Physics-Informed Digital Twin and Reconstruction Pipeline for Under-Lead Ultrasonics in Lead-Cooled SMRs

Marko Budimir
INETEC Ltd.
Dolenica 28, Donji Stupnik, Croatia
marko.budimir@inetec.hr

Dejan Strbad
The Lisbon Council
PC-Résidence Palace, 155 Rue de la Loi, 1040 Brussels, Belgium
dejan.strbad@lisboncouncil.net

Mihaela Orić, Filip Novoselnik
Protostar Labs Ltd.
Strossmayerova 341, Osijek, Croatia
mihaela.oric@protostar.ai, filip@protostar.ai

Ante Bilušić, Fabjan Jozić, Josip Žunić, Rita Rudelić, Toma Šiklić
University of Split, Faculty of Science
Ruđera Boškovića 33, Split, Croatia
ante.bilusic@pmfst.hr, fjozic@pmfst.hr, jzunic@pmfst.hr, rrudelic@pmfst.hr, tsiklic@pmfst.hr

ABSTRACT

Reliable operation of lead-cooled small modular reactors (SMRs) depends on in-service diagnostics capable of operating in optically opaque heavy liquid metal. Under-lead ultrasonics is one of the few credible modalities for this task, but its performance is constrained by long-path propagation in molten lead, elevated temperature, wetting-dependent coupling at steel/lead interfaces, and the sensitivity of coherent imaging to phase and timing distortions. These effects influence bandwidth, amplitude, and the reliability of reconstructed images.

This paper presents a framework that connects inspection physics, synthetic data generation, and reproducible reconstruction into a traceable development chain. The contribution is structured around two linked assets: a parameterised digital twin and scenario library aligned to lead-cooled SMR constraints, and a versioned DSP/reconstruction pipeline built around regression-testable software artefacts. The reference concept is informed by experimentally credible molten-lead conditions reported in the literature, including temperatures around 380 °C, controlled atmosphere for oxygen management, and wetting stabilisation via thin Ni coatings on steel interfaces. A five-case matrix spans baseline propagation, attenuation and degradation reference cases, simple localisation tasks, geometric image-sensitivity studies, and reconstruction-comparison scenarios.

The modelling framework combines acoustic simulation, physically motivated degradation operators, and flow-aware screening based on the Pridmore-Brown formulation to bound the regime in which mean-flow effects may be treated perturbatively. The pipeline standardises A-scan ingestion, metadata handling, time-of-flight extraction, and image reconstruction using delay-and-sum and total focusing method (TFM) processing. Initial benchmarking of learned restoration models shows that signal-level denoising and image-level reconstruction quality are not equivalent objectives. In the current test set, WaveNet achieved the strongest overall TFM performance, while a larger 1D U-Net provided stronger point-wise A-scan denoising without translating this into superior reconstructed image quality. The framework is intended to support future experimental anchoring, scalable HPC-based dataset generation, and more defensible operation and maintenance arguments for lead-cooled SMRs.

Keywords: lead-cooled SMR, under-lead ultrasonics, wetting, Pridmore-Brown, k-Wave, A-scan, DSP, TFM, reproducibility

1 INTRODUCTION

Lead-cooled small modular reactors (SMRs) create a difficult inspection environment: critical components are immersed in optically opaque heavy liquid metal, access is restricted, and diagnostic information must be extracted from signals propagating through molten lead under elevated temperature and uncertain interface conditions. In this setting, under-lead ultrasonics is one of the few credible modalities for structural awareness, localisation, and future defect-sensitive inspection [1].

Recent experiments have shown that ultrasonic imaging in molten lead is feasible at about 380 °C when wetting is stabilised and signal transfer across the steel/lead interface is maintained [1]. Feasibility, however, is not yet the same as a traceable basis for trustworthy imaging software or defensible inspection arguments.

The difficulty is that under-lead ultrasonic data are shaped by reflector geometry, by impedance mismatch, propagation loss, beam spreading, wetting-dependent coupling, possible coolant-motion effects, numerical settings, and the reconstruction method itself. Flow-related effects are considered here only as a bounded screening question, consistent with the Pridmore-Brown formulation later used in the paper [2]. Neither raw A-scans nor reconstructed images should therefore be treated as direct physical ground truth.

This paper addresses that problem by coupling a physics-informed digital twin with a reproducible reconstruction pipeline. The simulation layer is grounded in established k-space and wave-propagation methodology [3–8], while the reconstruction layer is anchored in full matrix capture and total focusing method (TFM) practice [9]. The purpose is to provide a traceable framework in which propagation, degradation, restoration, and image formation can be studied under declared assumptions.

Two observations motivate the approach. First, numerical discipline matters: discretisation, time step, Courant–Friedrichs–Lewy (CFL) choice, and excitation definition affect both runtime and practical observables [3–6]. Second, reconstruction-level evaluation is essential because waveform improvement alone does not guarantee inspection-relevant image quality. A model may improve a waveform locally while degrading the phase and timing relationships needed for coherent focusing.

Accordingly, the paper is built around two linked assets: a parameterised digital twin with a staged scenario library, and a versioned digital signal processing (DSP) and reconstruction pipeline for reproducible data handling, degradation, restoration, and image formation.

2 PHYSICAL BASIS AND INSPECTION CONSTRAINTS

Under-lead ultrasonics in lead-cooled SMRs is governed by coupled acoustic, interfacial, thermal, and geometrical constraints distinct from conventional contact ultrasonic testing. Because the inspection medium is molten lead at elevated temperature, the measured response depends on reflector geometry, on material properties, interface transmission, propagation loss, beam spreading, and preservation of phase information required for coherent reconstruction [1], [7], [10].

In the present framework, molten lead is represented by sound speed $c_{\mathrm{Pb}} \approx 1.7 \times 10^3\,\mathrm{m/s}$ and density $\rho_{\mathrm{Pb}} \approx 1.05 \times 10^4\,\mathrm{kg/m^3}$, while steel is represented by longitudinal-wave speed $c_{\mathrm{steel}} \approx 5.9 \times 10^3\,\mathrm{m/s}$. The corresponding acoustic impedance,

$$Z = \rho c, \qquad (1)$$

governs reflection and transmission at the steel/lead boundary. Even with moderate bulk attenuation, impedance mismatch, interface condition, propagation distance, and bandwidth narrowing can significantly reduce the information available for delay-and-sum and TFM reconstruction.

