ANNALS OF GEOPHYSICS, 69, 2026; doi:10.4401/ag-9557
OPEN ACCESS

WigglePy: a QGIS-integrated workflow for converting raster legacy seismic profiles into SEG-Y format

Francesco E. Maesano[1]

(1) Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy

Article history: received March 16, 2026; accepted May 6, 2026

Abstract

Legacy seismic reflection profiles are a critical subsurface resource, essential for seismotectonic and basin modeling, yet many remain accessible only as raster images. This paper presents WigglePy, an open-source QGIS plugin that integrates the entire raster-to-SEG-Y conversion workflow – from image calibration and cleaning to geometry assignment – within a single GIS environment. WigglePy provides an interactive interface utilizing fast extraction algorithms combined with a noise attenuation strategy. A synthetic benchmark shows that the implemented extraction strategies (Fast Fourier Bandpass, Savitzky-Golay morphological reconstruction, and Wiener deconvolution) achieve reconstruction errors comparable to those of matrix-based inversion methods, with improved computational performance. By integrating iterative quality control, spatial data management, and post-extraction spectral spike filtering, WigglePy offers an efficient solution for the scientific rescue and revitalization of vintage seismic archives.

Keywords: Legacy seismic data; Raster-to-SEG-Y conversion; QGIS plugin; WigglePy

1. Introduction

Large archives of vintage seismic reflection data are becoming progressively more important for academic geoscience. Although originally acquired for hydrocarbon exploration, these datasets are essential for constructing 2D-3D geological models, which are also fundamental for identifying and characterizing seismogenic structures onshore and offshore. Globally, the interpretation of seismic reflection data has been pivotal for mapping blind thrusts in densely populated regions like the Los Angeles Basin (Shaw and Shearer, 1999), defining deep fault geometries in slow-deforming intraplate settings (Hao et al., 2015), and identifying tsunamigenic mega-splay faults in subduction zones like the Nankai Trough (Park et al., 2002). Recently, the systematic integration of legacy seismic and bathymetric data has led to the creation of the first comprehensive offshore seismogenic database for Taiwan (Chen and Shyu, 2026).

In the Italian context, seismic reflection profiles are now increasingly relevant for structural geology and seismotectonics (Buttinelli et al., 2022; Feriozzi et al., 2024; Cicala et al., 2024), basin evolution (Amadori et al., 2019; Maesano and D'Ambrogi, 2016), geo-resources and subsurface energy applications (Conti et al., 2022; Diviacco et al.,

2015), and hazard assessment (Antoncecchi et al., 2020). The 3D geological models produced using these data proved essential for the understanding of the seismogenic process, highlighting the role of structural complexity and fault reactivation in the Central Apennines (Buttinelli et al., 2021) and for providing a fast response to the identification of the potential causative fault during a seismic sequence in the Adriatic offshore (Maesano et al., 2023). Moreover, this type of data is fundamental for studying and quantifying the slip rates of buried faults, providing crucial input for seismic hazard assessment (Maesano et al., 2024; Panara et al., 2021).

The Italian case offers a particularly favorable context for raster seismic digitization because large volumes of historical industrial seismic data are available through public declassification initiatives such as ViDEPI (https://www.videpi.com), which include more than 80,000 km of 2D seismic lines. A substantial fraction of this information survives only as scanned raster images. It is important to emphasize that raster seismic data, particularly in wiggle-trace displays, represent the original, processed data; vectorization efforts therefore aim to replicate rather than implicitly improve the signal. However, traditional graphical rendering of seismic profiles introduces static elements – such as timelines, baselines, and annotations – that can obstruct the correct reading and interpretability of the data. Since these elements are fixed on the physical or raster support, they cannot be modified to suit specific interpretive requirements. Transitioning from scanned raster versions to digitally vectorized SEG-Y files is thus essential for two main reasons. First, it significantly enhances readability and integration within modern interpretation platforms, allowing data to be placed in its proper geographical context and co-interpreted with other multiscale input datasets. Second, a vectorized support enables the application of various post-processing filters and seismic attribute analysis (Ercoli et al., 2023), which facilitate the identification of geological features of interest – such as subtle faults (Feriozzi et al., 2024), discontinuities, or the complex 3D architecture of subduction systems (Maesano et al., 2017) – that would otherwise remain obscured in the original paper versions.

The need to convert vintage raster representing seismic reflection profiles is not new. Besides commercial applications in use by hydrocarbon companies, the first academic tools date back twenty years (Miles et al., 2007; Farran, 2012) and were widely used by the scientific community (Owen et al., 2015; Araque-Pérez et al., 2024; Cicala et al., 2024; Conti et al., 2022), whereas recently, standalone software initiatives such as SegyRecover (Pertuz et al., 2025, 2026) have implemented Python-based workflows with similar algorithmic goals to handle scanned seismic sections. These developments align with a broader trend in geosciences toward the systematic recovery and digitization of historical monitoring records (Lorenz et al., 2026) and analogue datasets through open-source, interactive software solutions (Randazzo and Pesci, 2026; Bogiatzis and Ishii, 2016; Lemenkova et al., 2022). Sopher (2018) established a key reference point for this problem by focusing on variable-fill wiggle-trace displays and a band-limited linear inversion framework. This approach was implemented in the WIGGLE2SEGY MATLAB tool and operationally adopted within the SISMOLAB-3D (Laboratory of Seismic Reflection Profiles, https://sismolab3d.ingv.it/), within the INGV, demonstrating high-quality results across various Italian geodynamic settings (Buttinelli et al., 2022).

This paper presents WigglePy, a Python-based platform operating as a QGIS plugin that pursues a similar objective to WIGGLE2SEGY and SegyRecover, but with a different operational philosophy. Unlike previous tools, WigglePy embeds the entire profile reconstruction logic within QGIS. This integration is critical because seismic conversion is rarely done in isolation: many operators work within a GIS project where base maps, CDP locations, and structural layers are already available and must be integrated into a single workflow. Instead of an inversion-centered, standalone workflow, WigglePy provides an interactive, GIS-native pipeline that allows the operator to manage spatial context and trace extraction simultaneously.

This paper aims to: (i) present the WigglePy architecture; (ii) describe the extraction methods exposed to the user; (iii) compare it with the WIGGLE2SEGY workflow at a methodological level; (iv) evaluate its performance via synthetic benchmarks focusing on signal reconstruction limits, and (v) present a comparison of SEG-Y-converted legacy data using different methodological approach.

2. Materials and methods

This section describes the primary features of WigglePy, starting with the graphical interface and workflow, and illustrates the principal processing step, including the methodological strategies used to convert the signals represented in the raster image of a seismic profile.

