Superior Ultra-High Resolution Seismic Deghosting through Deep Learning with a Custom Autocorrelation Loss

M.O. Faouzi Zizi[1], B. Farmani[1], O. Baramony[1], N. Halbert[1]

[1] TGS

Summary

This paper introduces an innovative deep learning (DL) approach for deghosting Ultra-High Resolution Seismic (UHRS) data, addressing the limitations of traditional physics-based methods, which are computationally expensive due to the shallow receiver depths and fine temporal sampling in UHRS surveys. DL methods have also been developed for fast-track UHRS deghosting but risk learning residual ghosts if the ground truth data is poorly parameterized. To overcome these challenges, the paper proposes a custom loss function that combines the Mean Absolute Error (MAE) with an autocorrelation penalty, designed to suppress ghost residuals during training without relying heavily on perfect ground truth data. This approach enables efficient deghosting, even when ground truth data cannot be fully optimized. Results from a 2024 UHRS survey demonstrate that the proposed method (ACL-DL) outperforms conventional DL models, achieving superior ghost removal and cleaner seismic outputs, while minimizing the need for precise parameterization. This method provides a robust, automated, and fast alternative to deghost UHRS data, offering improved performance in real-world data conditions.

Superior Ultra-High Resolution Seismic Deghosting through Deep Learning with a Custom Autocorrelation Loss

Introduction

In seismic data acquisition, ghost reflections from the sea surface introduce unwanted artifacts that degrade the quality of subsurface images. Over the years, various processing-based methods have been developed to tackle receiver deghosting and improve imaging results. The delay in receiver ghost reflections is primarily influenced by the angle of incidence and the depth of the towed streamer. As a result, traditional deghosting techniques typically rely on plane-wave decomposition and are applied in domains such as the $f-k$ domain (Day et al., 2013) or the $\tau-p$ domain (Masoomzadeh and Woodburn, 2013).

3D Ultra-High Resolution Seismic (UHRS) data provides detailed imaging of the subsurface, capturing features like diffraction patterns and physical properties crucial for accurate windfarm site assessments. In comparison to conventional seismic surveys, UHRS surveys involve receivers that are deployed at shallower depths. These shallow receiver depths, combined with fine temporal sampling, make UHRS data more sensitive to variations in the receivers' positions. As a result, the dynamic fluctuations in receiver depths have a more pronounced effect on UHRS data quality. Consequently, conventional deghosting techniques often fall short when applied to UHRS data (Provenzano et al., 2020).

Inversion-based deghosting methods, such as the approach proposed by Bekara et al. (2024), estimate receiver depths and suppress ghost reflections using deconvolution in the τ-p domain. While effective, these methods require significant computational resources and precise parameter settings. Inaccurate receiver depth estimation can lead to suboptimal ghost removal. Recent advancements in Deep Learning (DL) offer a faster, automated alternative. Farmani et al. (2024) showed that DL models trained on a small portion of conventionally processed UHRS data can effectively deghost the remaining survey data, provided the acquisition configuration and receiver depth variations remain consistent. Unlike inversion-based methods, DL approaches are faster and do not require user-defined parameters. However, conventional DL models rely heavily on the ground truth used during training. If the receiver depths in new data differ significantly due to factors like rough sea conditions, the network’s performance declines, necessitating retraining. Moreover, generating perfect ground truth data is challenging as residual ghosts may persist due to imperfect parameterization during conventional deghosting. These limitations underscore the need for a more robust deghosting approach that minimizes reliance on ground truth parameterization and generalizes well across varying conditions.

To address these challenges, we propose a novel deep learning (DL) approach that utilizes a custom loss function composed of two terms: the Mean Absolute Error (MAE) - commonly used in conventional DL methods for UHRS deghosting, such as in Farmani et al. (2024) - and an autocorrelation penalty term based on the autocorrelation function (ACF). The autocorrelation penalty effectively suppresses unwanted ghost patterns, reducing the model’s dependence on ground truth parameterization. By penalizing residual ghost patterns during training, the model produces cleaner outputs and achieves superior deghosting performance. This approach maintains the speed and automation benefits of DL while improving robustness for real field data. Results demonstrate that our autocorrelation loss-based DL method (ACL-DL) outperforms conventional DL methods, providing more reliable deghosting for UHRS data.

