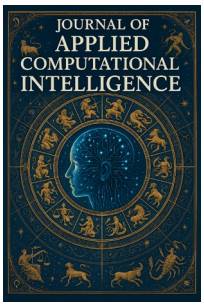
UNIVERSITY OF WEST ATTICA

eISSN: 3136-6937

Journal of Applied Computational Intelligence

Vol. 4, No. 2, April, 2026



Dielectric Prior Guided U-Net for Two-Dimensional Reconstruction of Railway Ballast Fouling from Ground Penetrating Radar

Marek Ciesielski[1,*], Robert Stępień[1]

[1] *Faculty of Mechatronics and Automation, Koszalin University of Technology, Koszalin, 75-453, Poland*

Received 23 December 2025; Revised 10 April 2026; Accepted 21 April 2026; Published 26 April 2026

Abstract

Spatially varying ballast fouling affects railway drainage and complicates ground-penetrating radar interpretation because moisture and fines produce overlapping electromagnetic responses. This study develops a dielectric prior-guided U-Net for reconstructing a two-dimensional section of continuous volumetric fouling fractions. Separate radar and dielectric encoders exchange information through confidence-gated feature fusion, while training combines supervised reconstruction, interface preservation and moisture-conditioned dielectric consistency. The reported evaluation comprises 1,200 partitioned ballast scenarios and 24,000 radar windows, with partitioning performed before window extraction. Mean absolute error decreases from 0.086 for a conventional U-Net to 0.048 for the proposed configuration, corresponding to a relative reduction of 44.2%. Structural similarity increases from 0.781 to 0.893, and intersection-over-union increases from 0.614 to 0.782 at a diagnostic fouling threshold of 0.25. Ablation comparisons associate the complete configuration with lower error than configurations omitting gating or dielectric consistency. Under a relative prior bias of 20%, error rises from 0.048 to 0.072, showing that reliability-weighted conditioning remains sensitive to calibration. These comparisons describe performance within the stated evaluation conditions; their interpretation depends on scenario provenance, reference support and independent calibration records. The method offers a framework for incorporating uncertain dielectric information into spatial ballast assessment. Transfer to operational inspection requires field validation across moisture states, antenna configurations and material domains.

Keywords: Ground Penetrating Radar, Ballast Fouling Reconstruction, Dielectric Prior Fusion, Confidence Gating, Spatial Inverse Learning

1. Introduction

Ballast fouling reduces drainage capacity and, in the presence of moisture, can accelerate settlement and deterioration of track geometry [1]. The composition of fouling material and its degree of saturation influence the strength and deformation of fouled railway ballast [2]. GPR-based ballast assessment relates changes in dielectric permittivity to fouling and the electromagnetic response of the trackbed [3]. GPR-based learning methods have distinguished moisture and mud-pumping defects in ballast beds using complementary time- and frequency-domain features [4]. These dependencies link the engineering requirement for spatial condition evaluation with the material uncertainty in sensing. Conventional sampling provides direct information at selected locations, while sampling spacing and depth determine the spatial continuity of the resulting condition description. Ground-penetrating radar provides closely spaced measurements along the survey line and is used to study local variations among the material sampling locations.

* Corresponding author: stepien.ro@tu.koszalin.pl

http://dx.doi.org/10.68406/jaci.2026.vol4iss2no4:44-58

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

Inversion of these observations to obtain a fouling distribution needs to consider antenna coupling, geometric interfaces, propagation losses and material contrasts. Therefore, a space-resolved reconstruction needs a specified target support and a justifiable link between waveform characteristics and material condition. Prior knowledge is particularly relevant at greater depths, where attenuation reduces the information retained in the measured signal. The reconstruction system needs to show its dependencies in the inputs, calibration assumptions and evaluation methods.

The main uncertainty concerns the combined effect of fines and water on the effective electromagnetic properties of ballast. Dielectric mixing descriptions relate the effective material properties to the constituent fractions and phase arrangement in granular railway materials. The dependence is conditional; fines composition, pore water and packing state may alter the radar response at the same fouling fraction. Different material states may produce similar results in the observation of a limited antenna band. Independently controlled combinations of water state and contamination are needed for the experimental separation of moisture and fouling effects. Direct interpretation of a radar image as a fouling map can obscure this ambiguity. Amplitude normalization may eliminate beneficial information on attenuation; excessive filtering can also reduce interfaces separating fine-rich layers. Selecting radar attributes reduces dimensionality but also limits the set of material patterns for the inverse model. The goal of reconstruction is spatial inference based on acquisition geometry, complementary material data and feasible electromagnetic behavior. The uncertainty of the conditioning information influences reconstruction alongside its nominal value. Confidence weighting is required for a dielectric estimation based on a single moisture assumption in the presence of local water state uncertainty. Therefore, a design is needed to incorporate prior knowledge in the reconstruction process while preserving the contribution of the radar observation.

Encoder-decoder networks are suitable for the purpose of combining local features and extended spatial context in this task. Ge et al. [5] combined convolutional features with self-attention in U-Net-based GPR inversion to represent long-range relationships within radar observations. Inverse reconstruction also needs constraints on the non-uniqueness of the observation-to-material mapping. A training-specific association between amplitude and fouling may become unreliable when moisture or antenna coupling changes. Recent reviews identify dataset dependence and limited generalization as continuing challenges for deep-learning-based GPR inversion [6]. A dielectric prior can reduce the set of feasible material states; reliability-aware fusion regulates the influence of conditioning according to externally supplied uncertainty. Similarly, a consistency loss can incentivize agreement with an incorrect prior unless its effect is regulated. The research gap addressed in this paper is the combined treatment of spatial reconstruction, conditional dielectric information and prior reliability. The aim is to achieve a continuous volumetric fouling ratio in the designated area. The operational interpretation of this fraction is determined by the inspection circumstances and relevant maintenance requirements. Additional evidence of drainage, loading and the relevant inspection criteria are needed for maintenance interpretation. This distinction enables the assessment of the reconstruction problem via continuous error and interface fidelity prior to setting an operating threshold for the output.

A dielectric estimate is suitable for fouling reconstruction if its material assumptions, moisture dependence and calibration uncertainty can be propagated through the learning process with the radar observations. Physics-constrained neural representations can be optimized through GPR full-waveform inversion to reconstruct subsurface electromagnetic parameters [7]. The current form employs a moisture-conditioned dielectric relation as a soft prior and explicitly expresses its uncertainty. This constraint connects the predicted material fraction to a calibrated dielectric relation in its specified range of applicability. Bayesian GPR waveform inversion represents subsurface parameters probabilistically and quantifies uncertainty in their estimation [8]. The contribution is the joint use of independently encoded dielectric mean and spread, externally scaled fusion reliability, and conditional material consistency within a specimen-partitioned reconstruction protocol. The protocol connects the reliability of dielectric conditioning with reconstruction quality in various material states and acquisition conditions. Section 2 reviews radar interpretation, learned reconstruction and dielectric conditioning. Section 3 introduces the reconstruction target and the prior representation. Section 4 presents feature fusion, training losses and inference. Design of experiments, model comparisons, ablations and prior sensitivity analysis are shown in Section 5. Section 6 presents the methodological contributions, limitations and necessary experiments for evaluation of the proposed method in railway operating conditions. This organization links the material assumptions, network design and evaluation criteria and maintains a single definition of the inverse task.

2. Related Work

2.1. Radar based assessment of granular track layers

Radar interpretation of railway ballast spans waveform inspection, attribute analysis, interface tracking and spatial classification. Comparisons of railway GPR approaches show that ballast contamination can be assessed through several signal attributes with different interpretive behavior [9]. Each method defines a distinct inference target. Interface-tracking estimates geometry; material classification labels a spatial support that may have a combination of conditions. Continuous fouling estimation needs an additional correlation of the electromagnetic response with the quantity of fines.

This correspondence depends on whether fouling is defined by dry mass, solid volume or occupied pore volume. Conversion between these definitions requires supporting material measurements.