Wetting is a central practical issue. Poor wetting reduces transmission efficiency, alters spectral content, and introduces variability that propagates directly into time-of-flight estimation and reconstructed image quality. Reported molten-lead experiments used thin Ni coatings to stabilise wetting at about 380 °C [1]. In the present model, this uncertainty is represented by an effective interface transmission factor T_{int}, used as a compact engineering proxy rather than a full interface model.

Propagation loss is introduced through

$$P(x) = P_0 e^{-\alpha x}, \qquad (2)$$

and, in k-wave-style modelling,

$$\alpha(f) = \alpha_0 f^y. \qquad (3)$$

In the present context, attenuation should be understood as a combined practical loss budget in which intrinsic absorption in molten lead is only one contributor, alongside beam spreading, diffraction, and wetting-dependent interface transmission losses. For Newtonian-fluid-like behaviour, $y \approx 2$ is a useful first approximation [3]–[5]. In practical under-lead inspection, however, attenuation should not be interpreted in isolation, since beam divergence, diffraction, wetting-dependent transmission, and interface losses may be of comparable importance [1], [4], [10]. While wetting governs signal transfer at the interface, a separate question concerns whether coolant motion along the propagation path introduces phase or timing distortions significant enough to affect reconstruction.

Possible coolant-motion effects are treated through a flow-aware screening layer based on the Pridmore-Brown formulation [2]. For the pressure perturbation amplitude $\tilde{p}(y)$,

$$\frac{d^2\tilde{p}}{dy^2} + \left(\frac{2k_x}{\Omega}\frac{du}{dy} - \frac{1}{\rho_0}\frac{d\rho_0}{dy}\right)\frac{d\tilde{p}}{dy} + \left[\left(\frac{\Omega}{c} + i\alpha\right)^2 - k_x^2\right]\tilde{p} = 0, \qquad (4)$$

with

$$\Omega(y) = \omega - k_x u(y), \qquad (5)$$

where $u(y)$ is the mean flow velocity, $\rho_0(y)$ the background density, and k_x the axial wavenumber. The corresponding Mach number is

$$M = \frac{u}{c}, \qquad (6)$$

and for representative lead-cooled conditions,

$$M \approx 5.8 \times 10^{-4}. \qquad (7)$$

This value is far below unity, so mean-flow effects are unlikely to dominate propagation but remain useful as a bounded screening layer. The corresponding representative velocity-gradient profiles are shown in Figure 1.

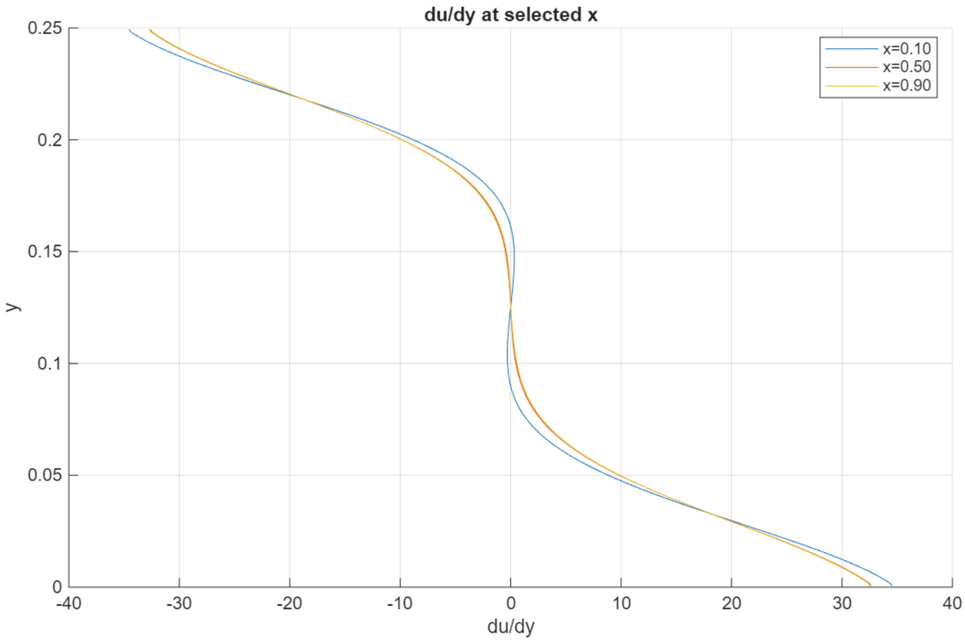


Figure 1. Representative velocity-gradient profiles du/dy across the molten-lead domain for selected axial positions. Although the corresponding Mach number is very small, the result supports the use of a bounded flow-aware screening layer.

From the imaging standpoint, the key requirement is preservation of amplitude consistency, phase stability, temporal fidelity, and usable bandwidth before reconstruction. Degradation is therefore introduced at the full matrix capture (FMC) signal level before restoration and TFM image formation. Under-lead ultrasonics should accordingly be treated as a constrained inverse problem whose outputs are only as credible as the physical assumptions embedded in propagation, degradation, and reconstruction.

3 REFERENCE INSPECTION CONCEPT AND SCENARIO LIBRARY

The reference inspection concept is defined as a controlled, simulation-first testbed for under-lead ultrasonics in lead-cooled SMRs. Its purpose is to establish a physically anchored and numerically reproducible environment in which propagation, degradation, localisation, and reconstruction can be evaluated under explicitly declared assumptions.

