI



Questions to investigate

1. Can we approximate the action of EPSI and SMRE via CNNs?

2. Supervised learning: What should be input/output pairs for training CNNs?

3. Can CNNs handle the intricacies of field data?

4. Can the trained CNN be applied to data from another seismic survey?

Questions to investigate

2. Supervised learning: What should be input/output pairs for training CNNs?
3. Can CNNs handle the intricacies of field data?

Answer:

- ► We conduct two experiments w/ different input/output pairs to train CNN

Training: Input/output pairs of CNN

Experiment 1:

- Input: shot records w/ multiples
- Output: predicted primaries by EPSI

Experiment 2:

- Input: shot records w/ multiples and a relatively poor prediction of the multiples
- Output: predicted primaries and multiples by EPSI

201 shot locations for training, by picking every other shot location.

Data augmentation by flipping the shot records with respect to the offset axis.

Mao X, Li Q, Xie H, Lau RY, Wang Z, Paul Smolley S. Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision 2017, pages 2794-2802.
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5967–5976, July 2017.

Training framework: GANs

$$\min_{\theta} \mathop{\mathbb{E}}_{\mathbf{x}\sim p_X(\mathbf{x}),\, \mathbf{y}\sim p_Y(\mathbf{y})} \left[\left(1 - \mathcal{D}_{\phi}\left(\mathcal{G}_{\theta}(\mathbf{x})\right)\right)^2 + \lambda \left\| \mathcal{G}_{\theta}(\mathbf{x}) - \mathbf{y} \right\|_1 \right],$$

$$\min_{\phi} \mathop{\mathbb{E}}_{\mathbf{x}\sim p_X(\mathbf{x}),\, \mathbf{y}\sim p_Y(\mathbf{y})} \left[\left(\mathcal{D}_{\phi}\left(\mathcal{G}_{\theta}(\mathbf{x})\right)\right)^2 + \left(1 - \mathcal{D}_{\phi}(\mathbf{y})\right)^2 \right].$$

$\{\mathbf{x},\ \mathbf{y}\}$ Input/output pairs, drawn from the probability distributions $p_X(\mathbf{x})$ and $p_Y(\mathbf{y})$

$\mathcal{G}_{\theta}(\mathbf{x})$ Generator

$\mathcal{D}_{\phi}$ Discriminator

ℓ_1-norm misfit term weighted by λ ensures that each realization of $\mathcal{G}_{\theta}(\mathbf{x})$ maps to a particular $\mathbf{y}$, i.e., $\mathbf{x} \mapsto \mathbf{y}$ rather than solely fooling the discriminator.

Quan, T. M., D. G. C. Hildebrand, and W.-K. Jeong, 2016, FusionNet: A deep fully residual convolutional neural net- work for image segmentation in connectomics: CoRR, abs/1612.05360.
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.

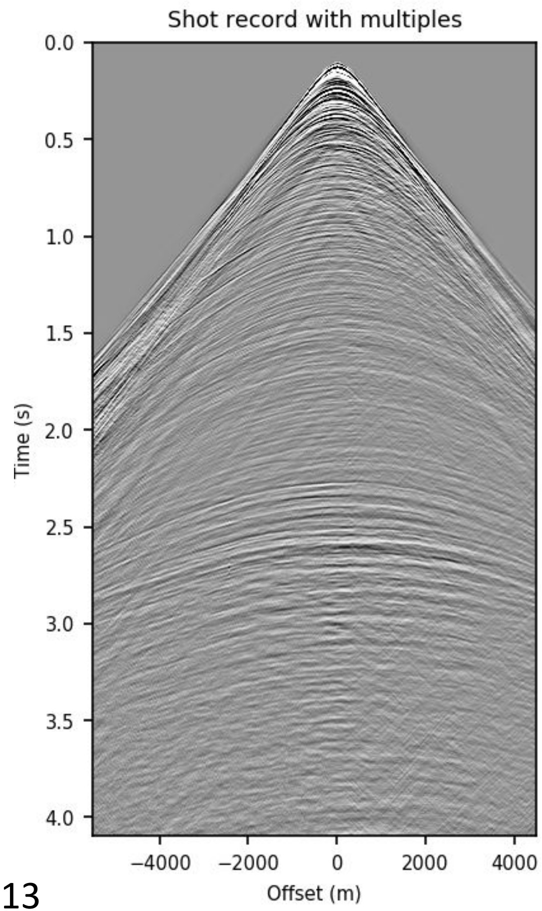
CNN architecture

Generator: modified CNN architecture introduced by Quan et al. (2016)

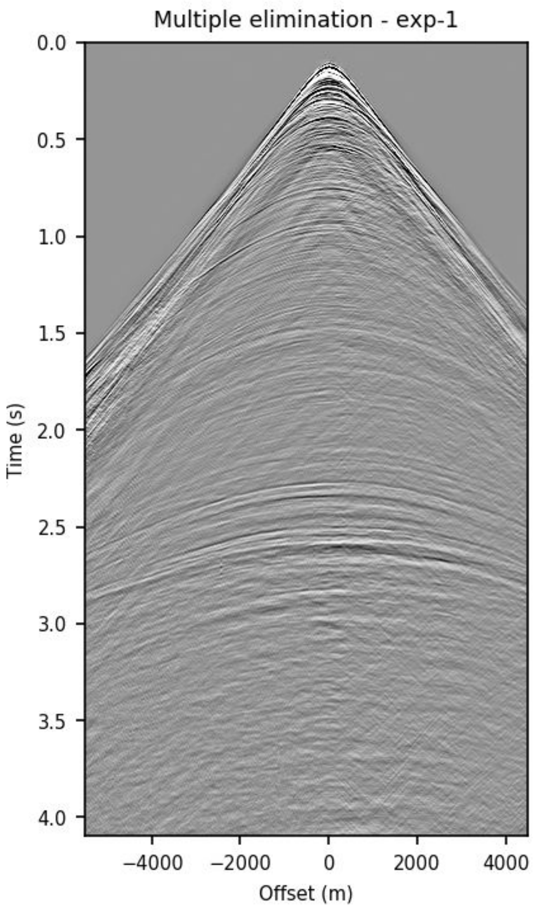
- 16 blocks, including eight encoding and eight decoding blocks
- Encoding blocks: a Residual Block (He et al., 2016), and a convolutional layer with stride two
- Decoding blocks: a Residual Block and a transposed convolutional layer with stride two
- For i = 1,2, ..., 7, the output of i'th block is concatenated with the output of (15-i)'th block

Discriminator: PatchGAN classifier (Isola et al. (2017))