2.1 Software architecture and setup

WigglePy is implemented in Python as a QGIS plugin, leveraging the native environment to manage spatial data and georeferencing and adopted the segyio (github.com/equinor/segyio) python library to export the SEG-Y files. Architecture reflects the sequential nature of the digitization process, organized into five main interactive modules. The workflow begins with the Setup and Calibration module (Fig. 1), where the user sets up the project by defining crop references for the signal-bearing portion of the raster and entering acquisition parameters extracted from the seismic profile header (e.g., sampling rate, time-variant filter). To ensure conversion repeatability, all setup information – including crop coordinates and header parameters – can be exported to and imported from a standardized JSON file.

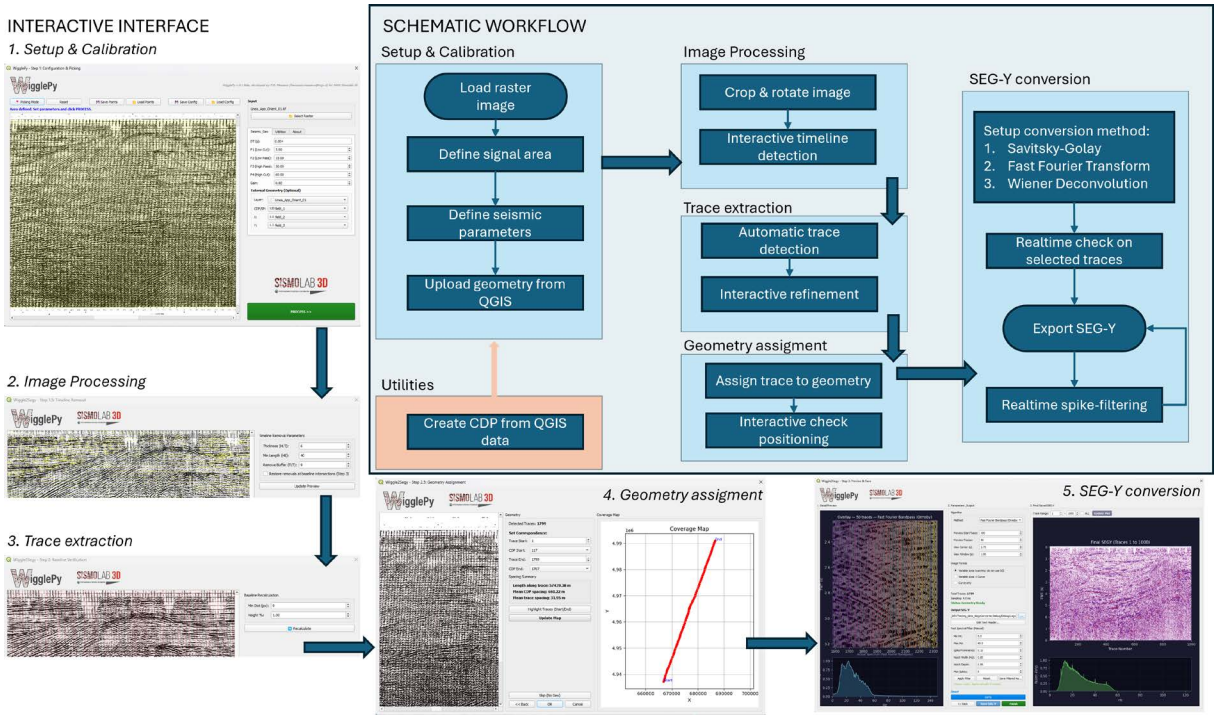

Figure 1. WigglePy schematic workflow and interactive modules appearance. (1) Project setup, including image cropping and input of seismic header acquisition parameters. (2) Timeline detection and suppression module with interactive tuning of line thickness, minimum sequence length, and vertical removal buffer, featuring optional pixel recovery at intersections. (3) Trace detection for modulating automatic identification based on pixel distance thresholds. (4) Geometry assignment module for associating line traces with Common Depth Points (CDP) and displaying a real-time survey coverage map. (5) Interactive SEG-Y conversion, integrating signal extraction and post-processing spectral spike filtering. Two additional utilities allow for creating the CDP geometry from a polyline layer in QGIS and to merge two SEG-Y files.

2.2 Image processing and multi-stage cleaning

The next stage focuses on suppressing non-seismic graphical elements by imaging processing (Fig. 1). WigglePy treats timeline removal as a dedicated interactive module. The algorithm implements a morphological filtering approach, allowing the user to fine-tune three key parameters in real time: the line thickness (in pixels), the minimum length of a pixel sequence to be classified as a line, and a vertical buffer for erasure. An optional feature restores signal pixels inadvertently removed at intersections between timelines and trace baselines before the extraction phase.

Following timeline suppression, the Trace detection panel enables automatic identification of individual seismic traces. This detection can be tuned by adjusting the minimum pixel distance between traces, ensuring correct signal separation even in high-density or low-resolution scans.

2.3 Geometry assignment and GIS integration

A distinctive feature of WigglePy is the geometry assignment module (Fig. 1), which enables the direct association of the raster line with its geographical position. This GIS-native approach allows the operator to link trace numbers to specific Common Depth Points (CDP) derived from georeferenced base maps. A real-time map viewer displays the portion of the seismic survey covered by the raster profile currently under processing, ensuring spatial consistency. A dedicated utility, accessible from the Main window, significantly simplifies the operations required for creating seismic line geometry. This tool enables the generation of navigation/positioning data directly from a digitized QGIS line. The utility allows the user to visualize the profile raster concurrently with the input phase. This concurrent visualization is essential for interactively determining the appropriate CDP/shot points spacing and numbering. It supports both regular-interval assignment and the accommodation of common irregularities often present in original vintage data, such as non-standard starting and ending CDP/shot-point values or intermediate variations across lines segmented across multiple raster sections.

2.4 Raster-to-signal extraction methods

The core of the WigglePy engine lies in transforming spatial pixel distributions into continuous 1D amplitude series (Fig. 1). This process is executed in two stages: raw pixel-to-amplitude mapping followed by numerical conditioning to restore the seismic waveform.