Hybrid GPR and machine-learning workflows have jointly estimated railway trackbed interfaces, ballast permittivity and layer thickness using numerical and field observations [10]. An amplitude descriptor of an individual trace has a different spatial extent than the label for a collected ballast sample. A mismatch between sensing and sampling supports can produce apparent prediction error even when both measurements are internally consistent. It may also increase agreement due to adjacent windows of the same physical specimen in both training and testing. Therefore, in the evaluation of related work, the independence of the specimens and the correspondence of labels to the sensing volume need to be considered. The problem that remains is how to retrieve a spatially-varying material property given that the radar response combines information over a finite area and has an extended sensing volume.

2.2. Learned reconstruction and spatial supervision

Learned reconstruction substitutes a hand-selected series of radar attributes with features optimized for an inverse target. Wu et al. [11] combined multiscale dilated convolution with hybrid attention gates in a two-stage GPR inversion network to reconstruct subsurface permittivity distributions. Skip connections can preserve fine-grained spatial details and deeper layers aggregate context; however, the interpretation of the result is still reliant on the training distribution. A network may learn to associate wet areas with severe fouling from the training data and thus produce the same connection even in the presence of wet clean ballast at deployment. This dependency concerns both supervision design and network architecture. Independent variation of material states in the supervision data is needed to isolate the underlying material effects.

Domain-adaptive generative adversarial networks have been used to address discrepancies between simulated and measured GPR images during frequency enhancement [12]. Changes in wavelet shape, time-zero, coupling and attenuation can alter the input distribution without modifying the target definition. Although data augmentation may introduce certain disturbances to the model, it is only beneficial if these disturbances are representative of the desired operating environment. A foreground defect loss has been introduced in semisupervised GPR inversion to improve reconstruction of subsurface defect boundaries [13]. Pixel-wise error promotes average agreement; interface-aware objectives focus on the location and sharpness of transitions. The balance between spatial regularity and preservation of diffuse fines transitions remains unresolved. A reconstruction method should specify the target support and evaluate both continuous error and boundary behavior.

2.3. Dielectric conditioning and prior reliability

Dielectric conditioning provides the material data which a waveform-only model would need to obtain indirectly. Dielectric mixing models relate effective permittivity to phase composition, constituent properties and assumptions about mixture geometry [14]. These models are dependent on phase fractions and constituent properties; however, for practical accuracy, calibration and arrangement of these phases are required. Transfer across fines mineralogies and moisture ranges requires evaluation of the fitted relationship in each target material domain. Therefore, a prior estimate based on this relationship should include information about its uncertainty, spatial resolution and calibration area. Treating the prior as a perfectly observed image obscures differences in calibration uncertainty and spatial support.

Multimodal fusion combines complementary information across data streams, while heterogeneous noise and missing information can limit its effectiveness [15]. The cited studies motivate two design choices for testing the interaction between material conditioning and prior reliability. First, the radar observation and dielectric estimation should have independent feature paths before fusion to retain the ability to represent discrepancies. Second, confidence should explicitly modulate the influence of the prior during feature fusion. The generalization of these design choices remains an empirical question. A controlled hypothesis is defined to be investigated by ablation and intentionally corrupted priors. The evidence gap concerns whether reliability-aware fusion enhances reconstruction in independent moisture states and acquisition domains while preserving local material boundaries. This question guides the model specification and the evaluation protocol in the following sections.

3. Reconstruction Target and Dielectric Prior

3.1. Spatial support and observation model

The target is a two-dimensional section through the ballast layer along a radar survey line. Each output cell represents the volumetric fraction of fine solids within the combined coarse and fine solid volume, excluding pore water from the denominator. This definition separates changes in solid composition from changes in water content. Conversion between this volumetric fraction and a dry mass fouling index requires the constituent densities. The reconstructed section covers a nominal horizontal length of 4.0 m and a ballast depth of 0.60 m. The output grid contains 128 horizontal cells and 64 depth cells. The nominal cell spacings are 3.125 cm horizontally and 0.9375 cm in depth; these are reporting intervals,

while physical resolution depends on the radar footprint and bandwidth. The input consists of 256 along track traces with 512-time samples per window. Acquisition geometry and a depth support mask accompany each window.

The observation is represented by a forward operator whose inputs include the dielectric field, conductivity, antenna coupling, and interface geometry. Their relationship is expressed through the observation equation:

$$y = \mathcal{F}(\varepsilon(f, m), \sigma, g) + n \tag{1}$$

Here, y denotes the radar observation, f is the fouling fraction, and m represents moisture state. The effective permittivity ε connects material composition to the electromagnetic response. Conductivity is denoted by σ, while g collects acquisition and geometric conditions. The additive term n represents measurement noise and residual processing artefacts. The additive residual also absorbs unmodelled forward effects; separating these effects from measurement noise requires independent acquisition and model-error records.

Figure 1 illustrates the observation and target construction workflow, from ballast constituents and acquisition geometry to a radar window, a dielectric prior, and an independently defined fouling section. The calibration branch supplies the prior independently of the material reference target. The prior estimation procedure uses information available during inference, while the reference fouling distribution is reserved for supervision and evaluation. This separation preserves the intended information boundary between material calibration and reconstruction assessment.

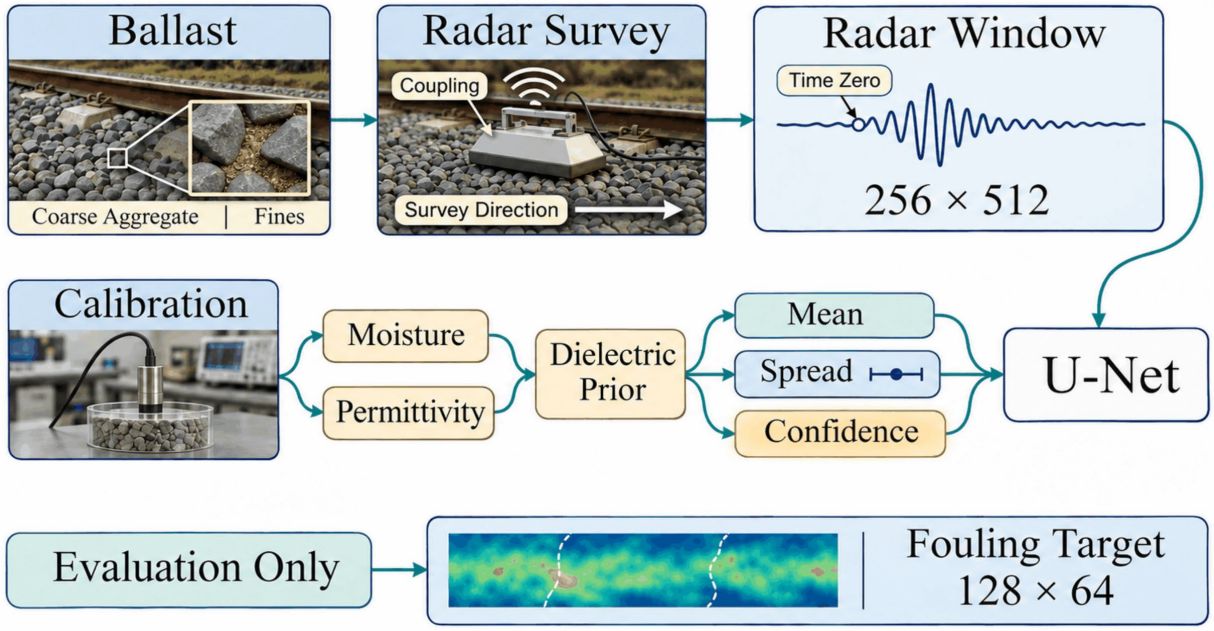


Figure 1. Observation and prior construction workflow with independent target support and a calibration boundary

3.2. Moisture conditioned dielectric representation

The dielectric prior is represented by a mean field, a spread field, and a confidence mask on the output support. Constituent volume fractions include coarse solids, fine solids, water, and air. The moisture descriptor is defined as the water share of pore volume, keeping saturation and solid packing as separately represented quantities. A square root mixing equation provides the conditional material relation:

$$\sqrt{\varepsilon} = v_c\sqrt{\varepsilon_c} + v_f\sqrt{\varepsilon_f} + v_w\sqrt{\varepsilon_w} + v_a\sqrt{\varepsilon_a} \tag{2}$$

The subscripts c, f, w, and a identify coarse solids, fine solids, water, and air, respectively. Each v denotes the corresponding fraction of total mixture volume, and each constituent permittivity is represented by ε with the matching subscript. The fractions are nonnegative and sum to unity. Porosity supplies the conversion between the solid-based target and total-mixture fractions: fine and coarse solids occupy the non-pore volume in their target proportions, while saturation divides pore volume between water and air. The relation supplies a conditional parameterization whose applicability to conductive or strongly dispersive fines requires separate calibration.