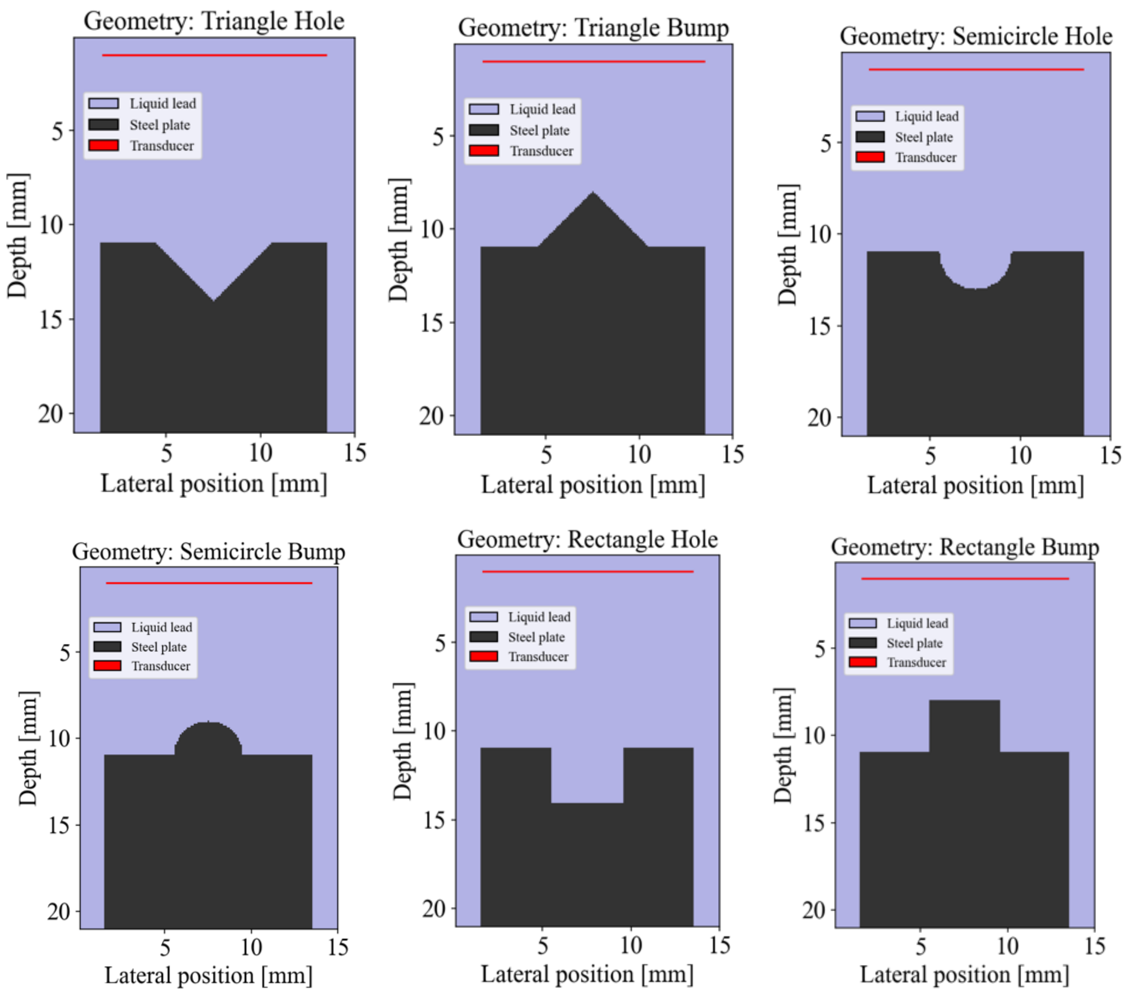
The concept is intentionally based on an experimentally credible under-lead inspection motif: a compact pitch-catch or FMC-capable arrangement operating over a steel structure immersed in molten lead, with practical constraints informed by reported molten-lead imaging conditions such as elevated temperature, controlled atmosphere, and wetting stabilisation by thin Ni coatings [1]. This provides a realistic starting envelope without introducing unnecessary geometric complexity. Such a staged approach is consistent with the broader inspection-development logic required for advanced lead-cooled reactor systems [10]–[12].

Table 1. Staged T1–T5 scenario library used to structure the under-lead ultrasonic digital twin and reconstruction benchmarking workflow.

Scenario	**Primary purpose**	**Representative target / condition**	**Main evaluation output**
T1 – Baseline propagation and coupling	Establish a clean reference case and verify basic waveform credibility	No target, planar reflector, or simple steel backwall under nominal interface conditions	Reference A-scan, spectral content, timing stability, baseline amplitude, reflection-coefficient checks

T2 – Attenuation and degradation reference	Quantify controlled signal corruption before reconstruction	Same simple geometry as T1, but with progressively degraded FMC data	Comparison of clean vs. degraded signals, calibration of degradation severity, bandwidth reduction, SNR loss
T3 – Single-target localisation	Test detectability and localisation of one simple reflector under controlled conditions	Single recess, protrusion, flat reflector, or isolated canonical defect surrogate	Target-position verification, interpretable B-scan/TFM response, simple localisation benchmark
T4 – Geometric sensitivity and image formation	Study the influence of target shape, position, and degradation severity on reconstructed image behaviour	Canonical triangular, semicircular, and rectangular holes or bumps	Comparative B-scan/TFM image set showing sensitivity to geometry, artefacts, and reconstruction robustness
T5 – Reconstruction comparison and pipeline performance	Compare alternative restoration and reconstruction paths on the same underlying case	Clean-reference, degraded-input, and learned-restoration-assisted reconstruction routes	TFM comparison, SSIM / PSNR / SNR-type metrics, evidence of current capability and model limitations

The geometry is intentionally canonical. A steel plate or backwall is interrogated through molten lead, while target surrogates are represented by simple triangular, semicircular, and rectangular holes or protrusions. These shapes are intended to provide controlled local scattering behaviour with unambiguous ground truth and clearly distinguishable B-scan and TFM signatures. This makes them appropriate for localisation tests, sensitivity studies, and reconstruction benchmarking. The corresponding canonical geometry/response set is shown in Figure 2.