2.4.1 Raw trace extraction (Mapping)

Before any spectral conditioning, the plugin converts the raster image into a numerical series using one of three mapping modes. To transform the discrete pixel distribution into a numerical signal, WigglePy implements a mapping process that operates on a row-by-row basis. For a binary image where foreground (trace) pixels are defined as $I(x, t) = 1$ and background as 0, the raw amplitude $A(t)$ at each sample t is calculated within an extraction window W according to one of the following modes:

- **Variable Area (VA)**: This mode treats the pixel distribution as a proxy for the signal envelope. The amplitude is defined by the horizontal pixel occupation relative to the baseline:

$$A(t) = \sum_{x \in W} I(x, t)$$

- **Curve-only (CO)**: This mode identifies the horizontal displacement of the wiggle edge $x_{edge}(t)$ relative to its local baseline x_0:

$$A(t) = x_{edge}(t) - x_0$$

- **Variable Area + Curve (VA+CO)**: This hybrid mode recovers the oscillatory nature of the original pulse from filled displays by applying a discrete numerical gradient (∇) to the VA count:

$$A(t) = \nabla \left(\sum_{x \in W} I(x, t) \right)$$

Amplitude Scaling and Normalization

Since the mapping from pixels to numerical values is inherently qualitative – being influenced by scanning resolution, line thickness, and original ink intensity – WigglePy uses a standardized amplitude domain through a global statistical normalization. After raw extraction, the entire dataset is scaled by the 99th percentile (P_{99}) of the absolute values of the section:

$$A_{norm}(t) = \frac{A(t)}{P_{99}(|A|)}$$

This approach ensures that the relative amplitude variations are preserved across the entire seismic line while mapping the variable pixel-scale measurements into a consistent, normalized range suitable for SEG-Y exportation and subsequent processing.

2.4.2 Signal conditioning and wavelet reconstruction

While the framework established in WIGGLE2SEGY (Sopher, 2018) relies on a formal band-limited linear inversion – treating the reconstruction as an inverse problem solved via large matrices for each trace – WigglePy adopts a direct signal-conditioning philosophy. The objective is to achieve equivalent waveform restoration using computationally lighter spectral and morphological operators that facilitate interactive, real-time adjustments within the GIS environment.

The input for all conditioning paths is the raw or normalized amplitude signal $d(t)$ (or its discrete form d_n) obtained from the mapping process described in Section 2.4.1 (i.e., $d(t) = A_{norm}(t)$). After mapping, this signal requires conditioning to suppress scanning artefacts and reconstruct the missing portions of the band-limited seismic wavelet (e.g., negative lobes in VA sections). WigglePy exposes three distinct conditioning paths:

1) Savitzky-Golay (Morphological path): This method applies a local least-squares polynomial smoothing (Niedźwiecki et al., 2021; Savitzky and Golay, 1964). Given the discrete input signal d_n derived from the pixel mapping, the smoothed output s_n is obtained via convolution with polynomial coefficients C_m:

$$s_n = \sum_{m=-(w-1)/2}^{(w-1)/2} C_m d_{n+m}$$

   By tuning the window length w and polynomial degree p, the operator can effectively attenuate high-frequency noise while preserving the relative amplitude and timing of reflectors (Liu et al., 2016).
   To strictly enforce band limitations, the plugin cascades this morphological smoothing with a hard frequency cutoff outside the user-defined $[f_1, f_4]$ range.
2) Fast Fourier Bandpass: This represents the most robust and computationally efficient reconstruction strategy. By transforming the mapped signal $d(t)$ into the frequency domain via Fast Fourier Transform (FFT), a trapezoidal bandpass filter (defined by four corner frequencies) is applied. The conditioned signal $y(t)$ is recovered as:

$$y(t) = \mathcal{F}^{-1}\{\mathcal{F}\{d(t) - \mu\} \cdot O(f)\}$$

   where μ is the signal mean, $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse discrete Fourier transforms, and $O(f)$ is the Ormsby mask defined by the corner frequencies (f_1, f_2, f_3, f_4). In VA mode, where negative lobes are absent from the raster, this frequency-domain filtering naturally forces the signal toward a band-limited zero-mean state. This process effectively "reconstructs" the missing negative lobes by redistributing the energy within the seismic passband (Ormsby, 1961).
3) Wiener Deconvolution: This path employs a wavelet-based sharpening strategy. Utilizing the theory of least-squares filtering (Robinson and Treitel, 2000; Treitel, 1969), WigglePy deconvolves the mapped signal $d(t)$ using a synthetic Ricker wavelet as the source signature (Ricker, 1953). The optimal inverse filter $H(f)$ is constructed in the frequency domain as:

$$H(f) = \frac{W^*(f)}{|W(f)|^2 + \lambda \max(|W(f)|^2)}$$

   where $W(f)$ is the spectrum of the Ricker wavelet $w(t)$, $W^*(f)$ is its complex conjugate, and λ is the user-defined Noise-to-Signal Ratio (NSR). The final trace $y(t)$ is obtained by applying the filter to the mapped data and reconvolving with the theoretical wavelet to maintain the desired output bandwidth:

$$y(t) = \mathcal{F}^{-1}\{\mathcal{F}\{d(t) - \mu\} \cdot H(f)\} * w(t)$$

This process aims to sharpen reflectors and improve temporal resolution, which is critical for interpreting closely spaced boundaries in vintage data. As a pitfall, this method may produce significant artefacts if the input parameters differ significantly from those of the acquisition source.

2.5 Post-extraction spectral spike filtering

WigglePy includes an interactive post-processing stage for residual noise attenuation (Fig. 1, panel 5). Despite rigorous cleaning, timeline fragments or periodic-scanning artefacts may persist as high-energy spikes in the frequency domain. By analyzing the average normalized amplitude spectrum, the user can interactively apply narrow-band notch filters to suppress residual noise without degrading the broad-band geological signal. This ensures that the final digital product is suitable for advanced seismic attribute analysis.

Post-extraction spectral spike filtering is a manual post-processing tool available after the SEG-Y generation. Its purpose is to identify and suppress narrow anomalous spikes in the average amplitude spectrum of the extracted section. The workflow operates in two stages: first, the plugin analyses the frequency spectrum of the saved section within a user-defined frequency band set by the Min Hz and Max Hz parameters; within this band, it automatically locates spectral peaks whose prominence exceeds the threshold defined by Spike Prominence. For each detected spike, a notch filter is applied, with a frequency width controlled by Notch Width (Hz) and an attenuation depth set by Notch Depth (for example, a value of 0.85 suppresses the amplitude by 85% at the notch center). The maximum number of spikes to be corrected in a single pass can be capped using Max Spikes. Once the filter is applied via the Apply Filter button, the filtered section can be inspected in the final view panel and, if the result is satisfactory, exported as a separate second SEG-Y file.

3. Performance evaluation: benchmarks and quantitative comparison

To evaluate the performance of the various extraction algorithms implemented in WigglePy, this section presents a synthetic benchmark using the WIGGLE2SEGY (Sopher, 2018) methodology as the reference framework (Fig. 2). The tests used two distinct input image representations: Variable Area (VA), in which only the positive lobes are filled (panel 2), and Variable Area + Wiggle (VAW), in which the full waveform is visible as a line (panel 3).

It is important to emphasize that the choice between VA and VAW is determined by the original data representation, not by the conversion algorithm. The benchmark results (Fig. 2, panels 4-8) highlight the trade-offs between reconstruction fidelity and computational efficiency across the different methodologies.