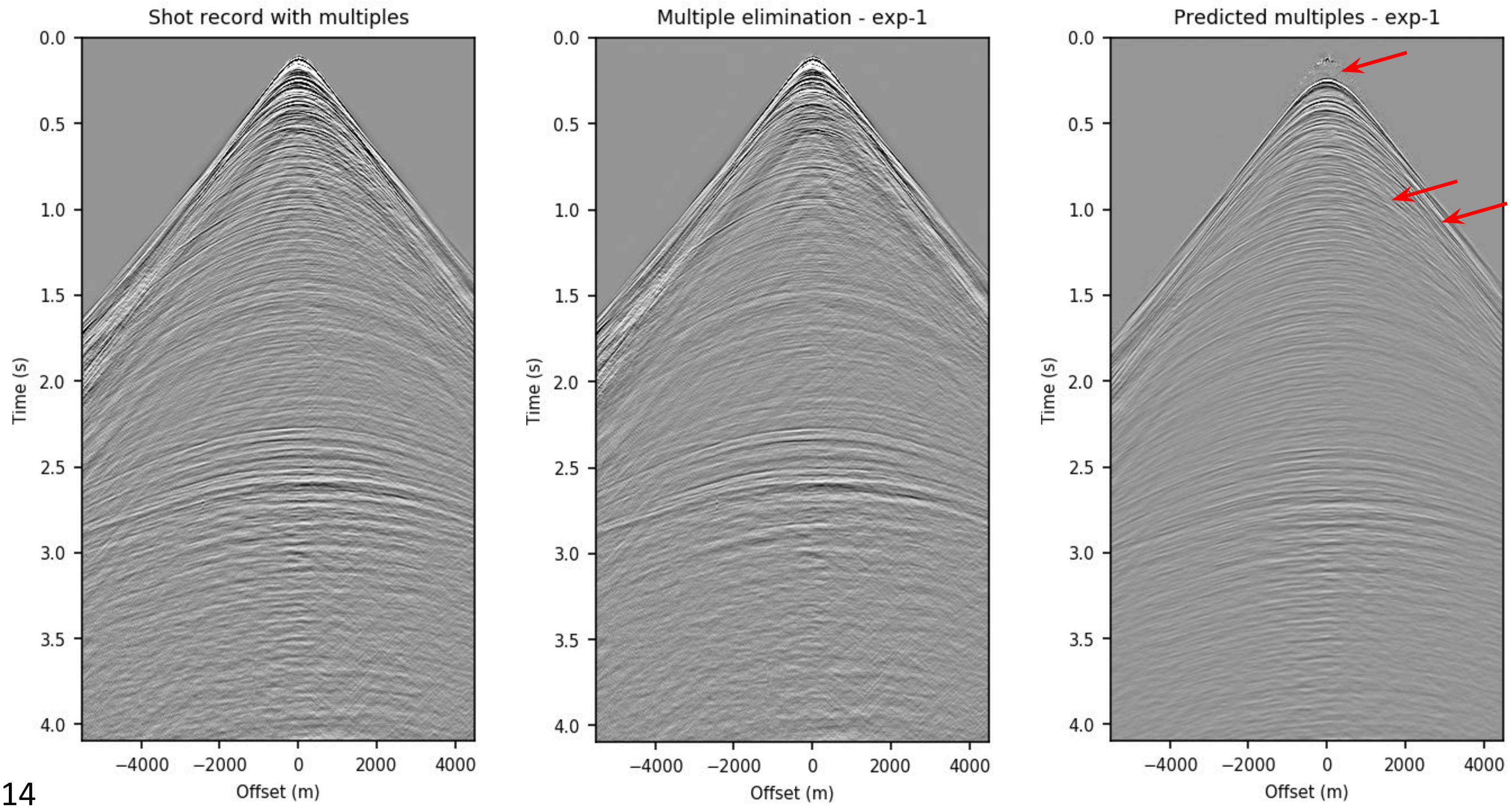
Multiple elimination - Experiment 1



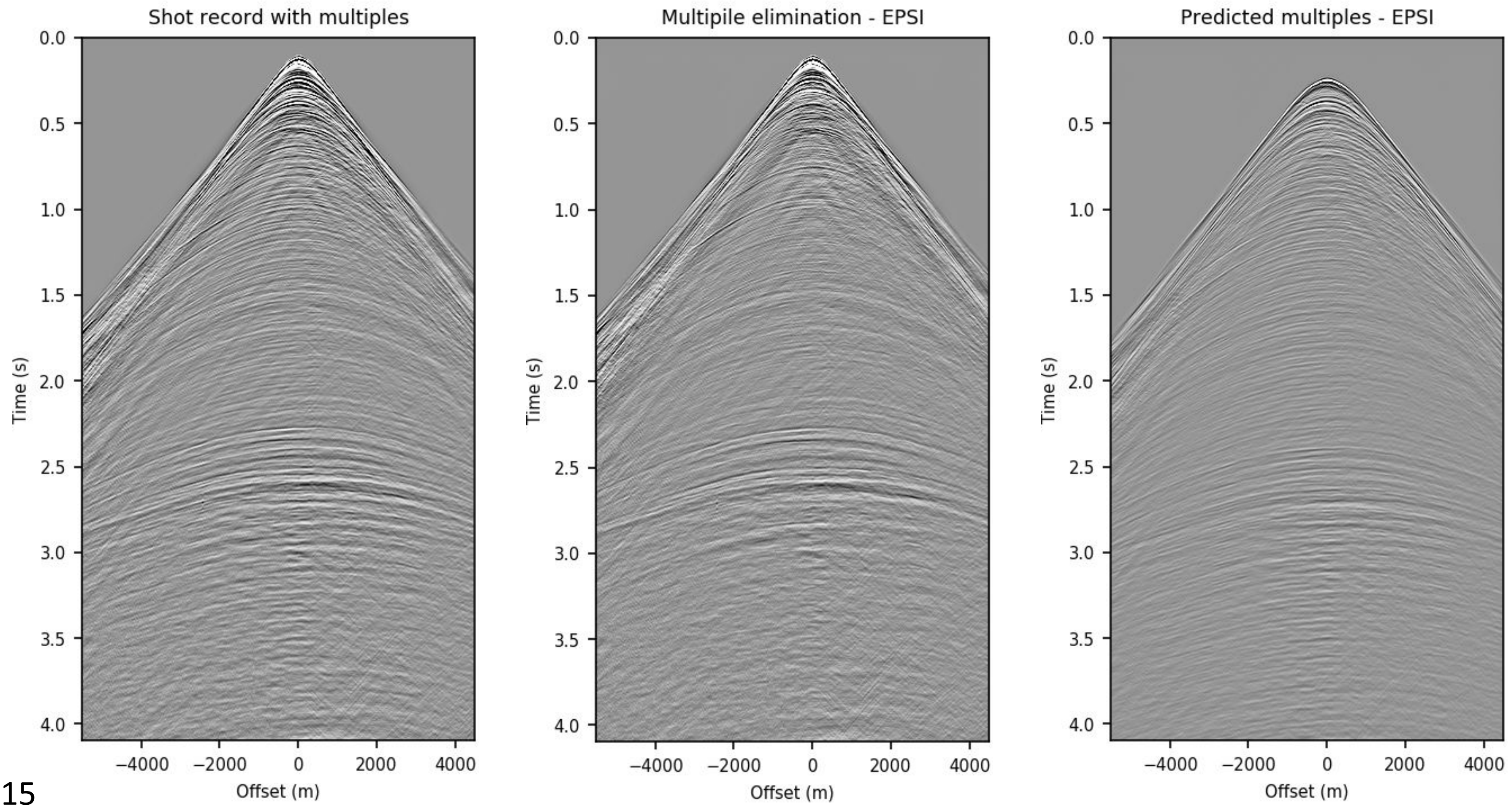
$$\longrightarrow \boxed{g_{\theta_1^*}} \longrightarrow$$



