An Augmented Machine Learning Framework for Deghosting Multi-Vintage Marine Seismic Data from the Krishna–Godavari Basin, India

G. Stock, D. Brookes, O. Brusova, A. Knowles, S. Mannick, M. Roberts, R. Yates J. Kumar, P. Lee (TGS)*

Introduction

Reprocessing multi-vintage marine seismic data to a consistent broadband standard remains a major challenge, particularly when acquisition parameters vary significantly between surveys and auxiliary acquisition information is incomplete or unavailable. Deghosting is a critical component of modern marine seismic reprocessing, enabling recovery of low- and high-frequency bandwidth, improved amplitude stability, and enhanced temporal resolution through the removal of surface-related ghost reflections. This abstract presents an augmented machine-learning (ML) deghosting workflow applied to a large multi-survey reprocessing project in the Krishna–Godavari (KG) Basin offshore India. The approach combines a generalized ML deghosting model with survey-specific adaptation derived from limited deterministic deghosting results. The workflow delivers stable broadband spectra, improved seismic imaging, and consistent results across surveys with varying tow depths and acquisition geometries, while maintaining efficient turnaround times.

The KG Basin mega-merge project involved the integration of several legacy 3D marine seismic surveys acquired between 2004 and 2021 in the Bay of Bengal offshore India. The objective was to generate a unified broadband seismic volume suitable for consistent basin-scale interpretation across fields. This required harmonization of spectral inconsistencies arising from differing acquisition configurations, where deghosting playing a key role. Sea-surface ghosts create spectral notches and bandwidth loss governed by source and receiver tow depths. Variations in these depths across surveys lead to differing notch frequencies, causing spectral inconsistencies in merged datasets if not properly addressed. While deterministic deghosting methods such as sparse Radon and tau-p approaches perform well when acquisition parameters are accurate and stable, vintage datasets often suffer from incomplete or uncertain depth information. This can result in residual ghosts, ringing artefacts, and spectral instability. These challenges drive the need for alternative or complementary methods that are robust to auxiliary data uncertainty while remaining physically consistent.

Augmented ML Deghosting Methodology

A machine-learning-based deghosting workflow was developed to address these challenges. The base ML model is trained on a large library of pseudo-synthetic seismic data representing a wide range of acquisition conditions. This enables the model to learn the mapping between ghosted and deghosted wavefields without requiring exact acquisition parameters during application. While the generalized model is robust, it may not fully capture survey specific characteristics, particularly when acquisition geometries differ significantly.

To overcome this, an augmentation step is introduced. A small subset of each survey is selected, where deterministic deghosting provides reliable results. These subsets are typically chosen in areas where acquisition conditions are close to nominal survey parameters. These high-quality outputs are used to finetune the general ML model. The resulting survey-specific model is then applied to the full dataset. Since it requires only limited deterministic processing, making it computationally efficient while improving spectral balance and stability.

Results

The workflow was tested on multiple surveys with varying acquisition geometries, including differences in tow depth and ghost characteristics. For surveys with shallow tow depths, ghost notches lie outside the main bandwidth. Both general and augmented ML models improve resolution and event continuity. However, the generalized model can introduce slight high-frequency over-enhancement, whereas the augmented model produces smoother and more balanced spectra. In deeper tow surveys, ghost notches are clearly visible in the input spectra. The generalized model partially corrects these features but may leave residual imbalance. The augmented model, in contrast, more effectively reconstructs energy at notch frequencies and produces stable broadband spectra, leading to improved reflector continuity and seismic character. Overall, the augmented approach consistently outperforms the generalized model in terms of spectral stability and physical plausibility.

Achieving consistency across surveys is critical in multi-vintage merging. An RGB spectral quality control (QC) attribute was used to evaluate spectral balance across the merged dataset. Before deghosting, strong spectral contrasts are observed between surveys - reflecting differences in acquisition. Application of the generalized ML model improves overall consistency but introduces a tendency for high-frequency bias. The augmented ML workflow delivers the best results, producing balanced spectra with minimal variation across survey boundaries. This is evidenced by both RGB QC maps and averaged amplitude spectra, demonstrating uniform broadband behavior across all surveys (Figure1).