Calibration requires specimen level information on constituent permittivity and uncertainty over the operating bandwidth. The calibration set is separated from the reconstruction test set by specimen identity. A dielectric estimate obtained through travel time analysis also requires an independently constrained interface depth because thickness and propagation velocity can compensate for one another. The workflow records whether the prior originates from direct dielectric measurements, a calibrated travel time estimate, or an auxiliary inversion. These sources can have different

error structures even when their nominal uncertainty is similar. Their performance should therefore be evaluated separately before pooled reconstruction results are reported.

The spread field is a nonnegative dispersion scale for effective permittivity, including propagated moisture uncertainty; its statistical interpretation must follow the calibration procedure used to obtain it. Where the moisture estimate is uncertain, its admissible range produces a wider dielectric spread. A spatial confidence mask identifies regions outside the calibration domain, near poorly localized interfaces, or beyond reliable signal support. Confidence expresses the admissibility of the conditioning information before reconstruction; final prediction error requires a separate assessment. Defining confidence from independent calibration information prevents agreement between the prior and prediction from becoming its own justification. The same prior construction procedure is applied to training and test windows.

3.3. Reliability scaling and identifiability

Dielectric spread is translated into a bounded reliability weight using an exponential scaling equation:

$$q = Mexp\left(-\frac{s_\varepsilon}{\tau}\right) \tag{3}$$

In this formulation, s_ε denotes the local dielectric spread, M is the spatial confidence mask, and τ controls the sensitivity of reliability to uncertainty. The scale parameter has the same units as the spread and is selected from calibration data before evaluation. Reproduction requires the selected scale and the spread estimator, which are not specified numerically in the present record. The mask is binary in the simplest implementation and can take fractional values where reliable coverage tapers. A zero weight suppresses the direct dielectric contribution at that location. The radar pathway remains available, with identifiability determined by the waveform information and the training assumptions.

The formulation separates reliability from the nominal prior value. Two locations can share the same dielectric mean and have different confidence because one lies within the calibration range while the other requires extrapolation. Training includes both situations. Randomly dropping the prior in selected windows exposes the architecture to incomplete auxiliary information. Spatially correlated perturbations additionally represent calibration errors that extend across neighboring cells. Their correlation length reflects the spatial support of calibration uncertainty and the smoothing introduced by auxiliary inversion. This treatment makes the perturbation structure consistent with the way dielectric information is obtained.

Identifiability is assessed against the available observation and conditioning information. When moisture and fouling produce overlapping dielectric responses within the measured bandwidth, the reconstruction retains the corresponding material ambiguity. The output is then a conditional estimate under the training distribution, with uniqueness requiring additional material information. The stress tests include moisture shifts, biased priors, and missing confidence regions to examine this dependency. Reconstruction sensitivity is evaluated as supporting information weakens, and spatial sharpness is interpreted alongside the available sensing support. These considerations motivate confidence gating and separate reporting of local error distributions.

4. Dielectric Prior Guided U-Net

4.1. Dual encoders and confidence gated fusion

The radar encoder uses four spatial scales with nominal channel counts of 32, 64, 128, and 256. At each scale, two convolutional blocks combine a small receptive field with normalization and a smooth activation. Downsampling expands the along track context while retaining local features through skip connections. A separate prior encoder processes the dielectric mean, dielectric spread, and confidence field using fewer channels. The radar input remains in time coordinates until a learned resampling stage maps its features to the declared depth support. This stage is conditioned on the nominal propagation range and fitted using training specimens. Its configuration is recorded explicitly because the conversion from time to depth determines the spatial interpretation of the reconstructed features.

At each scale, radar and prior features interact through the gated correction equation:

$$h_\ell = r_\ell + q_\ell \odot sigmoid\big(G_\ell(r_\ell, p_\ell)\big) \odot P_\ell(p_\ell) \tag{4}$$

Here, r_ℓ and p_ℓ denote radar and prior features at scale ℓ, while h_ℓ is the fused representation. The projection P_ℓ aligns the prior channel dimensions with those of the radar features. The learned mapping G_ℓ evaluates feature compatibility, and the resampled reliability field q_ℓ limits the auxiliary contribution. The circled-dot operator denotes elementwise multiplication; reliability is broadcast across feature channels after spatial alignment. The additive structure

preserves the radar pathway as confidence decreases. Learned compatibility and calibrated reliability therefore control different aspects of the same fusion process.

Figure 2 presents the architecture as two encoders feeding a depth aligned decoder, with confidence gates at the skip connections and a dielectric consistency branch during training. The decoder upsamples fused features and predicts a bounded fouling fraction. Bilinear upsampling followed by convolution provides a controlled increase in spatial support. A final mask restricts the training objective to valid ballast cells, excluding air, sleepers, and depths outside the defined reconstruction region. The mask requires independent validation because an incorrectly located layer boundary can alter the evaluated support and bias the reported reconstruction performance.

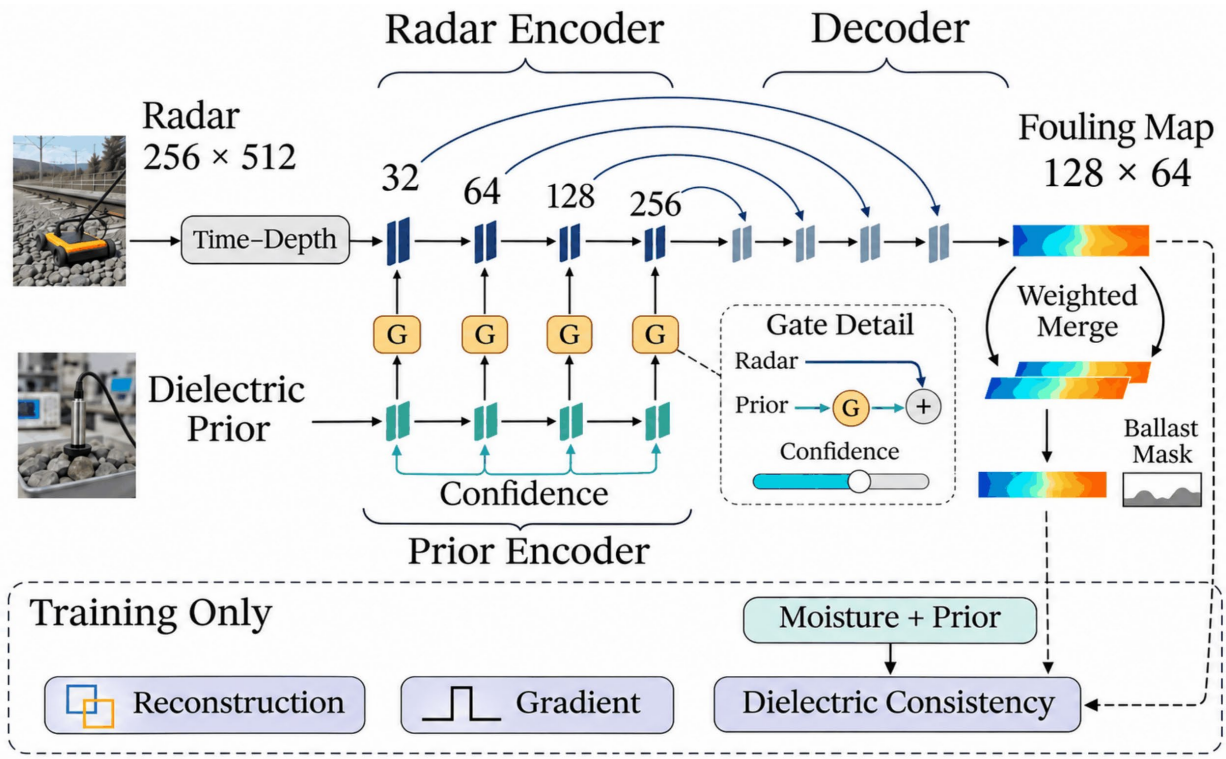


Figure 2. Dual encoder U-Net with confidence gated multiscale fusion and a training only dielectric consistency branch

4.2. Reconstruction and conditional consistency losses

The supervised objective measures fouling error over valid target cells. Its reconstruction component is intended to reduce sensitivity to isolated label discrepancies. The exact penalty and gradient discretization remain implementation specifications requiring disclosure before the objective can be reproduced. An interface component compares spatial gradients on the same support and uses a stabilizing offset to control the contribution of nearly constant regions. After normalization by valid support counts, the components enter a composite loss equation:

$$\mathcal{L} = \mathcal{L}_{rec} + \lambda_g \mathcal{L}_{grad} + \lambda_d \mathcal{L}_{diel} \tag{5}$$

The reconstruction, gradient, and dielectric terms are denoted by $\mathcal{L}_{\text{rec}}$, $\mathcal{L}_{\text{grad}}$, and $\mathcal{L}_{\text{diel}}$, respectively. Their relative contributions are controlled by λ_g and λ_d. Support normalization gives windows with different valid ballast areas a comparable contribution to optimization. Both weights are selected using validation specimens and fixed before final evaluation.