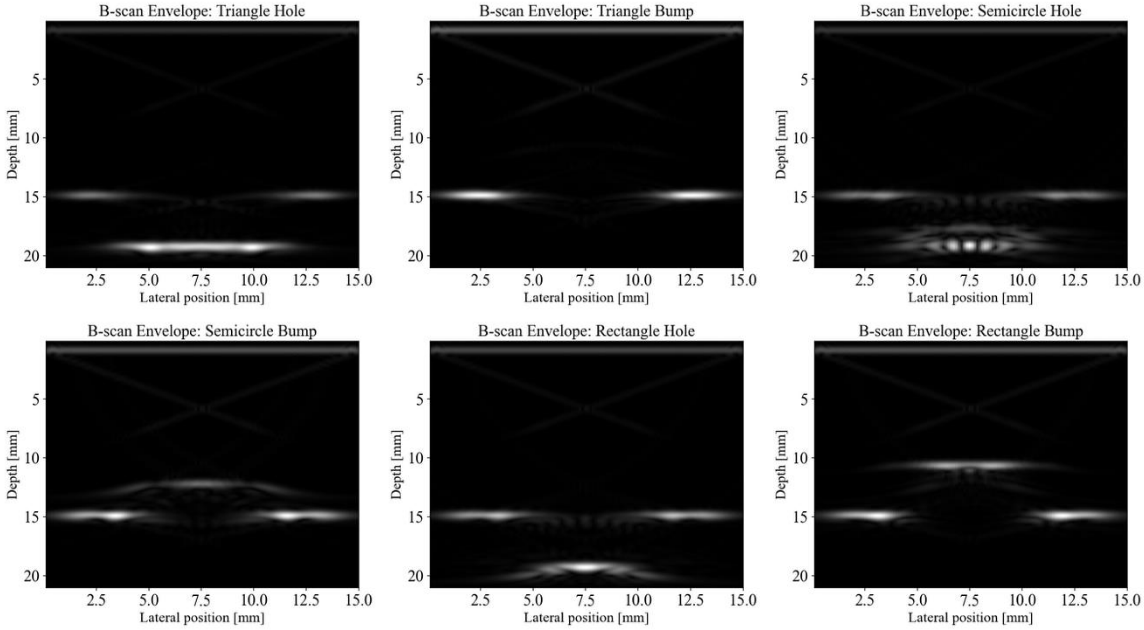


Figure 2. Canonical reflector geometries and corresponding simulated B-scan responses used in the staged scenario library. Rows correspond to geometry type, while columns distinguish hole- and bump-type perturbations and their associated B-scan responses. Even simple local perturbations generate distinct image signatures suitable for localisation and reconstruction benchmarking.

Time-of-flight remains the primary measurement observable. For a pulse-echo reflector at depth d,

$$t_{PE} = \frac{2d}{c}, \qquad (8)$$

and for a pitch-catch path of total length L,

$$t_{PC} = \frac{L}{c}. \qquad (9)$$

These idealised relations do not capture all practical effects, but they define the baseline timing structure against which localisation and reconstruction behaviour are interpreted.

The scenario library is organised as a staged T1–T5 matrix, summarised in Table 1. Its role is to impose methodological discipline: complexity is added only after the simpler level has established physical and numerical credibility. T1 and T2 provide baseline propagation and degradation references; T3 introduces simple localisation; T4 expands the geometry space through canonical targets; and T5 evaluates the complete restoration-and-reconstruction chain. In this sense, the scenario matrix is a planning aid, and the experimental design layer of the paper.

This staged structure serves two purposes. First, it prevents premature emphasis on visually attractive reconstructions before the underlying assumptions have been tested. Second, it provides a scalable template for future campaigns. The current local studies already show that the framework can generate synthetic FMC data, apply degradation, and benchmark restoration models; larger HPC-enabled studies would therefore extend the same scenario logic rather than replace it.

At its present maturity level, the reference inspection concept should be viewed as an intermediate layer between first-principles inspection physics and future experimentally anchored algorithm development. That is precisely the level needed for under-lead ultrasonics to progress from feasibility demonstration toward a credible inspection methodology.

4 DIGITAL TWIN FORMULATION AND IMPLEMENTATION

The digital twin developed in this work is not intended as a live replica of an operating plant, but as a parameterised inspection twin: a computational environment in which under-lead ultrasonic propagation, degradation, restoration, and reconstruction can be studied under controlled assumptions. Its role is to make the full development chain traceable rather than merely executable.

At the highest level, the twin is represented as

$$\mathcal{T}: \{\Theta_{phys}, \Theta_{num}, \Theta_{geom}, \Theta_{deg}\} \rightarrow \{S_{clean}, S_{deg}, I_{rec}, M\}, \qquad (10)$$

where the parameter groups define, respectively, the physical model, numerical configuration, inspection geometry, and degradation model. The outputs are the clean and degraded signal sets, the reconstructed image, and the associated metric set. The point of (10) is simple: the image is the endpoint of a declared chain.

The physical layer instantiates molten lead, steel, the inspection aperture, and the canonical target set introduced in Section 3. Material assumptions are consistent with published lead-property data and experimentally credible molten-lead imaging conditions [1], [7], [10], while the forward simulation is grounded in established k-space and wave-propagation methodology [3]–[5], [8]. In compact form,

$$S_{clean} = \mathcal{F}_{fwd}\big(\Theta_{phys}, \Theta_{geom}, \Theta_{num}\big), \qquad (11)$$

where $\mathcal{F}_{fwd}$ denotes the realised forward operator. This notation is intentionally economical: every A-scan entering the pipeline must be attributable to a known physical and numerical state.

A key design choice is that numerical settings are treated as part of the model rather than as hidden implementation detail. The supporting studies showed strong sensitivity of runtime and practical observables to spatial step, CFL choice, time step, and excitation definition, while reflection-coefficient estimates did not improve monotonically with finer discretisation alone. In particular, multi-cycle excitation approached the theoretical steel/lead reflection coefficient more credibly than a single-cycle case under several tested settings. For that reason, Δx, CFL, Δt, toneburst length, and PML settings are retained as explicit metadata rather than buried inside simulation defaults [3]–[6].