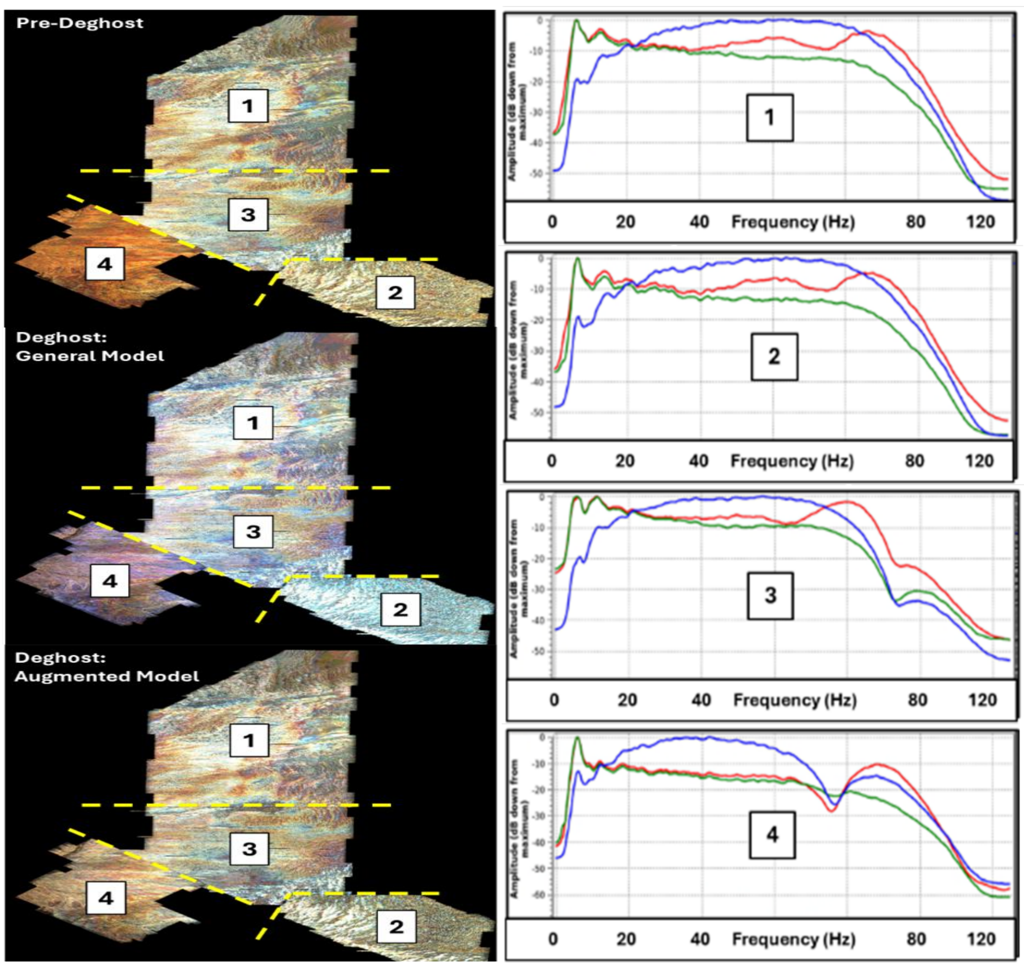

Figure 1: Left image shows RGB maps of merged 3D stacks of the four surveys, (pre-deghosting, deghosting with general model and augmented model) and right image shows Average spectra of each survey (blue = pre-deghosting, red = deghosting with general model, green = deghosting with augmented model)

Conclusions

Deghosting remains a critical challenge in multi-vintage seismic reprocessing, particularly where acquisition parameters are variable and metadata are uncertain. The augmented ML deghosting strategy presented here effectively combines generalized machine-learning models with limited deterministic calibration to address these challenges. Applied to the KG Basin mega-merge project, the workflow successfully produces a consistent broadband dataset suitable for regional interpretation. Furthermore, the adaptive training approach enables continuous improvement of the ML model, supporting future multi-vintage applications.

Acknowledgements

The authors thank the TGS Imaging team in Weybridge, UK, for their contributions to the processing, quality control, and delivery of the Krishna–Godavari Basin reprocessing project.