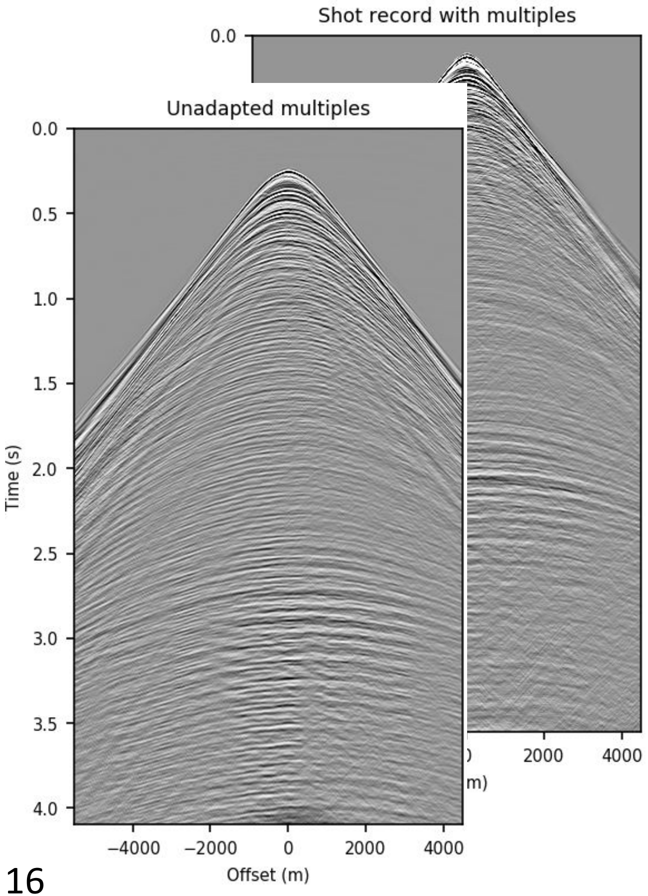
Results on test data - Experiment 1

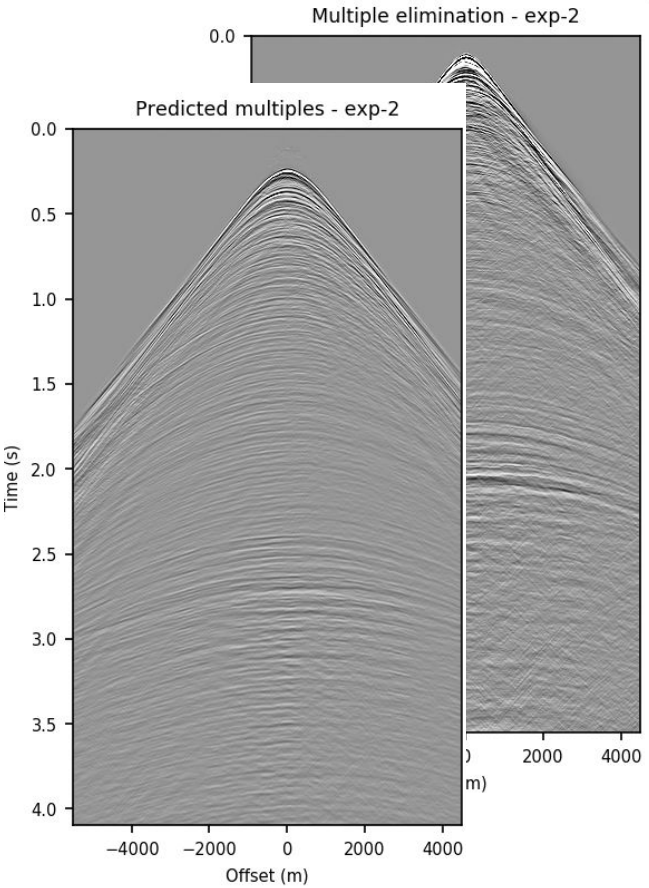


Multiple elimination w/ EPSI

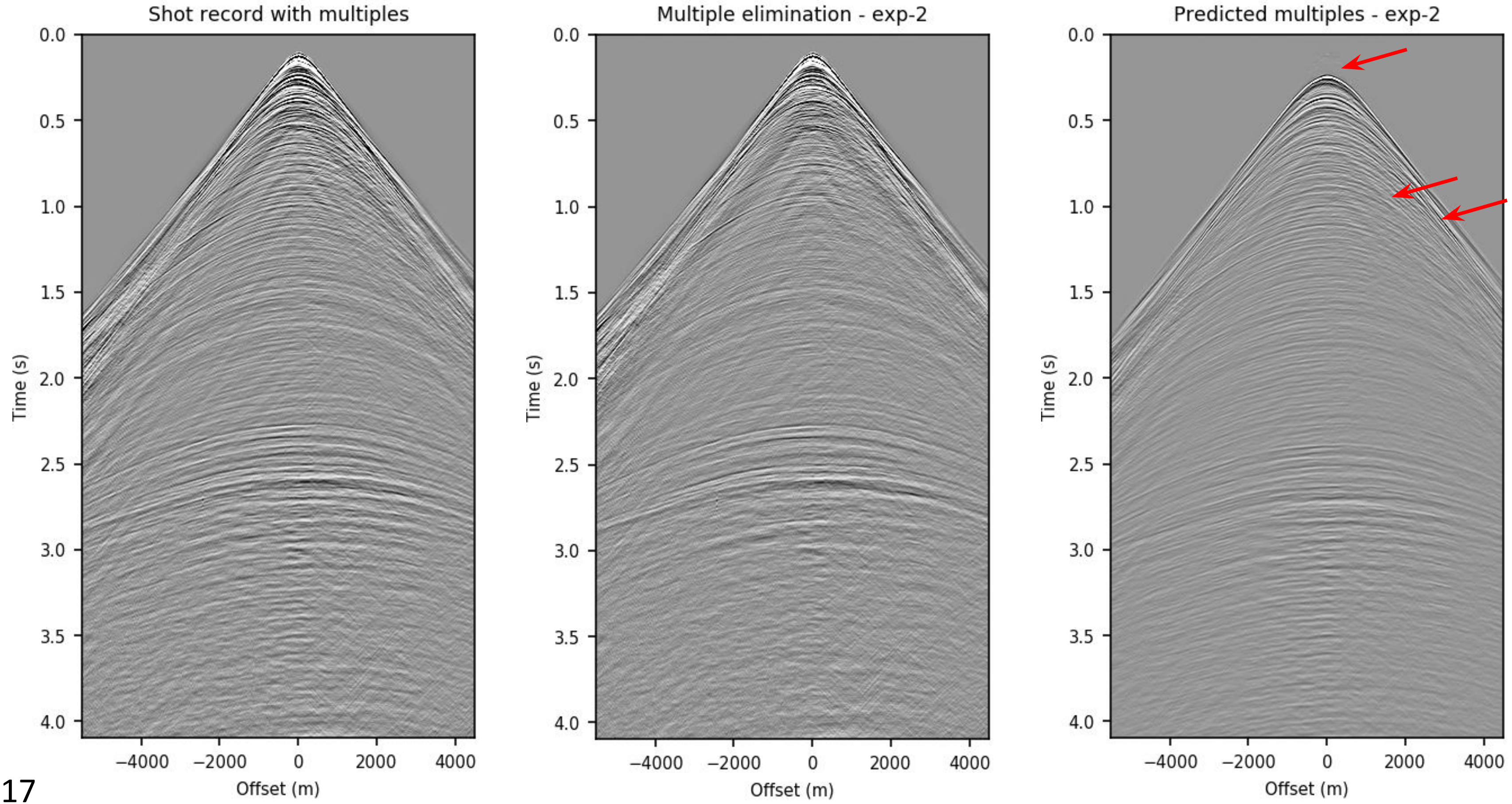


Multiple elimination - Experiment 2

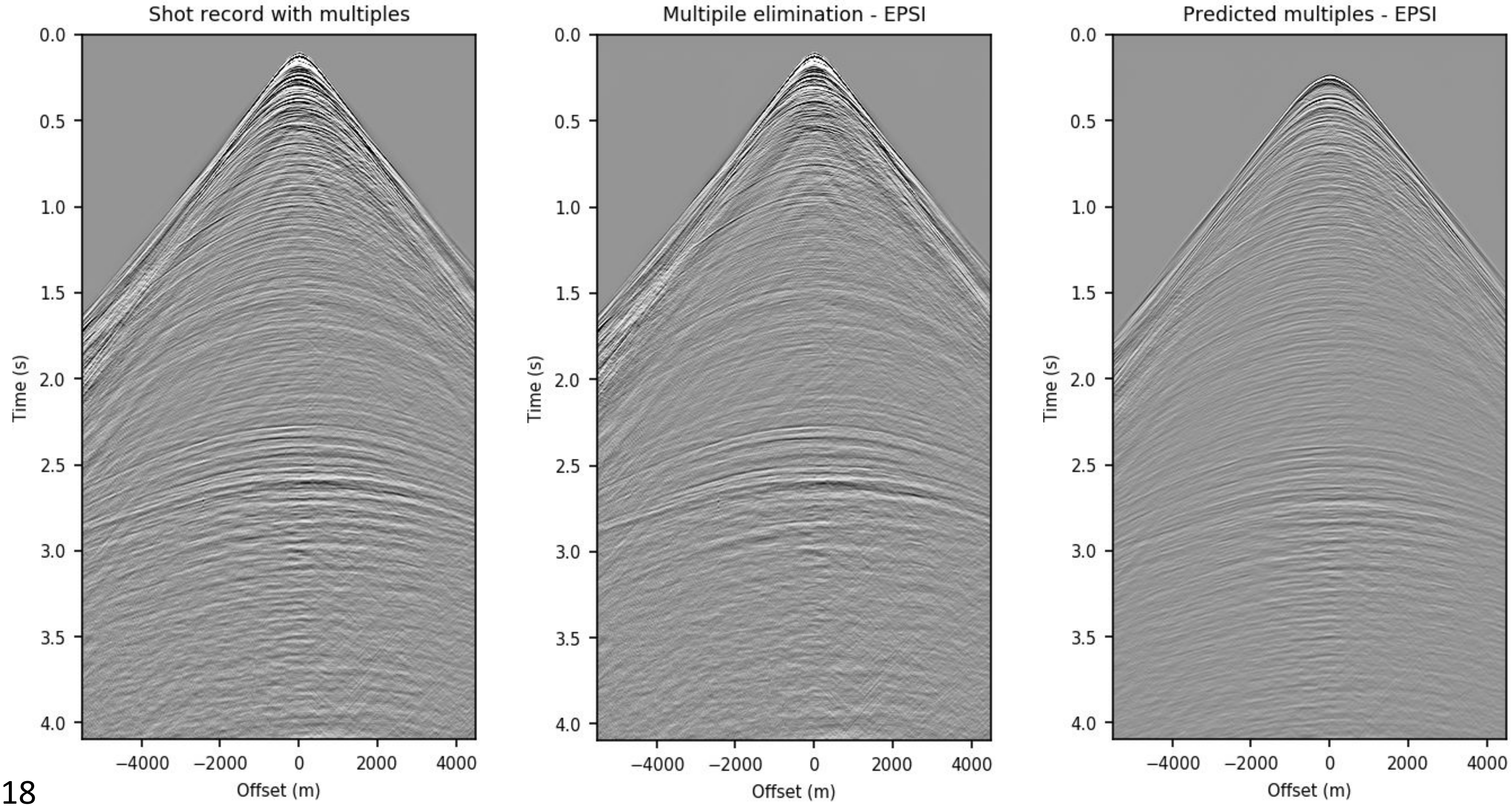