The twin also separates clean propagation from realistic inspection data. This distinction matters because attenuation in molten lead is only one contributor to signal loss; beam divergence, diffraction, interface transmission, and wetting state may be equally important in practice [1], [4], [10]. Accordingly, degradation is introduced as a broader operator,

$$S_{deg} = \mathcal{D}\big(S_{clean}; \Theta_{deg}\big). \qquad (12)$$

Here, Θ_{deg} parameterises the principal physics-motivated distortions applied at FMC level, including attenuation, bandwidth narrowing, additive noise, wetting-dependent amplitude loss, and reverberation-like effects.

Architecturally, the twin is layered but not verbose: a physics layer defines medium, geometry, and aperture; a signal layer converts simulated pressure histories into channel-organised A-scans; and a workflow layer applies degradation, restoration, reconstruction, and metric evaluation. This separation is mainly there to prevent false progress. If image quality changes, it should be possible to determine whether the cause lies in the propagation model, the degradation model, the restoration stage, or the metric layer, rather than in some undeclared change of signal formatting or timing convention.

The twin is reconstruction-aware from the outset. It is not a forward model whose outputs are generated specifically for coherent reconstruction and comparison. This is expressed as

$$I_{rec} = \mathcal{R}(S), \qquad (13)$$

where S may denote clean, degraded, or restored signals and $\mathcal{R}$ includes delay-and-sum and TFM logic [9]. The importance of (13) is methodological: restoration is only meaningful insofar as it improves the information available to reconstruction.

At the present operating point, the framework is already functional end to end, but its limits should be read correctly. The current studies used a small FMC dataset, truncated A-scan length, and desktop-class GPU resources. That is sufficient to demonstrate execution, degradation, restoration, and TFM benchmarking, but not to claim that the inverse problem has been solved. It is, however, sufficient to show that the framework exposes real trade-offs: some models improve point-wise denoising without improving reconstruction, and some numerical settings increase cost without proportionate gain in credibility.

For that reason, scalability is built in rather than appended later. The architecture is already prepared for larger synthetic campaigns and HPC-enabled training, but the conceptual structure would remain unchanged. What scales is the statistical depth with which the same inspection logic can be exercised.

Taken together, the digital twin serves three functions in the paper: it provides a controllable forward model, a degradation-aware benchmarking environment, and a scalable route toward experimentally anchored validation. That is its real contribution. In under-lead ultrasonics, credibility depends less on producing one attractive image than on preserving traceability across the entire chain that produced it.

The simulation and training pipeline is implemented in Python using PyTorch and executed on NVIDIA GPU hardware, with support for both local workstation runs and HPC deployment via PBS job scheduling on the Supek cluster of the University of Zagreb Computing Centre.

5 RECONSTRUCTION PIPELINE AND DATA HANDLING

The reconstruction pipeline is designed as a metadata-driven workflow that converts simulated or measured ultrasonic signals into reproducible image-level outputs. Its purpose is to ensure that every reconstruction can be traced to declared signal assumptions, timing conventions, and processing settings. This is essential in under-lead ultrasonics, where small undeclared changes in velocity model, sampling, cropping, or gate placement can produce visually persuasive but physically misleading differences.

Accordingly, waveform data are never treated as detached arrays. Each signal set is linked to explicit metadata, including sampling information, channel indexing, timing reference, assumed acoustic velocity, gate definition, reconstruction settings, and scenario identifier. This allows reconstruction to remain tied to a physical interpretation rather than to ad hoc post-processing.
At workflow level, the processing chain may be written as

$$I_{rec} = \mathcal{R}(\mathcal{E}(\mathcal{P}(\mathcal{V}(S)))), \qquad (14)$$

where S denotes the input signal set, $\mathcal{V}$ metadata and integrity validation, $\mathcal{P}$ optional preprocessing, $\mathcal{E}$ timing-feature extraction, and $\mathcal{R}$ the reconstruction operator.

Validation is intentionally strict. Before reconstruction, the pipeline checks metadata completeness, signal dimensional consistency, timing-axis validity, and compatibility between waveform data and declared processing parameters. This is less a software convenience than a safeguard against false gains caused by undeclared format or calibration changes.

Preprocessing is deliberately lightweight. Depending on the case, it may include detrending, band-limited filtering, or normalisation, but only insofar as these operations stabilise downstream analysis without changing the physical meaning of the signal. Timing extraction then identifies peaks, gated responses, or backwall-like indications and converts them into quantities interpretable under the declared velocity model. For simple reflectors, this remains anchored to the time-of-flight logic introduced in Section 3.

Reconstruction itself is based on transparent baseline methods. Delay-and-sum provides the simplest reference, while TFM is the main coherent imaging route because it tests whether the signal processing chain has preserved enough phase and timing structure for meaningful focusing [9]. This is why restoration is evaluated later at image level as well as signal level: a cleaner waveform is not necessarily a more useful one.

The pipeline is also reproducible by construction. Each run produces images and a machine-checkable processing state including software version, scenario label, preprocessing choice, gate definition, velocity assumption, and reconstruction parameters. In that sense, the pipeline is the mechanism by which the twin becomes benchmarkable.

6 BENCHMARKING STRATEGY, RESULTS, AND PERFORMANCE METRICS

The benchmark is designed to test whether restoration remains useful at both waveform and image level. This distinction is essential in under-lead ultrasonics: a model may improve local agreement with a clean A-scan while degrading the phase and timing structure required for coherent focusing. For that reason, all compared models are evaluated on the same degraded signals and reconstructed through the same TFM workflow, so that differences in outcome can be attributed to restoration rather than downstream inconsistency.

The metric set combines signal-domain and image-domain criteria,

$$M = \{M_{signal}, M_{image}\} = \{\mathrm{SNR}, \mathrm{SSIM}_{1D}, \mathrm{PSNR}_{TFM}, \mathrm{SSIM}_{2D}\}, \qquad (15)$$

where the signal metrics quantify local waveform fidelity and the image metrics quantify inspection-relevant reconstruction quality. In the present framework, the latter carry greater weight, because reconstruction is the actual endpoint of interest.