Methodology

In this work, we utilize the same RIDNet network used by Farmani et al. (2024). To enhance the ability of the model to suppress ghost residuals, we design a custom loss function by incorporating an autocorrelation penalty informed by ghost period calculations. Training data is generated using a data generator that extracts patches from the input seismic data and computes masks based on ghost periods. We compute the ghost period for each trace of the input data patches based on the receiver depth and acquisition geometry with the following equation:

$$\tau = \frac{2\, r_{depth} \cos(\alpha)}{v} \quad (1)$$

Here, r_{depth} is the estimated receiver depth, v is the water velocity, and α is the angle of incidence. These ghost periods indicate the likely locations of ghost reflections in the input data, allowing us to generate a mask m that highlights regions affected by ghost patterns. The resulting mask is in binary form, where regions within the ghost window are set to 1, and all other regions are set to 0. Let f_θ represent the RIDNet network parameterized by θ, and x be the input seismic data containing ghost energy. During training, input patches are processed by the network to produce a predicted ghost-free signal $\hat{s} = f_\theta(x)$. Our custom loss function $\mathcal{L}$ is defined as the sum of the MAE and an autocorrelation penalty term:

$$\mathcal{L} = \frac{1}{N}\sum_{1}^{N}|s_i - \hat{s}_i| + \lambda\, \frac{1}{N}\sum_{1}^{N}|\hat{s}_i^{autocorr}\, .\, m_i| \quad (2)$$

where: s is the ground truth signal, N is the number of data points, $\hat{s}^{autocorr}$ = ACF($\hat{s}$) is the autocorrelation of the predicted signal, m_i is a mask sample based on ghost sample locations, and λ is a scaling factor controlling the contribution of the autocorrelation penalty. The first MAE term measures the errors between the ground truth and the predicted patches. However, if the ground truth is generated using a parameterization that does not fully attenuate ghosts in all shot gathers — for example, due to ambiguity in receiver depth tracking in some gathers — the network may learn to retain residual ghost energy. To address this, the second term of our loss function applies an autocorrelation penalty to the predicted patches, weighted by the ghost period locations. The binary mask ensures that the loss penalizes errors only in the regions where ghost energy is expected. High autocorrelation in these regions indicates the presence of ghost residuals. By penalizing this autocorrelation, the model learns to suppress ghost residuals effectively. This approach allows the model to reduce ghost patterns independently of ground truth parameterizations, as the second term of the loss function (equation 2) does not rely on ground truth data. The combination of the MAE term, which ensures signal accuracy, and the autocorrelation penalty, which targets ghost suppression, maintains a balance between ghost removal and signal preservation. During training, the network parameters θ are updated via gradient descent, enhancing RIDNet's ability to outperform conventional DL methods in ghost removal.

Results

The data used in this study come from a 3D UHRS survey conducted in 2024. We trained both a conventional DL method and our ACL-DL method using the same input and ground truth data from an early-acquired sequence of the survey. Both models were then applied to the remaining survey data. The ground truth data were generated using an inversion-based $\tau - p$ method designed for UHRS data. The key difference between the DL methods lies in the loss function: the conventional DL method uses only the MAE term, while our ACL-DL method uses the custom loss function (equation 2) described in the methodology section. Figure 1 shows detail of a shot gather from a sequence not included in the