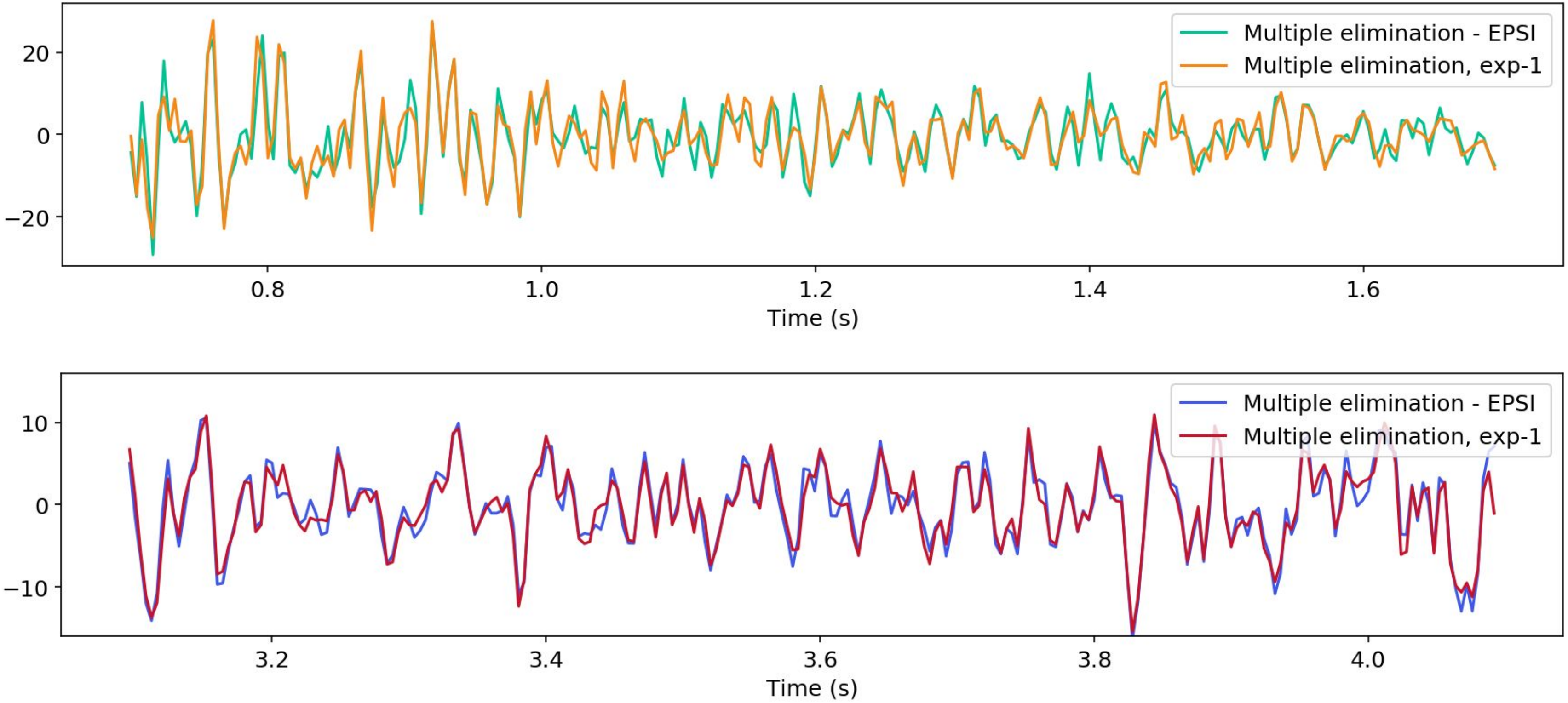
Results on test data - Experiment 2



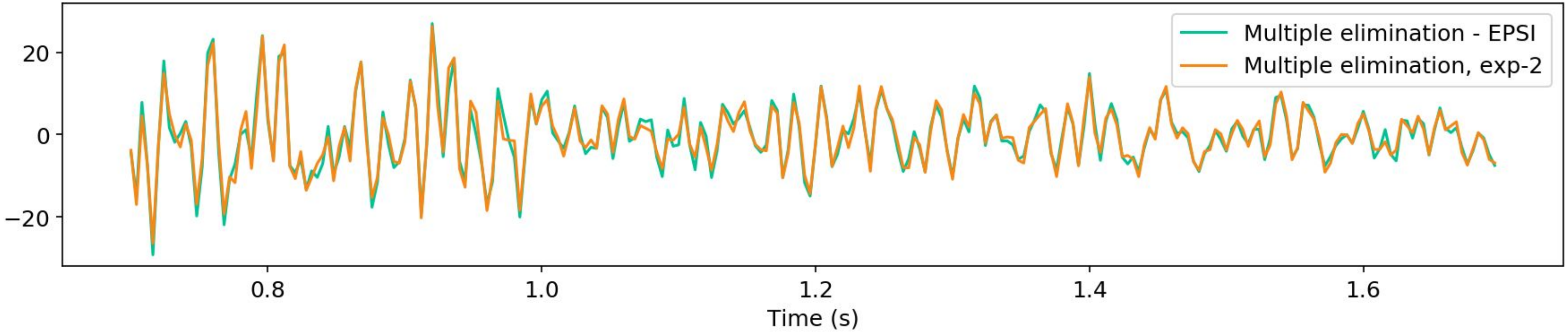
Multiple elimination w/ EPSI

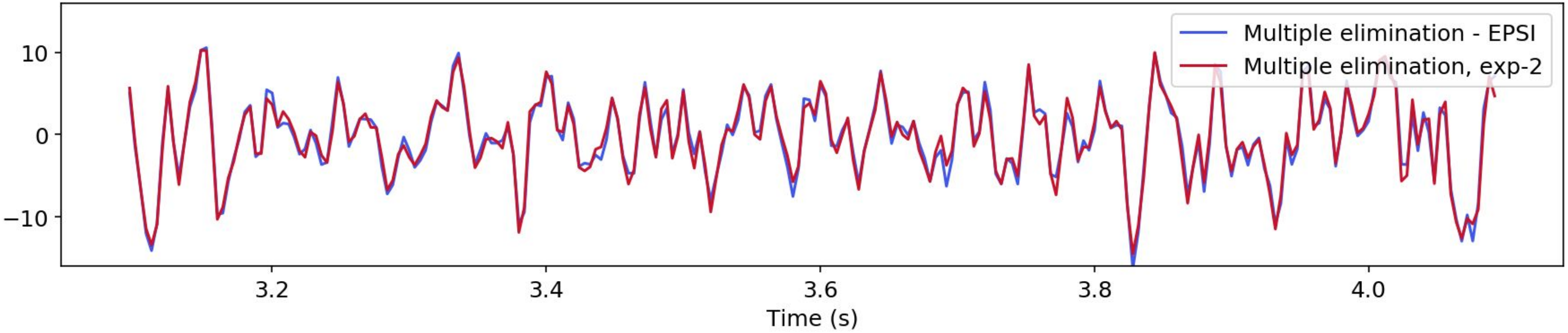


EPSI vs Experiment 1 - 150 m offset trace



EPSI vs Experiment 2 - 150 m offset trace





Questions to investigate

1. Can we approximate the action of EPSI and SMRE via CNNs?
2. Supervised learning: What should be input/output pairs for training CNNs?
3. Can CNNs handle the intricacies of field data?
4. Can the trained CNN be applied to data from another seismic survey?