Four restoration paths were compared against the degraded baseline: a 1D U-Net, WaveNet, a CNN-Transformer, and a 2D U-Net. Their architectural differences matter less here than the question they expose: whether better denoising at A-scan level produces better reconstructed images. The four models span different inductive biases: the 1D U-Net operates on individual A-scans through an encoder-decoder with skip connections, WaveNet uses stacked dilated convolutions to capture multi-scale temporal context, the CNN-Transformer combines convolutional encoding with self-attention, and the 2D U-Net processes full B-scans jointly.

A representative benchmark case corresponding to Sample 4 is shown in Figure 3. This sample contains three reflectors at different depths and lateral positions: a side-drilled hole, a rectangular bump, and a semicircular bump. The clean TFM image preserves the expected reflector structure, while the degraded input remains interpretable but weakened. The restored outputs then diverge clearly. The 1D U-Net suppresses noise strongly but also over-smooths deeper regions; the CNN-Transformer preserves structure less convincingly; the 2D U-Net recovers some visible features but remains weaker than the best case. WaveNet provides the most credible balance between denoising and structural retention and therefore the strongest TFM reconstruction in the representative comparison.

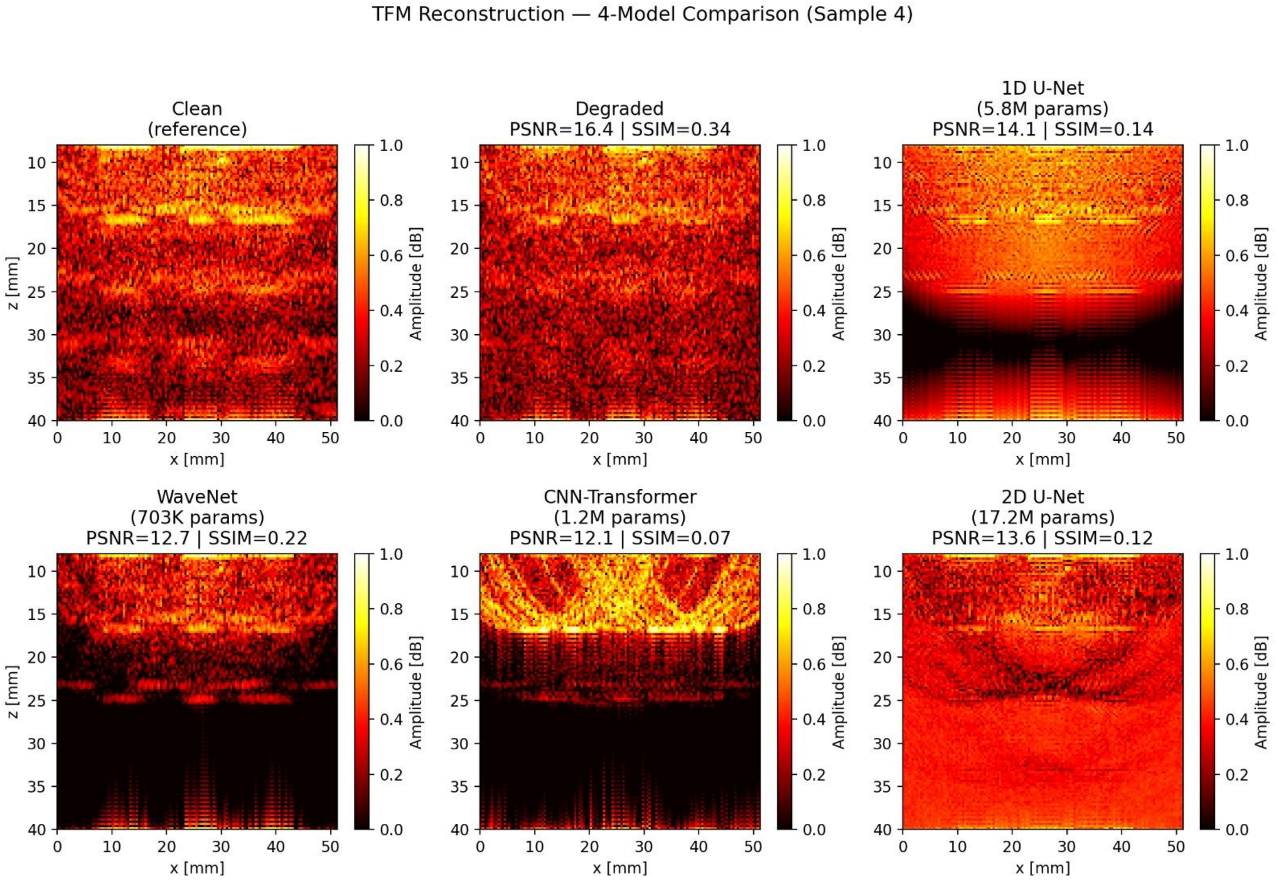


Figure 3. Representative TFM reconstruction comparison for a test sample containing three reflectors at varying depths and lateral positions: a side-drilled hole, a rectangular bump, and a semicircular bump. Stronger waveform denoising does not automatically produce superior image-level reconstruction.

The same conclusion is visible in Figure 4. At A-scan level, the compared models differ in residual noise and in how strongly they alter temporal fine structure and spectral content. Stronger smoothing is therefore not automatically desirable. A waveform may look cleaner and still become less useful for coherent reconstruction.