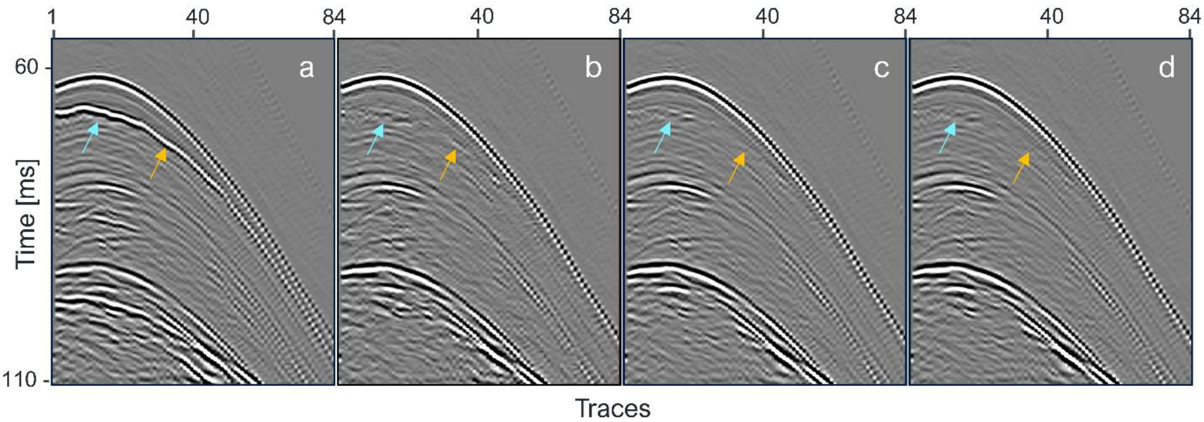

***Figure 1**. Detail of a shot gather before and after deghosting. a) Input with ghost, b) Deghosted with the $\tau - p$ method, which is used to generate ground truth data, c) Deghosted with the conventional DL, d) Deghosted with the ACL-DL method.*

training data, before deghosting (1-a) and after applying: the τ-p method (1-b), the conventional DL method (1-c), and our ACL-DL method (1-d). The τ-p method uses the same parameterization as the training data for a fair comparison, though better results could theoretically be achieved with optimal parameterization. This figure demonstrates that the conventional DL method generalizes well, achieving slightly better results than the τ-p method (orange arrows) despite being trained on a limited range of receiver depths from a different sequence. However, our ACL-DL method outperforms the conventional DL method, showing fewer ghost residuals (blue arrows) and delivering superior ghost removal.

Figure 2 shows a comparison of part of a QC stacked line before and after deghosting. The conventional DL method effectively reduces most ghost energy, but some residual patterns remain (blue arrows). In contrast, the ACL-DL method demonstrates superior ghost suppression, with better removal of the residual ghosts (green arrows). Figure 3 presents a full QC stacked line with ghost characteristics that significantly differ from those in the training data. After applying the conventional DL method, ghost energy is reduced, but a few residual ghost patterns persist (blue arrows). Our ACL-DL method, however, more effectively removes this ghost energy (green arrows), resulting in cleaner and more continuous seismic reflections. These results highlight the effectiveness of the autocorrelation penalty in our custom loss function. This penalty allows the model to generalize better and achieve more reliable ghost suppression, even when data conditions vary.