Siahkoohi A, Louboutin M, Herrmann FJ. The importance of transfer learning in seismic modeling and imaging. Geophysics. 2019 Jul 25;84(6):1-30
Yosinski, J., J. Clune, Y. Bengio, and H. Lipson, 2014, How transferable are features in deep neural networks?: Proceedings of the 27th International Conference on Neural Information Processing Systems, 3320–3328.

SLIM

Does it generalize?

CNNs maintain the quality of performance if,

- ► data from new survey is drawn from the same distribution as training data

Challenging, because of the Earth's heterogeneity and differing acquisition settings

Siahkoohi et al. (2019) demonstrate that transfer learning can be used to finetune a pre-trained network to the pertinent survey

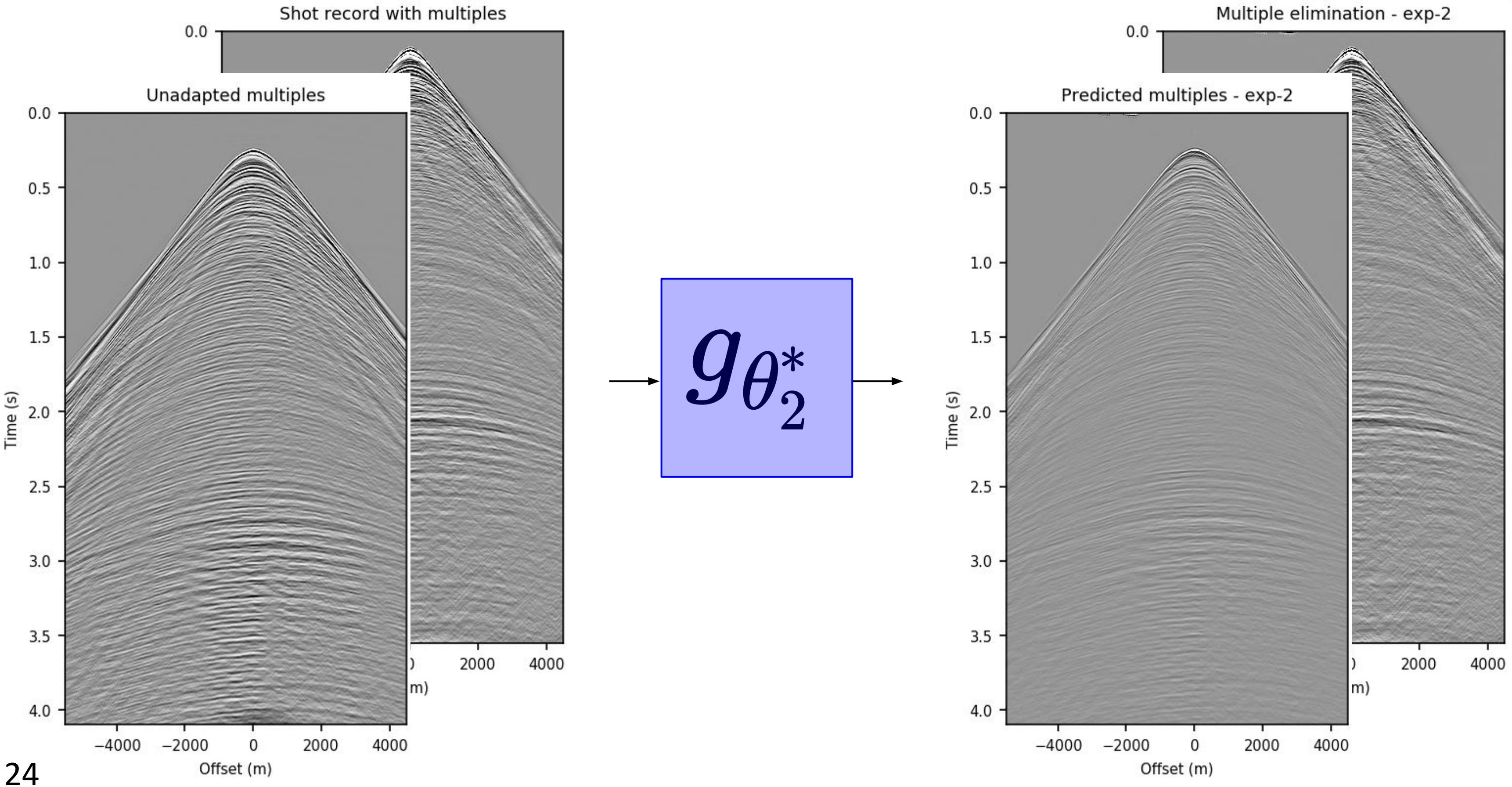
Does it generalize?