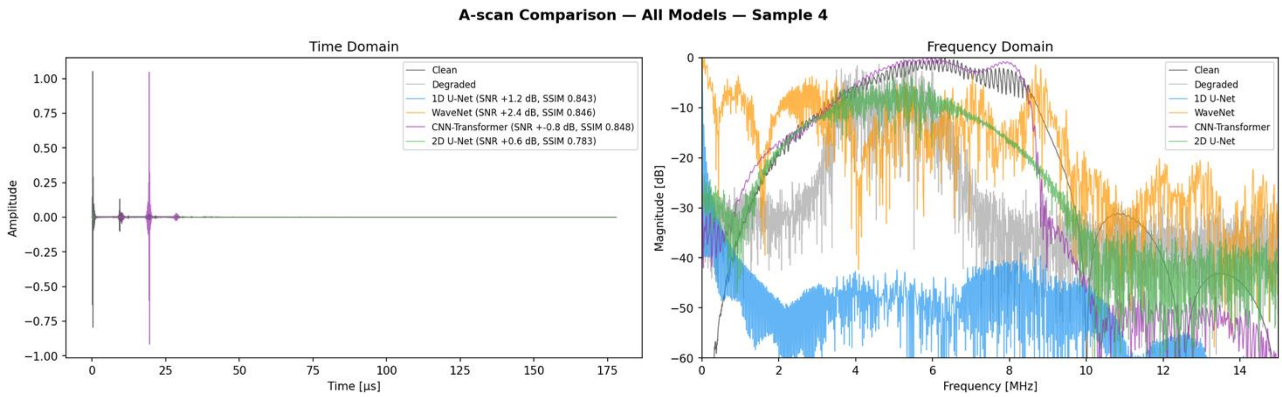


Figure 4. Representative A-scan comparison in time and frequency domain. The figure highlights the trade-off between aggressive smoothing and preservation of temporal structure relevant for coherent focusing.

The averaged benchmark over 23 test samples reinforces the same conclusion. WaveNet achieved the strongest overall image-domain performance, with TFM PSNR of 16.06 dB and TFM 2D SSIM of 0.35, despite not maximising signal-domain SNR. By contrast, the 1D U-Net remained stronger in point-wise denoising, while CNN-Transformer and 2D U-Net lagged in image-domain reconstruction quality. The central methodological result therefore remains unchanged: stronger point-wise denoising does not imply stronger coherent reconstruction.

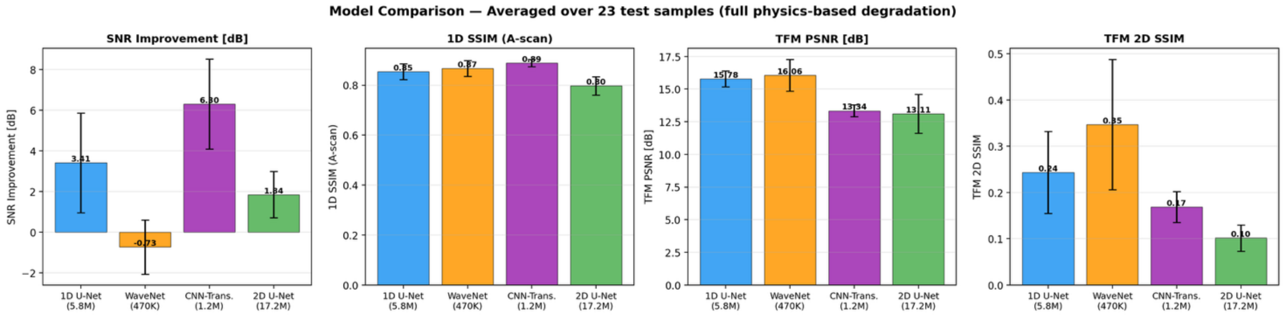


Figure 5. Averaged benchmark metrics across the 23-sample test subset. The best point-wise denoising model is not necessarily the best reconstruction model.

The broader significance of the benchmark is therefore that the framework already distinguishes between models that look promising in the signal domain and models that remain useful after coherent imaging. That is a more important result than a score alone. It shows that under-lead restoration should not be treated as generic denoising, and that model selection must be driven primarily by downstream reconstruction behaviour rather than by waveform appearance.

The benchmark should still be read with the correct scope. The dataset remains limited, and the present results do not establish universal superiority under all degradation regimes. They do, however, establish that the digital twin and reconstruction pipeline are already mature enough to expose real model trade-offs, real limitations, and a credible path for larger synthetic and experimentally anchored studies.

7 DISCUSSION

The main value of the present work is that it closes a traceable loop between inspection physics, numerical modelling, signal handling, and reconstruction benchmarking. In under-lead ultrasonics, this is essential because the inspection outcome is shaped by multiple coupled effects, including acoustic impedance mismatch, wetting-dependent transmission, attenuation, beam spreading, discretisation sensitivity, degradation, and the reconstruction method itself. A framework that leaves these effects implicit may still produce plausible images, but it cannot support technically defensible conclusions.

The benchmark confirms that under-lead restoration cannot be judged reliably from waveform quality alone. What matters is whether the recovered signal still supports coherent reconstruction and interpretable TFM images.

This has a direct implication for inspection development. If restoration models are judged only by A-scan metrics, there is a real risk of optimising toward signals that look cleaner but are less useful for reconstruction. The present results therefore support a stricter criterion: restoration should be assessed primarily by its downstream effect on TFM image quality, localisation, and interpretability.

A second important outcome concerns numerical discipline. The digital twin results show that discretisation, excitation length, and time stepping cannot be treated as neutral implementation details. Reflection-coefficient behaviour, runtime, and practical waveform fidelity depend on them in non-trivial ways. Trust in the twin must therefore be earned through controlled studies of the relevant observables [3]–[6].

The work also supports a more balanced interpretation of attenuation and coolant-motion effects. Under the representative lead-cooled conditions considered here, mean-flow effects are unlikely to dominate propagation, but retaining a bounded flow-aware screening layer remains scientifically useful because it defines when such effects may be neglected perturbatively [2]. Likewise, practical signal loss should be interpreted as a combined effect of attenuation, beam divergence, interface transmission, and wetting state [1], [10].

More broadly, the framework already separates three levels that are often blurred together: physical plausibility, numerical credibility, and reconstruction usefulness. A scenario may be physically plausible but numerically unreliable, or numerically stable yet still poor for reconstruction benchmarking. The staged T1-T5 scenario logic and the metadata-driven pipeline are valuable because they keep these distinctions explicit.

The present study should therefore be read as a foundation paper. Its strongest contribution is a disciplined environment in which under-lead ultrasonic inspection can be developed without confusing plausibility for proof. For a problem of this class, that is the necessary basis for credible experimental validation and later industrialisation.