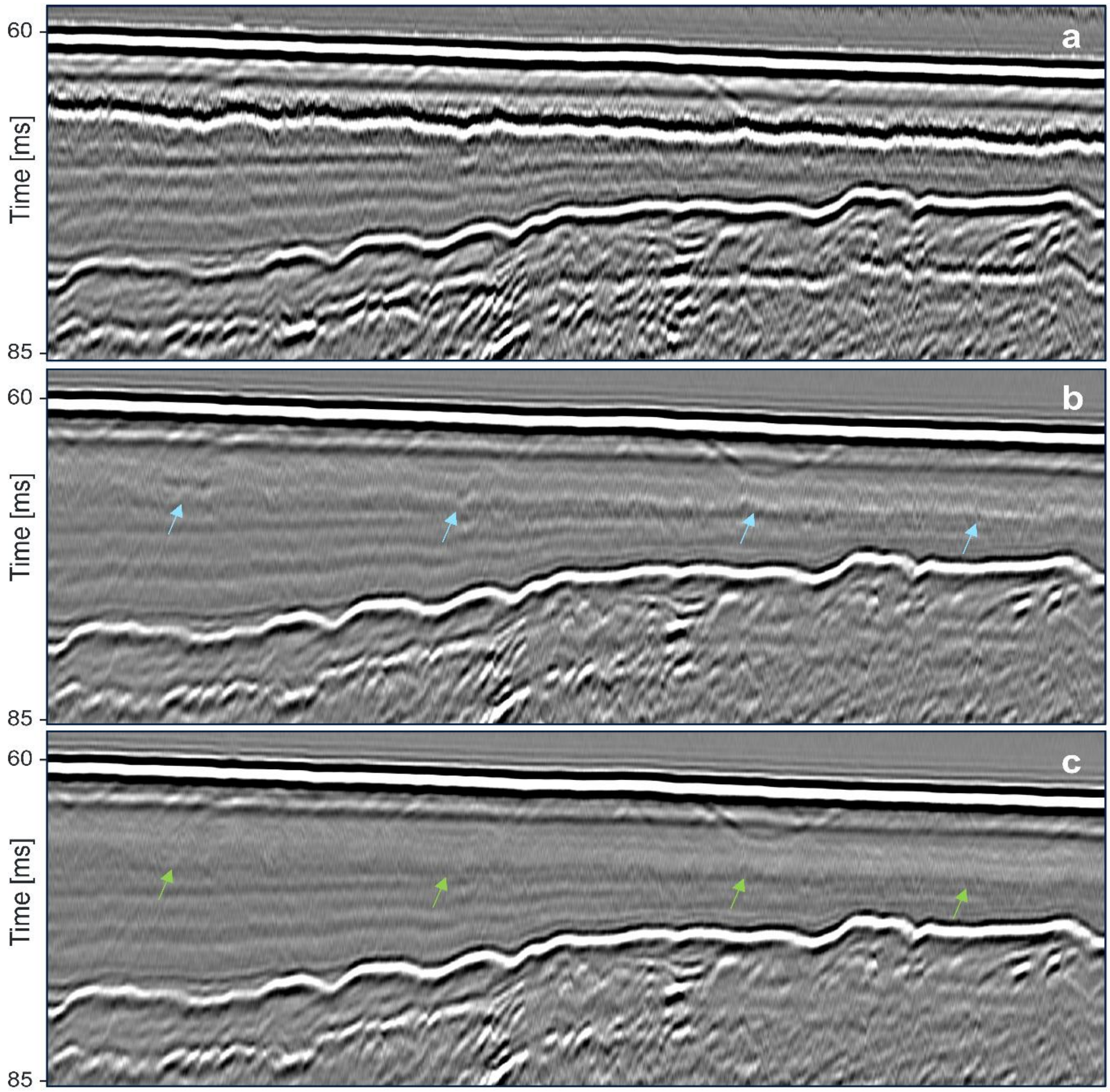

Figure 2. *Detail of a QC stacked line before and after deghosting. a) Input before deghosting, b) output after the conventional DL deghosting method, c) output after the ACL-DL deghosting method.*