To mimic this situation,

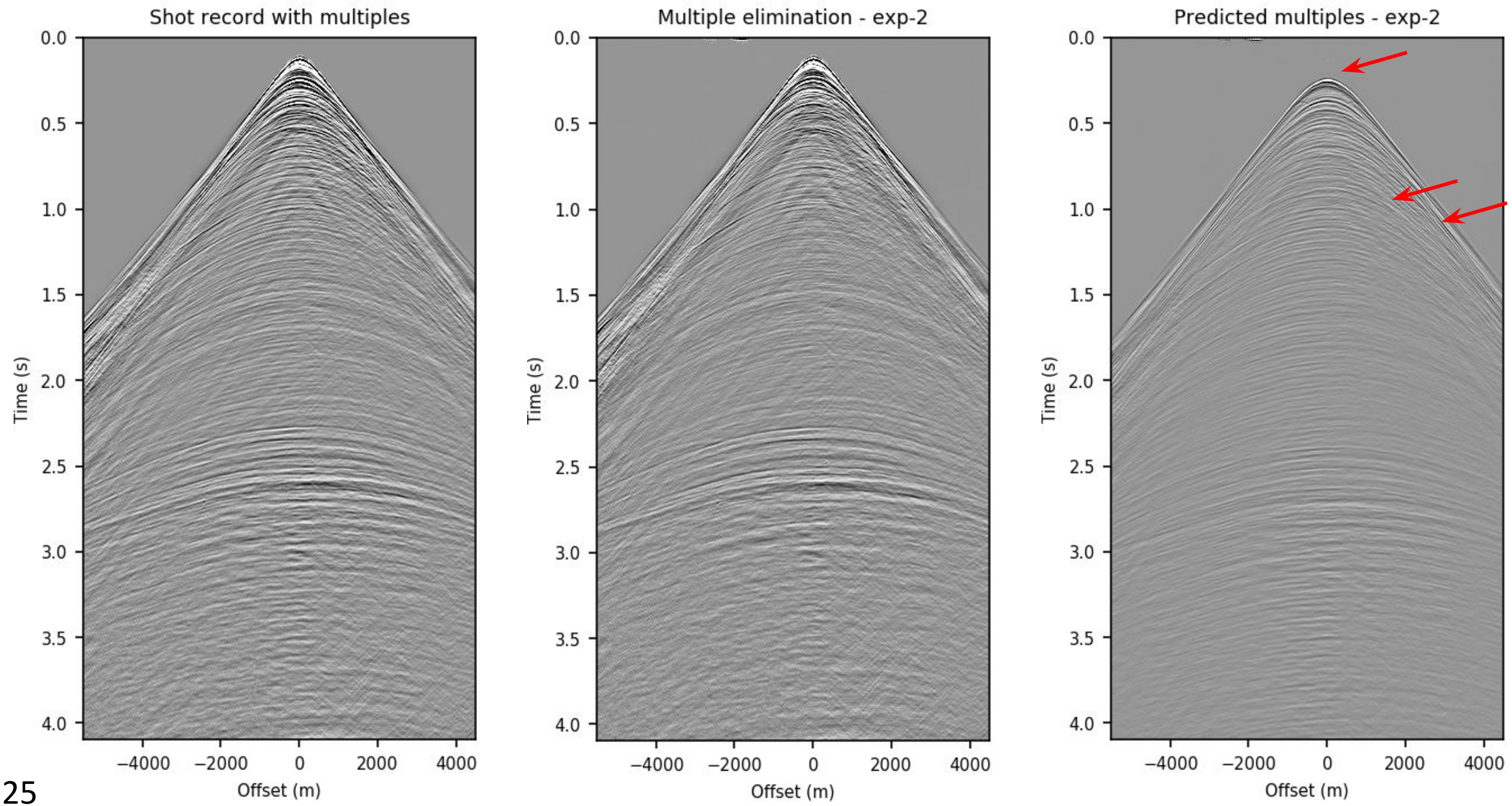
- we train on shots in first half of the seismic line, and
- we test on the rest fo the seismic line
- i.e., non-overlapping

To add variety to training data, we include flipped shot records w/ respect to offset axis during training

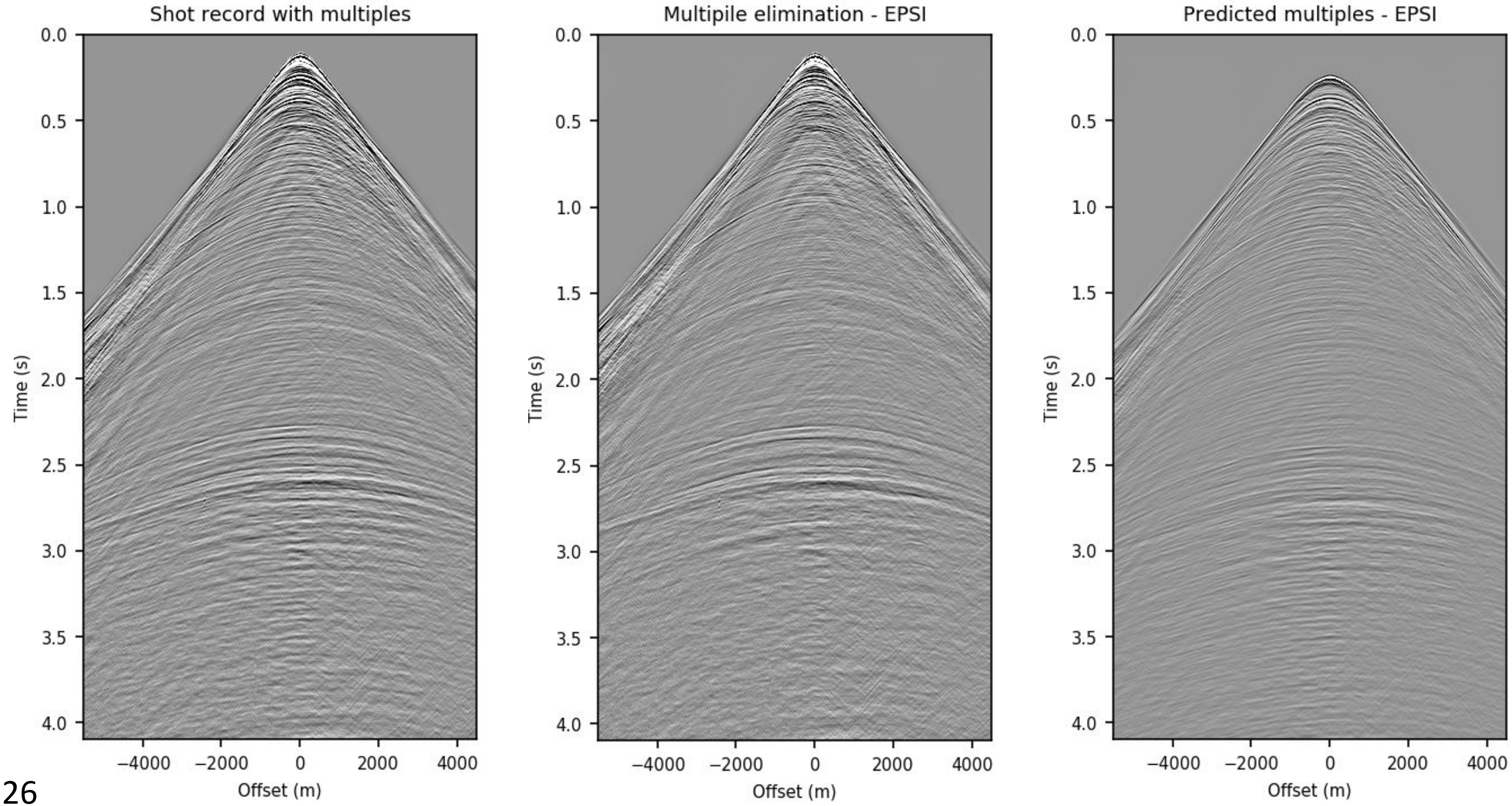
Experiment 2 - overlapping vs non-overlapping