Within this objective, dielectric consistency is evaluated through an uncertainty weighted equation:

$$\mathcal{L}_{diel} = mean_{\,valid}\left[q\,\frac{\left|\varepsilon(\hat{f}, m) - \varepsilon_p\right|}{s_\varepsilon + \delta}\right] \tag{6}$$

The predicted fouling fraction is denoted by $\hat{f}$, and ε_p is the supplied dielectric prior mean. The mixing relation converts the predicted material fraction and available moisture descriptor into an effective permittivity. The spread scales the mismatch, while reliability attenuates uncertain regions. The positive stabilizer δ controls the penalty when the estimated spread approaches zero. Differentiation passes through the material relation and predicted fouling fraction; the prior mean and spread remain fixed conditioning inputs. Reliability and inverse spread both attenuate uncertain locations, so calibration controls the effective strength of the dielectric penalty as well as feature fusion. The consistency

term constrains the material relation; waveform propagation enters through the observations rather than a differentiated wave-equation residual.

The loss weights control the balance between material consistency and spatial fidelity. A strong dielectric penalty can favor a smooth but biased material field when the moisture descriptor is biased. A strong gradient penalty can preserve transitions while amplifying weak boundaries. Validation therefore examines continuous error, structural similarity, and boundary displacement jointly. The starting weights are 0.15 for the gradient component and 0.10 for dielectric consistency after component normalization. Ablation and validation sensitivity analysis assess the balance among these terms. The training record includes learning curves and the response of validation performance to both weights.

4.3. Training controls and inference support

The training protocol partitions specimen identities before extracting radar windows. All windows from a given specimen remain in the same partition. The ratios of training, validation and test sets are 70%, 15% and 15%, respectively. Normalization parameters are obtained from the training data. Acquisition augmentation varies amplitude scale, time zero, band limitation and correlated noise within predefined ranges. The target fraction is unchanged in these transformations, and material augmentation necessitates the regeneration of the corresponding configuration. Separate prior perturbations introduce controlled disagreement of the radar observation and conditioning channels. This configuration exposes the model to uncertainty in both acquisition and auxiliary material information.

The starting learning rate for the optimizer is 0.0002, the batch size is 8, and the maximum number of epochs is 150. Early stopping monitors specimen-averaged validation error and has a patience of 20 epochs. Five independent optimization seeds are set for repeated training with fixed specimen partitions. Baselines are given the same target masks, preprocessing and evaluation support. A capacity-matched radar-only control is required to separate conditioning benefits from additional capacity; its performance is absent from the reported model comparison. The reproducibility record contains the checkpoint-selection rule, optimization seed, normalization parameters and prior preparation settings for each configuration.

Tapering weights that sum to one over the valid reconstruction support are used to combine overlapping windows during inference. Border regions receive lower weights because less spatial context is available. The retained outputs comprise the fouling estimate, supplied prior reliability and an indicator of whether the input lies outside the calibration domain. These quantities are stored separately in the inspection record. Predictive uncertainty of ensembles or repeated perturbations needs to be calibrated with respect to independent fouling reference errors. The reported processing breakdown covers preprocessing, prior projection, encoding, fusion and decoding; independent prior acquisition and overlap merging require separate timing. This accounting is relevant when auxiliary dielectric inversion contributes substantially to total processing cost. The final material estimate is provided in the native grid and spatial interpretation is controlled by the acquisition footprint and validated reconstruction support.

5. Experimental Design and Reconstruction Analysis

5.1. Dataset coverage and evaluation protocol

This section examines dataset coverage, reconstruction accuracy, and sensitivity to dielectric prior quality within a common evaluation framework. Field fouling assessment has combined dynamic cone resistance and geoendoscopic image characteristics, with laboratory material characterization used to interpret ballast condition [16]. The target follows the volumetric definition introduced in Section 3 and spans six fouling fractions from 0.05 to 0.55. The dataset specification contains 1,200 scenario groups and 20 radar windows per scenario, yielding 24,000 windows. Scenario independence requires provenance and grouping records; the counts alone establish the stated partition sizes. Training, validation, and testing receive 840, 180, and 180 scenarios, respectively. Profile-based soil mapping studies show that assigning correlated samples from one profile to both training and testing can inflate predictive accuracy, motivating profile-level partitioning [17]. Table 1 records these partitions together with the input dimensions, reconstruction support, moisture states, and noise conditions. Scenario identity determines the partition before window extraction. Independent auditing requires acquisition location or specimen identity, antenna bandwidth, reference measurement procedure, aggregate lithology, fines mineralogy, packing state and sampling depth. These records are outside the reported dataset specification, which limits interpretation of field transfer and calibration traceability.

Table 1. Dataset specification and information boundaries

Property	Specification
Scenario groups	1,200; 840 training / 180 validation / 180 test
Radar windows	24,000; 20 per scenario; 3,600 tests
Fouling fraction	0.05, 0.15, 0.25, 0.35, 0.45, 0.55
Pore-water saturation	0.05, 0.15, 0.25, 0.35, 0.45
Noise levels	5, 10, 15, 20, 30 dB
Input and output	256 × 512 radar samples; 128 × 64 target cells
Prior information	Mean, spread and confidence; no target derived test input

Figure 3(a) displays solid composition: fine-solid fractions of 0.05, 0.15, 0.25, 0.35, 0.45 and 0.55 correspond to coarse-solid fractions of 0.95, 0.85, 0.75, 0.65, 0.55 and 0.45. Figure 3(b) partitions pore volume between water and air at saturations from 0.05 to 0.45, with complementary air fractions from 0.95 to 0.55. These panels describe normalized constituent proportions. The stated test allocation contains 600 windows per fouling stratum and 720 per saturation group across the same 3,600 windows; window counts remain distinct from physical mixture fractions. Figure 3(c) converts signal-to-noise ratios of 5, 10, 15, 20 and 30 dB to noise-to-signal root-mean-square amplitude ratios of 0.562, 0.316, 0.178, 0.100 and 0.032 on a common signal support. Dielectric-based soil water calibration can be affected by differences in sensing volume and by site-specific soil conditions [18]. Equal marginal counts describe coverage within individual factors; joint scenario allocation is required to assess confounding among fouling, saturation and noise.