8 LIMITATIONS AND FUTURE WORK

The present study is a simulation-first development paper, not an experimental qualification campaign. Its main limitation is that the benchmark is still dominated by synthetic data, so its value depends on the assumptions embedded in the forward model, degradation layer, and reconstruction chain.

A second limitation is the present treatment of coupling and wetting. The effective interface transmission factor is a practical engineering proxy, but it remains a reduced-order description of a more complex interface state. Future work should therefore refine the representation of amplitude loss, spectral distortion, phase effects, and repeatability as functions of wetting condition [1].

The current dataset is also limited in scale. The benchmark already distinguishes between signal-level denoising and image-level usefulness, but it is not yet sufficient for definitive ranking of restoration architectures. More diverse synthetic campaigns, full-length A-scan handling, and HPC-enabled training are therefore more important than incremental model tuning.

The physics layer is intentionally simplified as well. Mean-flow effects are treated as a bounded screening layer, while attenuation and degradation are represented through compact but physically motivated operators rather than full multi-physics resolution. This is appropriate at the present stage, but should be revisited selectively as the experimental basis grows.

Finally, the reflector library remains intentionally canonical. This is methodologically correct for early benchmarking, but it means that the paper does not yet support component-specific qualification claims. The immediate next steps are therefore clear: experimental anchoring of simple scenarios, larger synthetic and training campaigns, refinement of wetting and degradation models, and broader benchmarking against localisation, repeatability, and robustness criteria.

These limitations do not weaken the present paper; they define its maturity level. The main outcome is the establishment of a traceable intermediate framework on which larger validation and later industrialisation can credibly build.

9 CONCLUSION

This paper has presented a physics-informed framework for under-lead ultrasonic inspection development in lead-cooled SMRs, built around two linked assets: a parameterised digital twin and a reproducible reconstruction pipeline. The purpose is not to claim solved inspection, but to provide a traceable environment in which propagation, degradation, restoration, and image formation can be studied under declared assumptions.

The main methodological result is clear: signal-level denoising and image-level reconstruction quality are not equivalent objectives. In the present benchmark, WaveNet provided the strongest overall TFM performance, while a larger 1D U-Net achieved stronger point-wise A-scan denoising without translating that advantage into superior reconstructed images. For coherent under-lead imaging, the decisive quantity is therefore not waveform smoothness alone, but preservation of the phase and timing structure required for focusing.

The second result is architectural. The digital twin proved useful as a propagation model, and as a controlled benchmarking environment in which physical assumptions, numerical settings, degradation modes, and restoration strategies can be varied without losing traceability. Together with the metadata-driven pipeline, this provides a practical basis for reproducible comparison and later experimental anchoring.

The present work should be read at the correct maturity level. It is not a qualification study, nor a final statement on AI-assisted under-lead reconstruction. Its contribution is more foundational and, for that reason, more durable: it establishes the development logic required for credible progress in a harsh, optically opaque, and physically coupled inspection environment.

For under-lead ultrasonics in lead-cooled SMRs, that logic matters as much as the models themselves. Better inspection performance will not come from denoising alone, but from tighter control of the full chain by which physically meaningful images are produced.

ACKNOWLEDGEMENT

The authors acknowledge partial financial support for this work through the project contract NPOO.C3.2.R3-I1.06.0302, funded under the call "Proof of Innovative Concept – Third Call" (reference number C3.2.R3-I1.06) within the Recovery and Resilience Mechanism, with the support of the Ministry of Science, Education and Youth and the Croatian Agency for SMEs, Innovation and Investments (HAMAG-BICRO).

REFERENCES

[1] J. Hawes, J. Knapp, R. Burrows, R. Montague, P. Wilcox, H.-T. Chien, J. Arndt, S. Walters, "Demonstration of an ultrasonic imaging system for molten lead," *Nuclear Engineering and Technology*, Vol. 56, No. 4, pp. 1460–1471, 2024.

[2] D.C. Pridmore-Brown, "Sound propagation in a fluid flowing through an attenuating duct," *Journal of Fluid Mechanics*, Vol. 4, No. 4, pp. 393–406, 1958.

[3] B.E. Treeby, B.T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of Biomedical Optics*, Vol. 15, No. 2, 021314, 2010.

[4] B.E. Treeby, J. Jaros, A.P. Rendell, B.T. Cox, "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method," *The Journal of the Acoustical Society of America*, Vol. 131, No. 6, pp. 4324–4336, 2012.

[5] M. Tabei, T.D. Mast, R.C. Waag, "A k-space method for coupled first-order acoustic propagation equations," *The Journal of the Acoustical Society of America*, Vol. 111, No. 1, pp. 53–63, 2002.

[6] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, Vol. 114, No. 2, pp. 185–200, 1994.

[7] V. Sobolev, "Database of Thermophysical Properties of Liquid Metal Coolants for GEN-IV," Scientific Report SCK-CEN-BLG-1069, SCK CEN, 2011.

[8] T.D. Mast, L.P. Souriau, D.-L. Liu, M. Tabei, A.I. Nachman, R.C. Waag, "A k-space method for large-scale models of wave propagation in tissue," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 48, No. 2, pp. 341–354, 2001.

[9] C. Holmes, B.W. Drinkwater, P.D. Wilcox, "Post-processing of the full matrix of ultrasonic transmit–receive array data for non-destructive evaluation," *NDT & E International*, Vol. 38, No. 8, pp. 701–711, 2005.

[10] OECD Nuclear Energy Agency, *Handbook on Lead-bismuth Eutectic Alloy and Lead Properties, Materials Compatibility, Thermal-Hydraulics and Technologies*, OECD/NEA, Paris, 2007.

[11] Nuclear Innovation and Research Office (NIRO), *Advanced Modular Reactors: Technical Assessment*, Department for Business, Energy and Industrial Strategy, July 2021.

[12] National Nuclear Laboratory (NNL), *SMR Techno-Economic Assessment Project 3: SMRs Emerging Technology Literature Review of Emerging SMR Technologies*, Report for the Department of Energy and Climate Change, 15 March 2016.