The Morphological ((Savitzky-Golay + FFT Cutoff, SG)) approach (panel 4) shows a reconstruction error (RMS) of 0.112 for VA inputs and 0.071 for VAW. While it provides extremely fast processing (0.2-0.3 ms), the higher error on VA data confirms that polynomial smoothing alone cannot fully reconstruct the frequency content of missing unipolar lobes without specific wavelet or spectral inference.

The Wiener Deconvolution method (WD) (panel 6) provides improved performance for VA inputs with an RMS error of 0.085. By using a known wavelet signature (30.0 Hz Ricker), it restores the wiggles' structure more effectively than simple morphological smoothing at a matching speed of 0.1 ms. However, it shows a higher sensitivity on VAW inputs (RMS: 0.104) due to spectral competition between the filled area and the tracing line.

The Fast Fourier Bandpass (FFT) (panel 5) emerges as the most robust reconstruction strategy among the direct filters. It achieves RMS errors of 0.073 (VA) and 0.037 (VAW), values that are nearly identical to those obtained via the formal matrix-based inversions of Sopher (2018): SOPHER AMPINV2 (panel 7) yielded 0.071 (VA) and 0.038 (VAW), while SOPHER GRADINV2 (panel 8) yielded 0.066 (VA) and 0.041 (VAW).

Crucially, the benchmark reveals a significant disparity in processing time. While the inversion-based methods (panels 7-8) require 102.9-113.3 ms per trace due to the computational overhead of solving large linear systems, all WigglePy direct methods complete the extraction in approximately 0.1 ms per trace. This three-order-of-magnitude increase in efficiency is the fundamental enabler of WigglePy's interactive, real-time nature within the GIS environment.

The results demonstrate that while both FFT-based and inversion-based methods must rely on mathematical inference to reconstruct unrepresented signal portions from VA images, the lighter spectral-conditioning strategy of WigglePy achieves comparable fidelity with reliable noise management under controlled synthetic conditions.

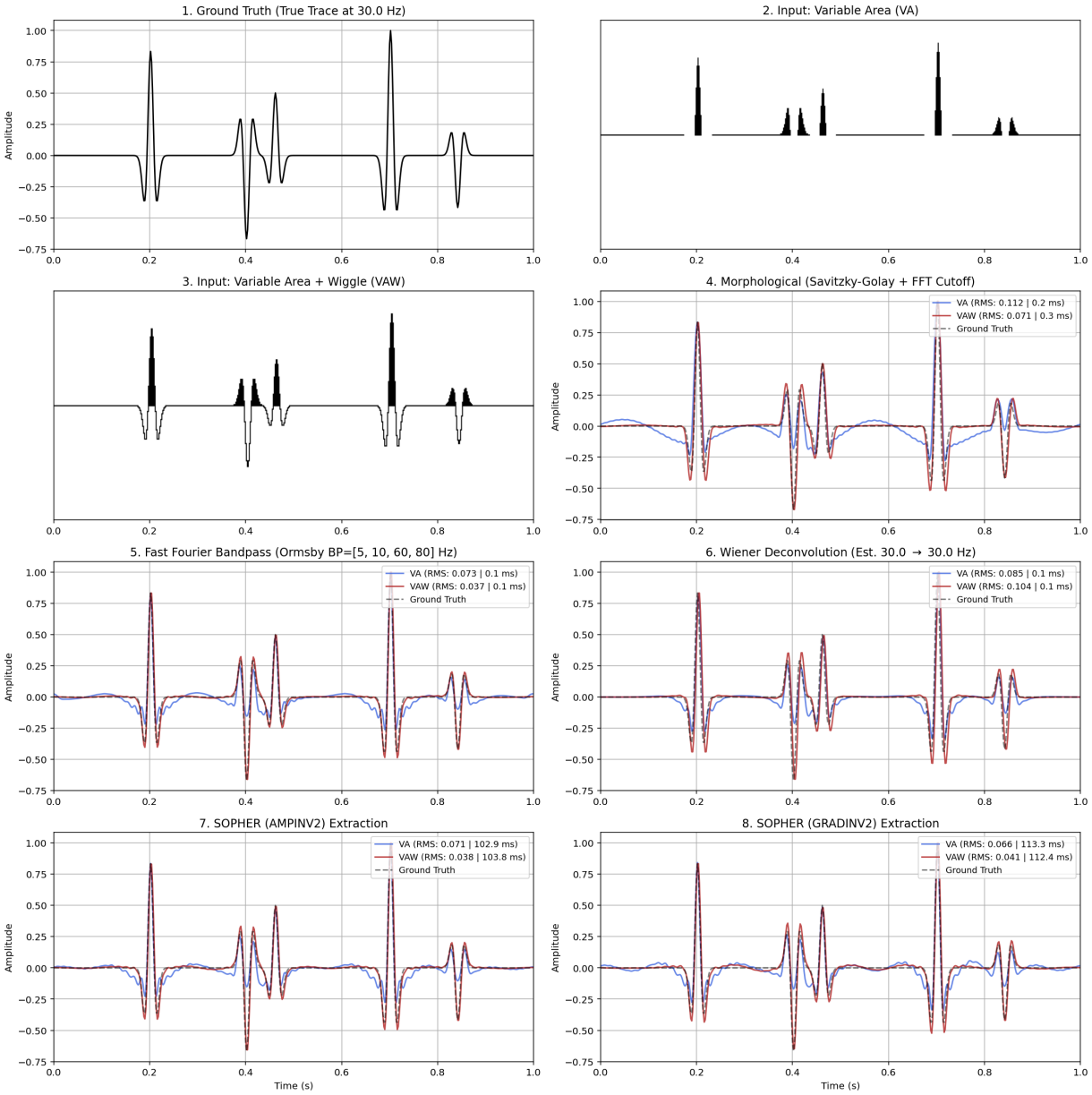

Figure 2. Synthetic benchmark and signal reconstruction limits. (1) Ground Truth seismic trace generated at 30.0 Hz. (2) Input image in Variable Area (VA) format. (3) Input image in Variable Area + Wiggle (VAW) format. (4-8) Comparison of reconstruction results using different algorithms: Savitzky-Golay (Morphological), Fast Fourier Bandpass (Ormsby FFT), Wiener Deconvolution, and Sopher-style inversion (AMPINV and GRADINV). RMS error values and processing time are reported for both VA and VAW inputs for each methodology.

To further quantify the fidelity of WigglePy on real-world datasets, a robust comparative analysis was conducted between the WigglePy outputs and the reference WIGGLE2SEGY results using a regional profile from the Adriatic Sea (profile B-403; Fig. 3) already presented as a test case for WIGGLE2SEGY applications on the publicly available data for the Italian territory (Buttinelli et al., 2022). The input raster reports only the positive variable area (VA) component of the wiggle. The analysis used three key metrics:

- Zero-Lag Cross-Correlation (CC0): measures the sample-by-sample linear relationship between traces after global RMS normalization (Robinson and Treitel, 2000). CC0 values consistently above 0.85 indicate high waveform fidelity.
- Mean Spectral Coherence: assesses the degree of linear association in the frequency domain using Welch's method (Welch, 1967). WigglePy methods show coherence values near 1.000 in the core passband (10-40 Hz), confirming that the spectral content is effectively preserved.