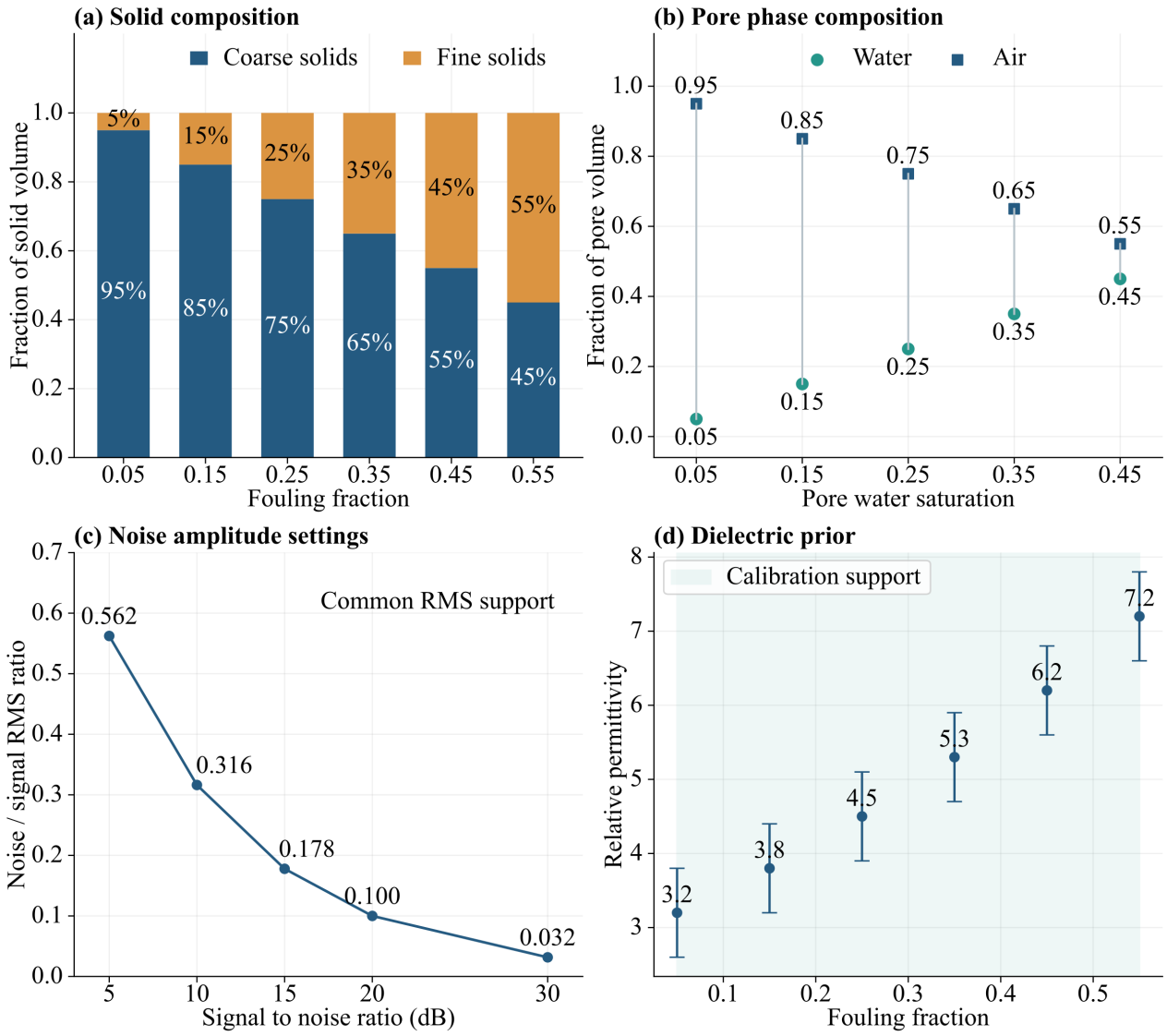


Figure 3. Material and acquisition settings with (a) solid composition, (b) water and air fractions of pore volume, (c) noise-to-signal RMS amplitude ratios, and (d) dielectric prior means and spread

Figure 3(d) gives dielectric means of 3.2, 3.8, 4.5, 5.3, 6.2, and 7.2 across the six fouling fractions, with a spread of 0.6 units. Successive mean increments increase from 0.6 to 1.0, indicating a nonlinear conditioning relation at the central moisture state. Adjacent uncertainty ranges overlap substantially, particularly at lower fouling fractions. The prior therefore supplies a graded material constraint whose discriminating power depends on both the local slope and the uncertainty spread. At the lower end of the range, a dielectric perturbation of fixed size represents a larger ambiguity in fouling fraction than it does where the relation is steeper. The shaded calibration support covers fractions from 0.05 to 0.55 and identifies the interval within which this conditional relation is specified. Moisture variation introduces an additional source of displacement relative to that central relation. Comparing the prior spread with the spacing of the means makes the uncertainty visible at the stage of material conditioning, before its influence is transformed by feature fusion or the dielectric loss. An independently calibrated experiment would estimate these ranges for each relevant moisture state and acquisition bandwidth.

Structural similarity evaluates luminance, contrast and structural information, while remaining sensitive to geometric misalignment between reference and reconstructed images [19]. The evaluation fixes the data range at one and applies the same ballast mask to every model. Mean absolute error and root mean square error retain the dimensionless scale of the volumetric fouling target. A threshold of 0.25 defines the affected region for diagnostic overlap analysis, while boundary displacement measures transition geometry on the same support. This threshold and the 3 cm displacement guide are evaluation settings whose maintenance relevance requires independent justification. Resampling spatial data requires attention to autocorrelation because ignoring spatial dependence can underestimate variability and overstate statistical significance [20]. Accordingly, scenario-level averages are the comparison unit, with all windows belonging to a scenario retained together during partitioning and any resampling. The 180 test scenarios and the five optimization seeds represent different sources of variation. Missed transitions are retained as a separate count because displacement is defined only for matched boundaries. When a reference sample integrates material over a finite depth interval, the prediction is averaged over that interval before continuous error is calculated. This support matching is particularly relevant near diffuse transitions, where a pointwise prediction and a bulk material reference can represent different quantities. The evaluation record retains the averaging interval, valid mask, and specimen identity for each reported error.

5.2. Reconstruction comparisons and spatial behavior

Signal-envelope and frequency features extracted from GPR measurements have been used with regression models to estimate railway ballast fouling indices [21]. The evaluation compares this baseline with a conventional U-Net, an attention U-Net, a concatenated prior model, and the confidence gated model. Table 2 lists mean absolute errors of 0.121, 0.086, 0.074, 0.062, and 0.048 in that order. Root mean square errors are 0.158, 0.116, 0.101, 0.087, and 0.069. Relative to the conventional U-Net, the gated configuration has an absolute error decrease of 0.038 and a relative decrease of 44.2 percent in the comparison. The larger root means square errors show the influence of more pronounced local discrepancies within the error summaries. GPR inversion architectures have incorporated attention gates and convolutional block attention to emphasize features associated with geological anomalies [22]. The corresponding comparison controls should preserve specimen partitions, preprocessing, target masks, and optimization budgets. The stated parameter counts of 7.8, 8.4, and 8.6 million for the conventional, concatenated, and gated models introduce a capacity difference that requires a matched radar only control under matched evaluation conditions. Acquisition of the auxiliary prior also belongs in the computational accounting.

Table 2. Model reconstruction performance

Model	MAE	RMSE	SSIM	IoU
Attribute regression	0.121	0.158	0.684	0.492
U-Net	0.086	0.116	0.781	0.614
Attention U-Net	0.074	0.101	0.817	0.663
Concatenated prior	0.062	0.087	0.851	0.718
Confidence gated	0.048	0.069	0.893	0.782

Figure 4(a) shows ten specimen error values per model, with medians decreasing from 0.121 for attribute regression to 0.048 for confidence gated fusion. The conventional and gated distributions extend approximately from 0.071 to 0.101 and from 0.038 to 0.058, respectively. These ten displayed values describe the plotted cases; their selection rule and relationship to the full test set require disclosure before they can characterize population variability or seed uncertainty. Repeated evaluations and appropriate statistical tests are recommended when comparing machine-learning models [23]. Scenario identity is therefore retained in the paired distribution record. Figure 4(b) relates structural similarity to thresholded overlap, with the conventional model at 0.781 and 0.614 and the gated model at 0.893 and 0.782. The

absolute changes are 0.112 in structural similarity and 0.168 in overlap. Their common direction indicates agreement between continuous spatial appearance and affected region localization within this comparison. The intermediate concatenated model reaches 0.851 and 0.718, leaving a further change of 0.042 and 0.064 when confidence gating and the full objective are introduced. The paired metrics permit improvements in texture or smoothing to be distinguished from improvements in thresholded region geometry.