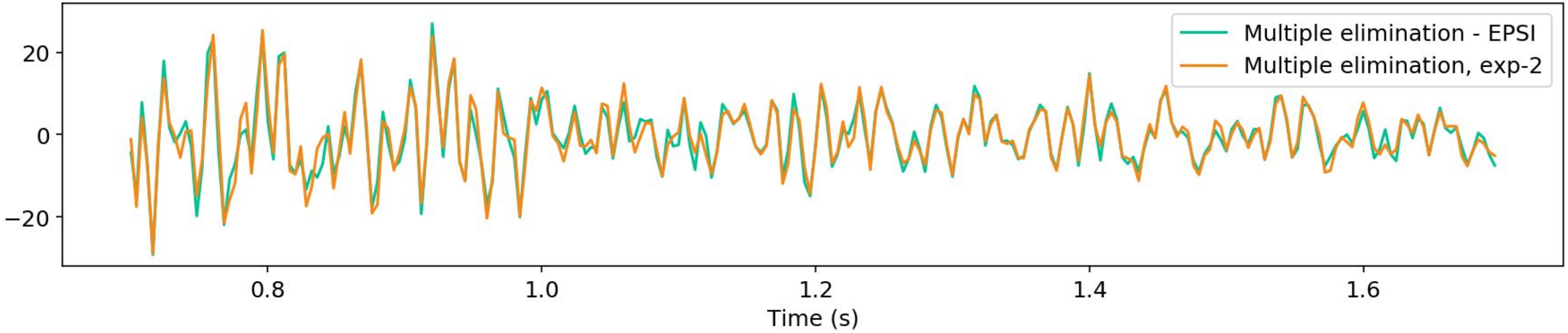
Experiment 2 - non-overlapping train/test set

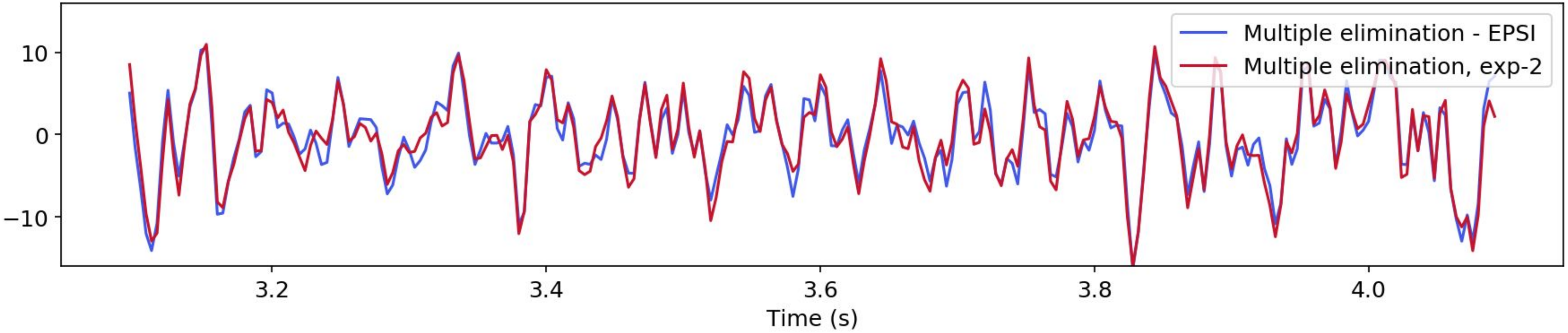


Multiple elimination w/ EPSI

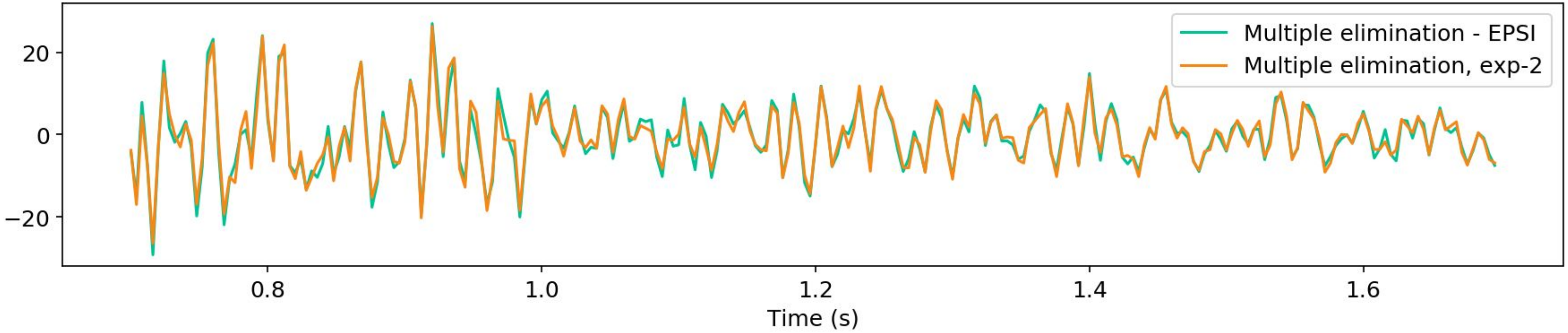


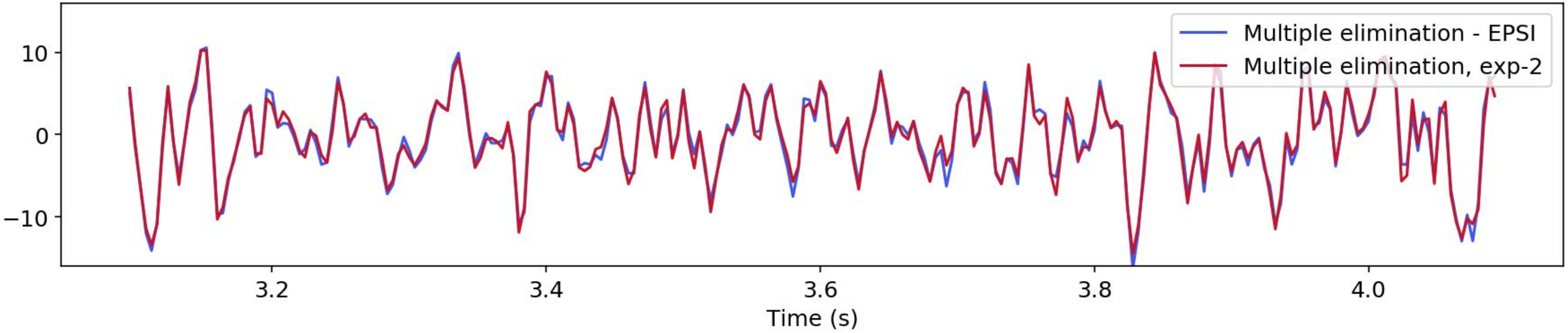
EPSI vs Experiment 2 - non-overlapping





EPSI vs Experiment 2 - overlapping





Conclusions

Providing the CNN with a relatively cheap prediction of multiples leads to better results

CNNs are able to approximate the action of EPSI while preserving the intricate details in field data

The CNN generalized, to some extent, we we used non-overlapping training/testing sets

Future directions: Pre-training a CNN on neighbouring survey + finetuning with small percentage of processed data from pertinent survey