– Envelope Correlation: assesses the wiggles structure preservation by extracting instantaneous amplitude envelopes via Hilbert transform (Taner et al., 1979).

The three reconstruction workflows (FFT, Savitzky-Golay, and Wiener deconvolution) were compared against the WIGGLE2SEGY reference section (Fig. 3). All methods preserve the main reflector geometry and maintain high trace-by-trace agreement, with zero-lag cross-correlation consistently above the 0.85 acceptance threshold. Spectral coherence is near-unity in the core passband for all cases (core = 0.999), while transition-band coherence differentiates the methods (FFT: 0.945; SG: 0.991; WD: 0.981). Envelope-correlation statistics indicate that FFT achieves the highest similarity (mean 0.993, median 0.993), followed by WD (mean/median 0.983/0.984) and SG (mean/median 0.965/0.967). A key practical result is computational efficiency: compared with the reference generation time (46.26 s), all tested methods run in under 1 s (0.77-0.98 s), indicating substantial speed-up with minimal structural loss. The high performance of the FFT approach in envelope similarity confirms its robustness in recovering the overall wavelet energy. At the same time, purely morphological methods (paired with an FFT bandpass) like SG show superior stability in the transition frequency range.

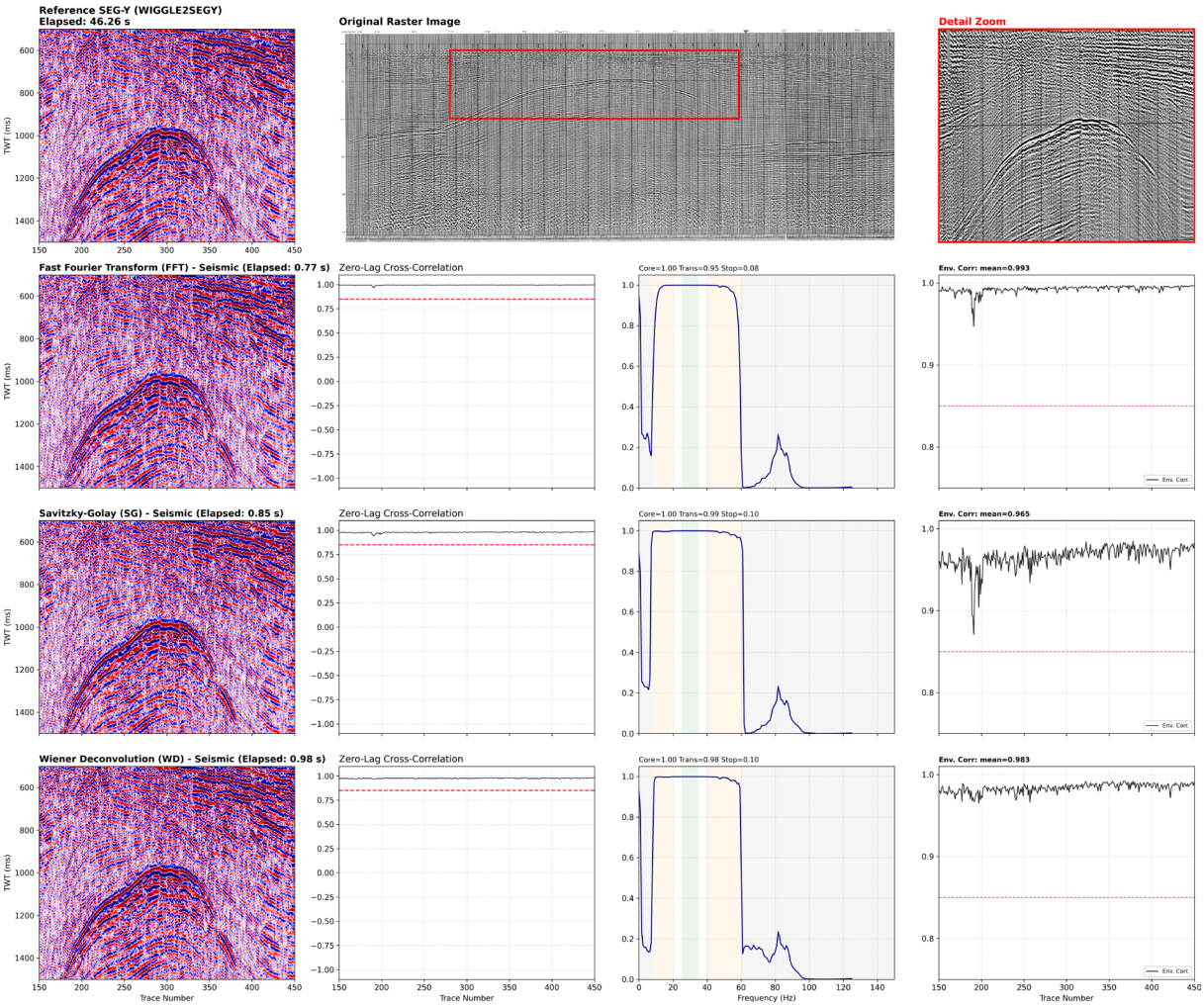

Figure 3. Quantitative performance evaluation of WigglePy methodologies against the WIGGLE2SEGY reference applied to the B-403 regional seismic line (Northern Adriatic Sea). The top panels show the original raster and the WIGGLE2SEGY reference. The three lower rows compare the WigglePy direct methods (FFT, Savitzky-Golay, Wiener Deconvolution) with reference. For each method, the columns display (from left to right): (i) the reconstructed seismic profile with its computation time; (ii) Zero-Lag Cross-Correlation (CC0) across all traces; (iii) Mean Spectral Coherence plot with passband statistics; and (iv) Trace-wise Envelope Correlation plot.

4. Discussion

WigglePy should be understood as an alternative operational workflow that prioritizes speed and interactivity within a GIS environment. WIGGLE2SEGY (Sopher, 2018) remains the scientific reference for formal band-limited inversion, but WigglePy shows that comparable fidelity can be reached with more direct spectral methods (Fig. 3). WigglePy does not natively include the inversion strategy implemented by Sopher (2018) to avoid possible interface lag in the real-time preview and to keep its originality. The multi-stage approach to signal cleaning represents a key operational improvement. By combining initial morphological timeline suppression with a final post-extraction spectral spike filter (Spectral Spike Filter), WigglePy provides robust protection against the inherent limitations of legacy datasets. Residual timeline fragments, which often survive purely automated filters (clearly visible as spectral spikes in the comparison in Fig. 4), can be suppressed in the final preview stage, ensuring a clean digital record ready for advanced interpretation.