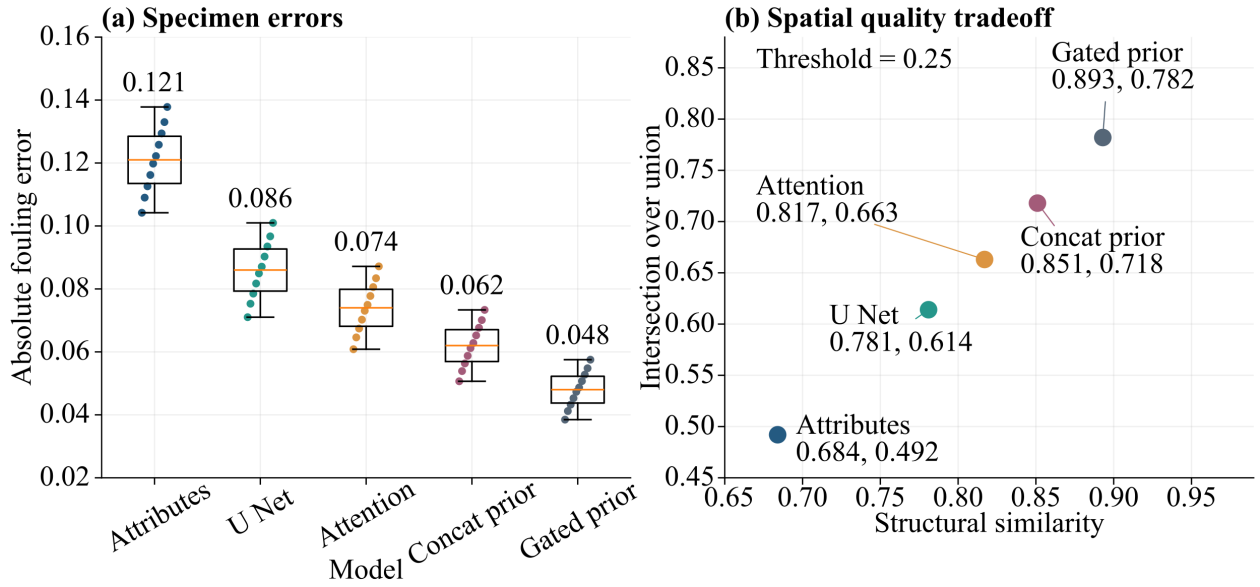


Figure 4. Model comparison with (a) specimen error distributions and (b) structural similarity versus thresholded overlap

Figure 5(a) compares four depth profiles at eight sampling positions, with target fractions of 0.08, 0.10, 0.18, 0.31, 0.43, 0.47, 0.32, and 0.16. The conventional prediction overestimates the shallow low fouling region and underestimates the central peak, producing a compressed profile range. At the target maximum of 0.47, its value is 0.40, compared with 0.44 for concatenated conditioning and 0.46 for the gated model. The resulting peak discrepancies are 0.07, 0.03, and 0.01. This selected profile localizes error around one peak; its spatial behavior should be interpreted alongside the aggregate metrics, because a single transect provides limited evidence about other fouling geometries. The crossing of the 0.25 threshold also depends on the shape of the rising and falling transitions. A prediction can preserve the average fraction while changing the depths assigned to an affected region. Near the bottom of the profile, the conventional estimate of 0.22 exceeds the target of 0.16, whereas the gated estimate is 0.17. The shallow and deep departures therefore occur on opposite sides of the central underestimation. Their spatial arrangement illustrates the value of retaining signed profile residuals alongside a pooled absolute error, particularly when a transition is spread across several reconstruction cells.

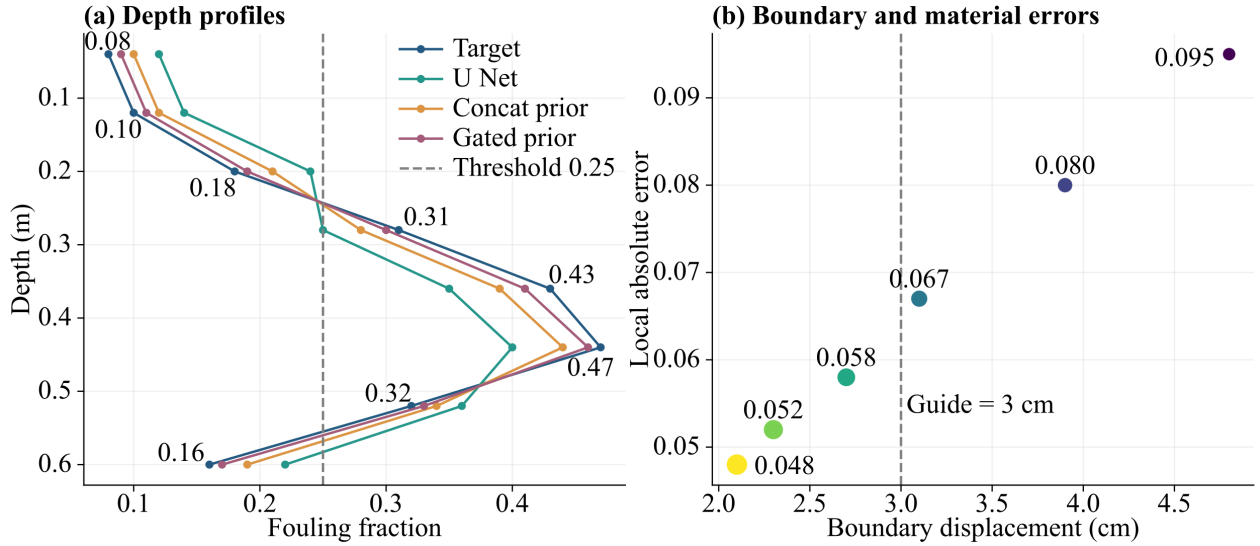


Figure 5. Spatial reconstruction with (a) depth profiles and (b) boundary displacement versus local error

Antenna frequency imposes a trade-off between GPR penetration depth and spatial resolution, affecting the interpretation of reconstructed interfaces [24]. Figure 5(b) pairs boundary displacements of 4.8, 3.9, 3.1, 2.7, 2.3, and

2.1 cm with local errors of 0.095, 0.080, 0.067, 0.058, 0.052, and 0.048. Across these points, the displacement decreases by 2.7 cm and the local error decreases by 0.047. The 3 cm guide separates the three smaller displacement cases from those with larger offsets. The error decrease across the nearest points on either side of that guide, from 0.067 to 0.058, is gradual, consistent with its role as a diagnostic marker. The plotted association relates interface localization to local material reconstruction error. Its physical interpretation depends on the precision of the reference boundary and on whether the transition is sharp or diffuse. For a bulk sample spanning several output cells, an uncertain reference depth can contribute to both displacement and local fraction error. Retaining the sampled depth interval and transition width would allow those contributions to be examined at matched spatial supports.

5.3. Ablation and prior mismatch stress tests

Ablation experiments in GPR inversion have assessed the contributions of multiscale convolution, attention modules and loss-function design [25]. Figure 6(a) assigns errors of 0.086, 0.062, 0.057, 0.054, and 0.048 to radar only inference, concatenation, removal of consistency, removal of gating, and the full model. Removing consistency increases the full model error by 0.009; removing gating increases it by 0.006. These contrasts support separate examination of the material constraint and reliability mechanism. Moisture in ballast contributes to electromagnetic absorption and preferential loss of high-frequency radar energy with increasing propagation depth [26]. Figure 6(b) gives errors of 0.044, 0.046, 0.048, 0.056, and 0.067 as saturation rises from 0.05 to 0.45. The final saturation increment produces an error increase of 0.011, compared with 0.002 for the first increment. This curvature concentrates the degradation toward wetter conditions and crosses the 0.060 diagnostic guide between the two highest states. Figure 6(c) allocates 2, 4, 6, 8, and 12 ms to preprocessing, prior projection, radar encoding, fusion, and decoding. Decoding occupies 37.5 percent of the 32 ms total, while fusion occupies 25 percent. The 32 ms subtotal excludes independent dielectric acquisition and overlap merging. Hardware, batch size and repeated timing records are required before interpreting this subtotal as deployment throughput.