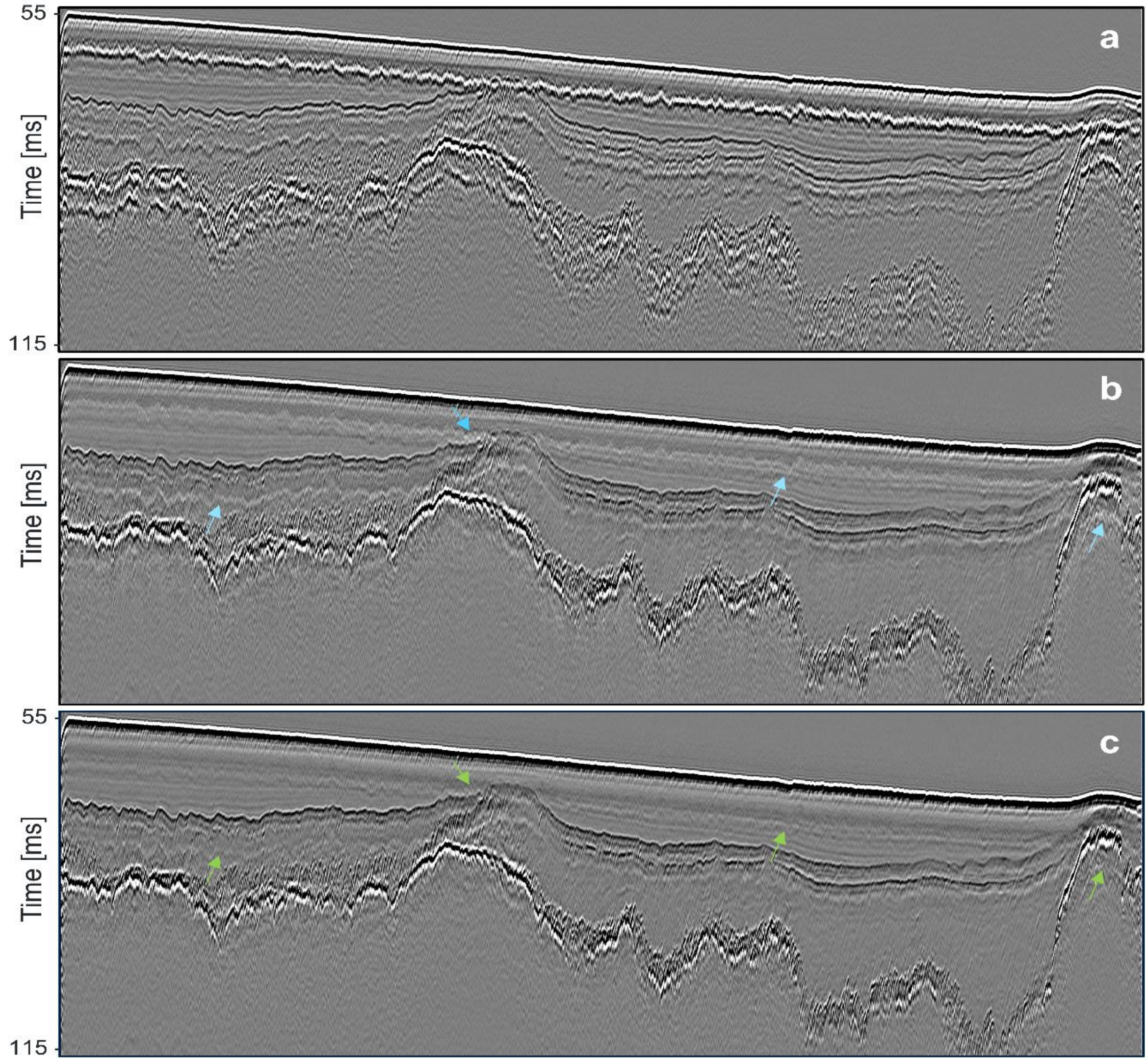

Figure 3. *Full QC stacked line before and after deghosting. a) Input before deghosting, b) output after the conventional DL deghosting method, c) output after the ACL-DL deghosting method.*

Conclusions

Our deep learning approach, with a custom loss function incorporating an autocorrelation penalty, overcomes the limitations of conventional DL methods for UHRS deghosting. When perfect training data is difficult to generate, this method effectively suppresses ghost residuals and reduces reliance on precise ground truth parametrization. Results show that ACL-DL maintains the advantages of DL while delivering cleaner seismic outputs, outperforming conventional DL methods across varying conditions.

References

Bekara, M., Davison, C., Lange, M., and Limonta, L. [2024]. Deghosting of UHR seismic data via dynamic ghost tracking. *85th EAGE Annual Conference & Exhibition*, Jun 2024, Volume 2024.

Day, A., Klüver, T., Söllner, W., Tabti, H., and Carlson, D. [2013]. Wavefield-separation methods for dual-sensor towed-streamer data. *Geophysics*, 78, WA55-WA70.

Farmani, B., Hlebnikov, V., Lange, M., and Lesnes, M. [2024]. Deghosting of Ultra-High Resolution Seismic Data with Deep Learning. *Fifth EAGE Global Energy Transition Conference & Exhibition (GET 2024)*, Nov 2024, Volume 2024, p.1 – 5.

Masoomzadeh, H., and Woodburn, N. [2013]. Broadband processing of conventional streamer data Optimized deghosting in the tau-p domain. *75th EAGE Annual Conference and Exhibition*.

Provenzano, G., Henstock, T.J., Bull, J.M., and Bayrakci G. [2020] Attenuation of receiver ghosts in variable-depth streamer high-resolution seismic reflection data. *Marine Geophysical Research*, 41.