This discussion will focus on evaluating the performance of WigglePy, including the Spectral Spike Filter functionality, in a real-world challenging scenario. With this aim, a composite seismic section from the Po Plain (Northern Italy), available through the ViDEPI project (Linea App_Orient_01), was processed. This regional profile crosses the epicentral area of the 2012 Emilia earthquake sequence (Govoni et al., 2014), providing critical imaging of the blind thrust systems involved in the event. This specific line was also recently incorporated into a multidisciplinary study on seismic source inversion (Improta et al., 2023). The raster seismic profile is particularly challenging for WigglePy because the original image is of relatively low quality and exhibits significant distortion along the trace.

This comparison confirms the significant increase in operational speed provided by WigglePy. Numerically, the extraction is a direct spectral operation that avoids the high computational overhead of solving the large systems of linear equations associated with band-limited inversion. While inversion requires the definition and processing of large matrices for each trace – a process that scales significantly with signal length and resolution – direct spectral methods operate with much lower computational complexity, providing near-instantaneous trace recovery even on standard hardware.

As illustrated in Fig. 4, the qualitative comparison of the Po Plain composite section demonstrates that all extraction methods successfully recover the primary seismic reflectors. However, distinct differences emerge in the spectral domain and noise handling. The WIGGLE2SEGY reconstruction (Fig. 3, panel 1) exhibits characteristic spectral spikes, associated with residual timeline fragments that the automated morphological filters failed to suppress fully. The spectral-spikes are visible also in the SEG-Y extracted with WigglePy but are less pronounced both in terms of relative amplitude in the spectral domain and in the visual inspection of the sampled data (Fig. 4). After the post-extraction spectral spike filtering, the FFT-based extraction provides a smoother spectral response within the user-defined corner frequencies, thereby reducing the impact of these high-energy artefacts. The Savitzky-Golay approach prioritizes continuity and noise suppression, albeit at the cost of a slight reduction in high-frequency sharpness and some anomalous recovery of low-frequency signals near the lower cutoff. In contrast, Wiener deconvolution yields a broader spectrum and sharp reflector definition but introduces a higher level of incoherent background noise. These results confirm that while inversion-based methods offer high theoretical fidelity, WigglePy's multi-stage conditioning provides a flexible environment for testing alternative solutions when dealing with highly degraded images.

Operationally, the integration within QGIS eliminates the need for data transfer between disparate software environments (e.g., MATLAB and GIS). The primary advantage is the convenience of managing geometry reconstruction and SEG-Y export within the same spatial environment. This ensures that the recovered digital data is natively consistent with the line traces, base maps, and CDP positions already managed in the GIS project. Among the utilities implemented in WigglePy, a dedicated tool for automatically generating CDPs from linear tracks significantly streamlines the workflow. This interactive utility allows simultaneous visualization of line geometry and CDP numbering, enabling the user to account for both regular internal intervals and common irregularities, such as non-standard starting and ending CDP values. Moreover, the optional utilities included in WigglePy allow for merging multiple SEG-Y files after the conversion, allowing the possibility to recreate the original length of the data within a single environment. The availability of open-source tools integrated in a GIS environment is a valuable asset for the ongoing initiative within the EPOS (European Plate Observing System) aimed at facilitating the discovery of legacy seismic reflection data (Lorenz et al., 2026).