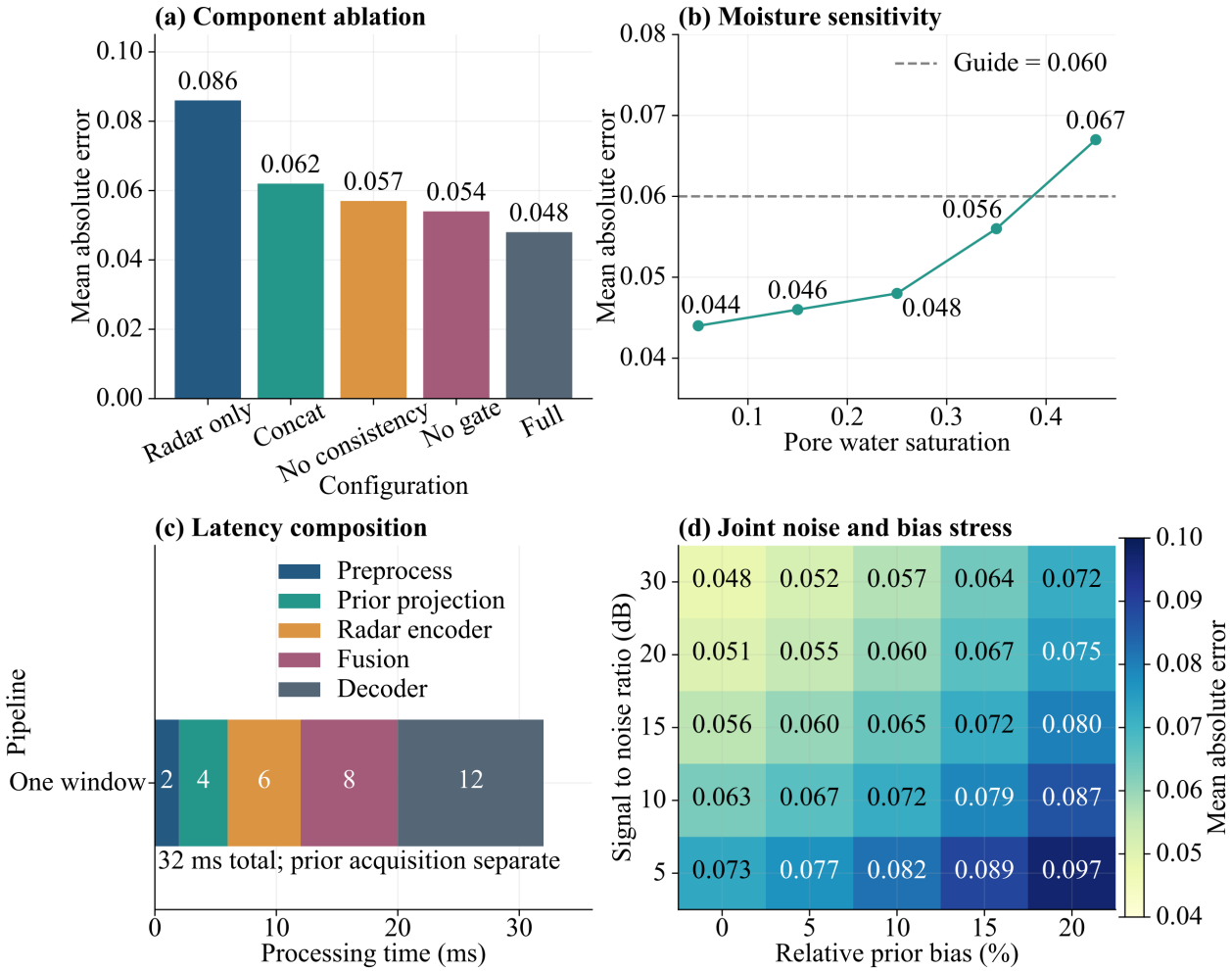


Figure 6. Component and condition analysis with (a) ablation errors, (b) moisture dependence, (c) latency composition, and (d) the noise and prior bias error surface

Figure 6(d) exposes the joint effect of five noise levels and five prior biases through 25 errors. At zero bias, error increases from 0.048 at 30 dB to 0.073 at 5 dB. At 20 percent bias, the corresponding values are 0.072 and 0.097. The

bias penalty is therefore 0.024 at both noise extremes, and the low signal penalty is 0.025 at both bias extremes. The equal endpoint penalties are consistent with an additive pattern within the displayed grid; evidence for interaction requires uncertainty estimates or an explicit interaction analysis. Interaction would appear as a changing bias penalty across noise levels, particularly where a weakened radar signal increases reliance on an incorrect prior. Incomplete multimodal learning explicitly addresses robustness when some input modalities are unavailable during inference [27]. Figure 7(a) compares those bias effects for concatenated and gated fusion. Their errors rise from 0.062 to 0.098 and from 0.048 to 0.072 over biases from zero to 20 percent. The increases are 0.036 and 0.024, respectively. The separation between models grows from 0.014 to 0.026, showing lower bias sensitivity for the gated configuration over the displayed range. The dependence of that pattern on the supplied confidence values would require testing with deliberately miscalibrated reliability masks.

Figure 7(b) shows median errors of 0.048, 0.051, 0.055, 0.061, and 0.068 as missing prior coverage rises from zero to 40 percent. The absolute increase is 0.020, equivalent to 41.7 percent of the central error with complete coverage. The displayed extents widen from half widths of 0.014 to approximately 0.022; their quantile or interval definition is unspecified, so they describe plot extent rather than calibrated confidence coverage. The upper extent reaches approximately 0.090 at the largest coverage loss, indicating how selected cases can depart from a modest group median in the distributions. Coverage is removed in contiguous regions, so the perturbed input contains a spatial gap whose size interacts with the receptive field of the network. A small gap may retain useful neighboring dielectric context, while a larger gap leaves more of the reconstruction dependent on radar features and learned spatial regularity. The widening distributions associate larger gaps with greater variation in reconstruction error. Recording gap position and depth permits cases near the bottom interface to be separated from those in well illuminated shallow ballast. The same missing area can affect different target structures depending on whether it intersects a transition or a relatively uniform layer.

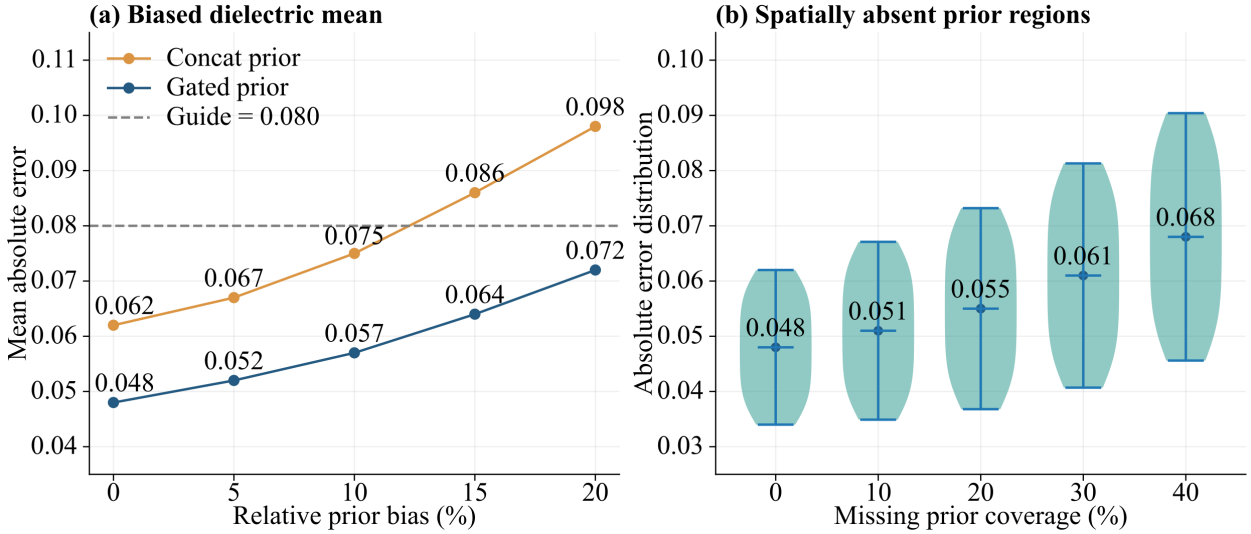


Figure 7. Prior robustness with (a) bias response and (b) errors under spatially missing prior coverage

Spatial regression can exhibit biased estimates or invalid inference when explanatory variables are associated with spatially correlated residuals [28]. Residual assessment therefore considers both spatial structure and distributional departures. Figure 8 compares 120 standardized residual values with normal reference scores and marks the upper region beyond 1.5. The largest plotted residual is about 2.90 and its normal score is around 2.64; thus, the upper departure is about 0.26. The central residuals remain closer to the reference, while the departure is concentrated above the indicated score. The departure is concentrated in the positive standardized tail. A quantile plot alone leaves spatial dependence unresolved; interpretation as fouling overestimation additionally requires the residual sign and standardization scale, which are absent from the reported diagnostic. The upper-tail cases can then be examined against material condition, interface position and prior coverage. A conductive fines pocket, an uncertain reference depth and a gap in the conditioning field represent distinct sources of error. Sample support also affects the diagnostic; predictions pooled in a bulk reference interval may have a different residual distribution than those of individual grid cells. The retained record would connect each residual to its specimen, depth interval, acquisition settings, moisture descriptor and calibration domain, thus enabling the study of these sources of variation independently.