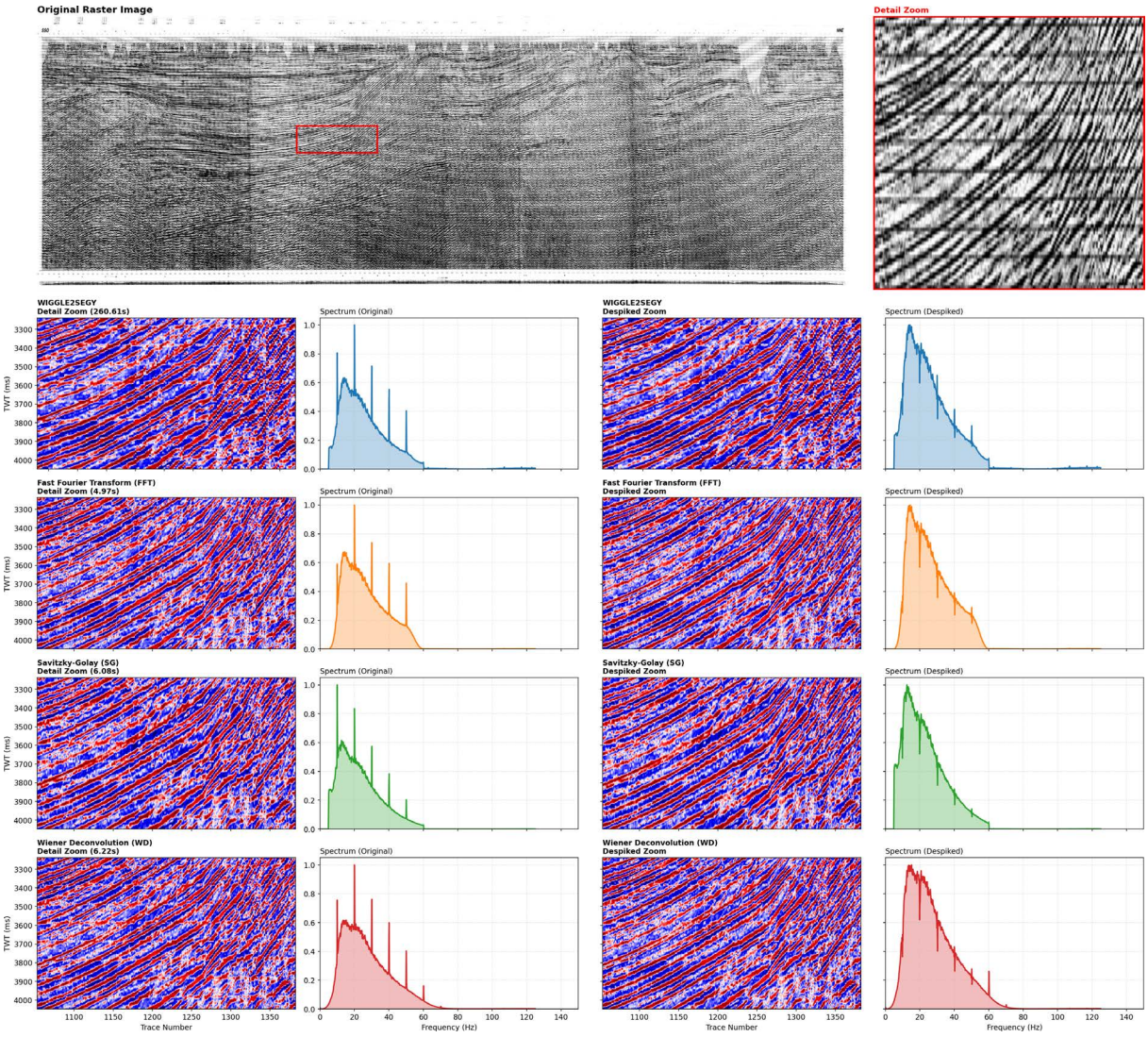

Figure 4. Algorithm comparison on a legacy seismic profile. The dataset consists of a regional composite seismic section from the Po Plain (Northern Italy, VIDEPI archive, Line App_Orient_01). The top panel displays the original raster image and a close-up view of the signal representation. Subsequent panels compare the digital reconstruction results of WIGGLE2SEGY (1) with the three primary WigglePy extraction methods: FFT (2), Savitzky-Golay (3), and Wiener Deconvolution (4). All the panels show a sample of the original SEG-Y and its frequency content (for all the traces) and the results after the application of the post-extraction spectral spike filter highlighting the noise attenuation of the spectral spikes related to residual timelines.

5. Conclusions

WigglePy provides a scientifically credible and operational alternative for converting legacy seismic images into digital SEG-Y data. WigglePy is intended as an academic product to facilitate and accelerate researchers' work.

By integrating image cleaning, extraction, and geometry assignment into QGIS, it streamlines workflows for researchers and professionals. The FFT-based extraction yields reconstruction errors comparable to those of established inversion methods while offering greater operational speed and a robust multi-stage noise-attenuation pipeline. In the context of Italian legacy archives (ViDEPI), WigglePy represents an effective solution for revitalizing historical data for modern seismotectonic and georesources investigations.

Data availability statement. WigglePy code and documentation can be downloaded at: https://github.com/framae80/WigglePy.

The original raster images used for testing can be downloaded at:

- B-403_02 (Fig. 3): https://www.videpi.com/deposito/sismica/ZB/Linea_B-403_02.pdf
- App_Orien_1 (Fig. 4): https://www.videpi.com/deposito/videpi/sismicatitoli/Linea_App_Orient_01.pdf

Acknowledgements. The author wishes to thank the Editor of Annals of Geophysics and two anonymous reviewers for the comments which significantly improve the original manuscript. This work was possible thanks to the infrastructure of INGV's Reflection Seismology Laboratory "SismoLab-3D" (https://sismolab3d.ingv.it/). The author thanks Giuseppe Vico, Roberto Basili, Mauro Buttinelli and Roberta Maffucci for the testing phase and for providing useful feedback during the writing phase. The author acknowledges the use of generative artificial intelligence (AI) during the coding and writing phases of this study. The author used Antigravity (Gemini 3.1 Pro, Claude Sonnet 4.6, and Claude Opus 4.6) and Codex within Visual Studio Code (GPT 5.3-Codex and GPT 5.4). All AI-generated outputs, including code and text, were critically reviewed and edited by the author, who maintains full accountability for the integrity and accuracy of the work. WigglePy can be downloaded at (https://github.com/framae80/WigglePy) under the GNU General Public License v3. It is provided as experimental software. The author makes no representations or warranties regarding the accuracy or completeness of the results produced by this plugin. When used in scientific publications, please refer to this paper. Project funded by the Ministry of University and Research under the PRIN 2022 Call – RePop – Project Code 20225MLCRS – CUP D53C24004280006- (Resp. UR. INGV Francesco E. Maesano).

References

Amadori, C., G. Toscani, A. Di Giulio, F. E. Maesano et al. (2019). From cylindrical to non-cylindrical foreland basin: Pliocene-Pleistocene evolution of the Po Plain-Northern Adriatic basin (Italy). Basin Res. 31, 991-1015. doi:10.1111/bre.12369.

Antoncecchi, I., F. Ciccone, G. Dialuce, S. Grandi et al. (2020). Progetto SPOT – Sismicità Potenzialmente Innescabile Offshore e Tsunami: Report integrato di fine progetto. 1. Ministero dello Sviluppo Economico, Roma, Italia. doi:10.5281/zenodo.3732887.

Araque-Pérez, C. J., T. Teixidó, F. De Lis Mancilla and J. Morales (2024). Reprocessing and interpretation of legacy seismic data using machine learning from the Granada Basin, Spain, Tectonophysics, 885, 230414, doi:10.1016/j.tecto.2024.230414.

Bogiatzis, P. and M. Ishii (2016). DigitSeis: A New Digitization Software for Analog Seismograms, Seism. Res. Lett., 87, 726-736, doi:10.1785/0220150246.

Buttinelli, M., L. Petracchini, F. E. Maesano, C. D'Ambrogi et al. (2021). The impact of structural complexity, fault segmentation, and reactivation on seismotectonics: Constraints from the upper crust of the 2016-2017 Central Italy seismic sequence area, Tectonophysics, 810, 228861, doi:10.1016/j.tecto.2021.228861.

Buttinelli, M., F. E. Maesano, D. Sopher, F. Feriozzi et al. (2022). Revitalizing vintage seismic reflection profiles by converting into SEG-Y format: case studies from publicly available data on the Italian territory, Ann. Geophys. 65, DM538, doi:10.4401/ag-8883.

Chen, C. H. and J. B. H. Shyu (2026). Seismogenic Structures Offshore Taiwan and Their Earthquake Potential from the Taiwan Earthquake Model (TEM) Project, Bull. Seismol. Soc. Am. 116, 26-51, doi:10.1785/0120250009.

Cicala, M., F. De Giosa, A. Piscitelli, G. Scicchitano and V. Festa (2024). Conversion of vintage seismic reflection profiles of the ViDEPI dataset crossing the Gondola Line seismogenic fault (offshore Apulia, Adriatic Sea, Southern Italy) to SEG-Y, Data in Brief, 55, 110705, doi:10.1016/j.dib.2024.110705.

Conti, A., R. Maffucci and S. Bigi (2022). The use of public vintage seismic reflection profiles, in: Interpreting Subsurface Seismic Data, Elsevier, 127-156, doi:10.1016/B978-0-12-818562-9.00003-0.

Diviacco, P., N. Wardell, E. Forlin, C. Sauli et al. (2015). Data rescue to extend the value of vintage seismic data: The OGS-SNAP experience, GeoResJ, Rescuing Legacy Data for Future Science, 6, 44-52, doi:10.1016/j.grj.2015.01.006.

Ercoli, M., F. Carboni, A. Akimbekova, R. B. Carbonell and M. R. Barchi (2023). Evidencing subtle faults in deep seismic reflection profiles: Data pre-conditioning and seismic attribute analysis of the legacy CROP-04 profile, Front. Earth Sci. 11.

Farran, M. (2012). IMAGE2SEGY: Converts Rasters to SEG-Y Format. http://gma.icm.csic.es/ca/image2segy.

Feriozzi, F., L. Improta, F. E. Maesano, P. De Gori and R. Basili (2024). The 3D crustal structure in the epicentral region of the 1980, Mw 6.9, Southern Apennines earthquake (southern Italy): New constraints from the integration of seismic exploration data, deep wells, and local earthquake tomography, Tectonics, 43, e2023TC008056.

Govoni, A., A. Marchetti, P. De Gori, M. Di Bona et al. (2014). The 2012 Emilia seismic sequence (Northern Italy): Imaging the thrust fault system by accurate aftershock location. Tectonophysics, 622, 44-55, doi:10.1016/j.tecto.2014.02.013.

Hao, Y., K. McIntosh and M. B. Magnani (2015). Long-lived deformation in the southern Mississippi Embayment revealed by high-resolution seismic reflection and sub-bottom profiler data, Tectonics 34, 555-570, doi:10.1002/2014TC003750.

Improta, L., A. Cirella, G. Pezzo, I. Molinari and A. Piatanesi (2023). Joint Inversion of Geodetic and Strong Motion Data for the 2012, Mw 6.1-6.0, May 20th and May 29th, Northern Italy Earthquakes: Source Models and Seismotectonic Interpretation, J. Geophys. Res.: Solid Earth, 128, e2022JB026278, doi:10.1029/2022JB026278.

Lemenkova, P., R. De Plaen, T. Lecocq and O. Debeir (2022). Computer Vision Algorithms of DigitSeis for Building a Vectorised Dataset of Historical Seismograms from the Archive of Royal Observatory of Belgium, Sensors 23, 56, doi:10.3390/s23010056.

Liu, Y., B. Dang, Y. Li, H. Lin, and H. Ma (2016). Applications of Savitzky-Golay Filter for Seismic Random Noise Reduction, Acta Geophys., 64, 101-124, doi:10.1515/acgeo-2015-0062.

Lorenz, H. and EPOS CSS community (2026). EPOS CSS – Facilitating the discovery and reuse of Controlled Source Seismic data. doi:10.5194/egusphere-egu26-5358.

Maesano, F. E. and C. D'Ambrogi (2016). Coupling sedimentation and tectonic control: Pleistocene evolution of the central Po Basin, Ital. J. Geosci., 135, 394-407, doi:10.3301/IJG.2015.17.

Maesano, F. E., M. M. Tiberti and R. Basili (2017). The Calabrian Arc: three-dimensional modelling of the subduction interface, Sci. Rep., 7, 8887, doi:10.1038/s41598-017-09074-8.

Maesano, F. E., M. Buttinelli, R. Maffucci, G. Toscani et al. (2023). Buried Alive: Imaging the 9 November 2022, Mw 5.5 Earthquake Source on the Offshore Adriatic Blind Thrust Front of the Northern Apennines (Italy), Geophys. Res. Lett., 50, e2022GL102299, doi:10.1029/2022GL102299.

Maesano, F. E., C. Zuffetti, A. Abbate, C. D'Ambrogi and R. Bersezio (2024). Quaternary slip-rates probabilistic estimation for the Northern Apennines frontal thrust in the Po Plain (Northern Italy) by integrating surface and subsurface data, Tectonophysics, 873, 230227.

Miles, P. R., M. Schaming, and R. Lovera (2007). Resurrecting vintage paper seismic records, Mar. Geophys. Res., 28, 319-329, doi:10.1007/s11001-007-9034-5.

Niedźwiecki, M. J., M. Ciołek, A. Gańcza and P. Kaczmarek (2021). Application of regularized Savitzky-Golay filters to identification of time-varying systems, Automatica 133, 109865, doi:10.1016/j.automatica.2021.109865.

Ormsby, J. F. A. (1961). Design of Numerical Filters with Applications to Missile Data Processing, J. Ass. Compu. Machin., 8, 440-466, doi:10.1145/321075.321087.

Owen, M. J., M. A. Maslin, S. J. Day and D. Long (2015). Testing the reliability of paper seismic record to SEGY conversion on the surface and shallow sub-surface geology of the Barra Fan (NE Atlantic Ocean), Marine Petrol. Geol., 61, 69-81, doi:10.1016/j.marpetgeo.2014.12.009.

Panara, Y., F. E. Maesano, C. Amadori, J. Fedorik et al. (2021). Probabilistic Assessment of Slip Rates and Their Variability Over Time of Offshore Buried Thrusts: A Case Study in the Northern Adriatic Sea. Front. Earth Sci. 9.

Park, J. O., T. Tsuru, S. Kodaira, P. R. Cummins and Y. Kaneda (2002). Splay Fault Branching Along the Nankai Subduction Zone, Science, 297, 1157-1160, doi:10.1126/science.1074111.

Pertuz, A., M. I. Benito, P. Llanes Estrada, P. Suarez-Gonzalez and M. García-Martín (2025). SEGYRecover: A Python Tool for Digitizing Vintage Seismic Sections into SEG-Y Format, doi:10.5281/ZENODO.15053412.

Pertuz, A., M. I. Benito, P. Llanes, P. Suarez-Gonzalez and M. García-Martín (2026). Remastering legacy seismic sections: A user-friendly toolkit for vectorization, velocity modeling, and processing, Appl. Comput. Geosci., 30, 100344, doi:10.1016/j.acags.2026.100344.

Ricker, N. (1953). The form and laws of propagation of seismic wavelets, Geophysics, 18, 10-40, doi:10.1190/1.1437843.

Robinson, E. A. and S. Treitel (2000). Geophysical Signal Analysis, Society of Exploration Geophysicists, doi:10.1190/1.9781560802327.

Savitzky, A. and M. J. E. Golay (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627-1639. doi:10.1021/ac60214a047.

Shaw, J. H. and P. M. Shearer (1999). An Elusive Blind-Thrust Fault Beneath Metropolitan Los Angeles, Science, 283, 1516-1518, doi:10.1126/science.283.5407.1516.

Sopher, D. (2018). Converting scanned images of seismic reflection data into SEG-Y format, Earth Sci. Inform. 11, 241-255, doi:10.1007/s12145-017-0329-z.

Taner, M. T., F. Koehler and R. E. Sheriff (1979). Complex seismic trace analysis, Geophysics, 44, 1041-1063, doi:10.1190/1.1440994.

Treitel, S. (1969). Predictive deconvolution-theory and practice, Geophysics, 34, 155-169, doi:10.1190/1.1440003.

Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoustics, 15, 70-73, doi:10.1109/TAU.1967.1161901.

***CORRESPONDING AUTHOR: Francesco E. MAESANO**,
Istituto Nazionale di Geofisica e Vulcanologia
e-mail: francesco.maesano@ingv.it