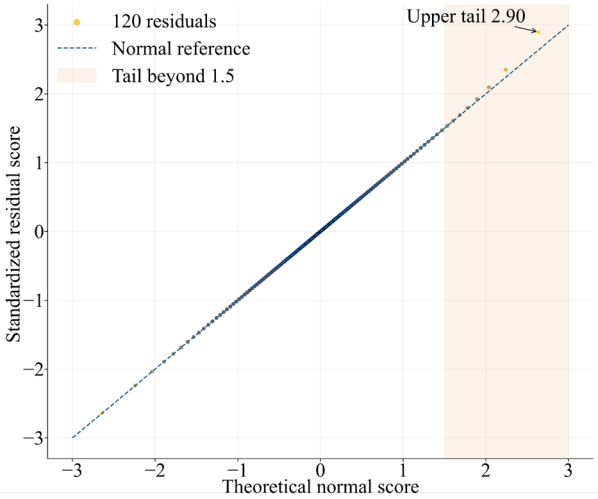


Figure 8. Residual quantile diagnostic

6. Conclusion

A dielectric prior-guided U-Net was proposed to reconstruct continuous ballast fouling distributions from ground-penetrating radar in this paper. Separation of radar features and auxiliary dielectric information, introduction of confidence-gated fusion, and coupling of supervised reconstruction with conditional material consistency are employed in this work. The target definition differentiates between solid fines and pore water, and explicitly shows the spatial support for reconstruction. The experimental design links continuous error, structural fidelity, boundary localization and prior mismatch in a common specimen-based evaluation system. The methodological contribution concerns the integration and evaluation of reliability-weighted material conditioning; the reported comparisons support this formulation within the stated scenario coverage.

The main limitations concern the calibration of dielectric uncertainty, assumed material relationship and spatial support for observation. The applicability of the assumed mixing relationship requires further assessment for conductive fines, frequency-dependent losses and changes in packing. Calibration of prior confidence is required with respect to independently determined dielectric uncertainty. A two-dimensional section also excludes lateral heterogeneity outside the survey plane, and the proposed target grid is finer than the resolution that may be supported by the radar footprint. These conditions bound the present interpretation. Practical use requires independently audited material references and a validated connection between reconstructed conditions and inspection criteria.

Future work should examine controlled specimens spanning independent combinations of moisture, fines content and fouling geometry. Compatible spatial supports should be used for the collection of dielectric data and material reference samples, and calibration specimens should be kept separate from the final evaluation. Priority should be given to independent reference measurements and confidence calibration, followed by frequency-dependent material models and transfer among antennas and track sites. Field validation should assess the stability of spatial estimates under changes in antenna coupling and the effect of localized errors on inspection decisions. Publishing acquisition settings, specimen identities, trained weights and evaluation scripts would enable an independent evaluation of the proposed conditioning mechanism and clarify the conditions for its application in providing useful material information.

7. Declarations

7.1 Author Contributions

Marek Ciesielski contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Robert Stępień contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

7.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

7.3 Institutional Review Board Statement

Not applicable.

7.4 Informed Consent Statement

Not applicable.

7.5 Date Availability

Not applicable.

7.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

8. References

[1] Charoenwong, C., Connolly, D. P., Alves Costa, P., Galvín, P., & Romero, A. (2024). The effect of ballast moisture content and fouling index on railway track settlement. Transportation Geotechnics, 45, Article 101193. https://doi.org/10.1016/j.trgeo.2024.101193

[2] Rahman, M. H., Bhuyan, M. H., & Hayano, K. (2025). Fouling conditions and associated deformation characteristics of railway ballast. International Journal of Rail Transportation, 13(3), 425–443. https://doi.org/10.1080/23248378.2024.2363919

[3] Koohmishi, M., Kaewunruen, S., Chang, L., & Guo, Y. (2024). Advancing railway track health monitoring: Integrating GPR, InSAR and machine learning for enhanced asset management. Automation in Construction, 162, Article 105378. https://doi.org/10.1016/j.autcon.2024.105378

[4] Bu, J., Jing, G., Long, X., Wang, L., Peng, Z., & Guo, Y. (2025). Intelligent classification of ballast bed defects using a bimodal deep learning model. Transportation Geotechnics, 50, Article 101464. https://doi.org/10.1016/j.trgeo.2024.101464

[5] Junkai, G., Huaifeng, S., Wei, S., Dong, L., Yuhong, Y., Yi, Z., ... & Shangbin, L. (2024). GPR-TransUNet: An improved TransUNet based on self-attention mechanism for ground penetrating radar inversion. Journal of applied geophysics, 222, 105333. https://doi.org/10.1016/j.jappgeo.2024.105333

[6] Fang, Y., Tong, Z., Hei, T., Wang, S., & Ma, T. (2026). Deep learning applications in ground-penetrating radar inversion: A review. Measurement, 258, Article 119399. https://doi.org/10.1016/j.measurement.2025.119399

[7] Sun, J., Liu, Y., Lin, Y., Xing, L., & Liu, H. (2025). Implicit multiparameter full waveform inversion of multioffset ground penetrating radar data. Geophysical Journal International, 240(2), 904–919. https://doi.org/10.1093/gji/ggae420

[8] Aziz, I., Soltanaghai, E., Watts, A., & Alipour, M. (2024). Bayesian inversion of GPR waveforms for sub-surface material characterization: An uncertainty-aware retrieval of soil moisture and overlaying biomass properties. Remote Sensing of Environment, 313, Article 114351. https://doi.org/10.1016/j.rse.2024.114351

[9] Shapovalov, V., Arkhipov, V., Okost, M., & Morozov, A. (2026). Comparison of approaches to assessing ballast layer contamination using ground penetrating radar. International Journal of Transportation Science and Technology, 21, 286–305. https://doi.org/10.1016/j.ijtst.2025.02.001

[10] Mbubia Tchoua, E., Tissier, J., Martin, A., Fargier, Y., & Ihamouten, A. (2026). The use of ground penetrating radar and artificial intelligence for automated railway trackbed stratigraphy and ballast fouling assessment. Transportation Engineering, 23, Article 100415. https://doi.org/10.1016/j.treng.2025.100415

[11] Wu, M., Liu, Q., & Ouyang, S. (2025). Two-stage GPR image inversion method based on multi-scale dilated convolution and hybrid attention gate. Remote Sensing, 17(2), Article 322. https://doi.org/10.3390/rs17020322

[12] Zhang, H., Hu, Y., Wang, Q., Qian, Z., & Liu, P. (2025). End-to-end frequency enhancement framework for GPR images using domain-adaptive generative adversarial networks. Computer-Aided Civil and Infrastructure Engineering, 40(22), 3466–3482. https://doi.org/10.1111/mice.13525

[13] Wang, Z., Jiang, H., Wang, W., Xu, J., Sui, Q., Gao, S., & Wang, J. (2025). A semisupervised deep learning method for ground-penetrating radar data inversion and concrete subsurface defect imaging. IEEE Transactions on Geoscience and Remote Sensing, 63, Article 4503511. https://doi.org/10.1109/tgrs.2025.3546212

[14] Awais, M., Li, L., Zaigham Abbas Naqvi, S. M., Zhang, W., Wu, J., Tlili, I., & Hu, J. (2026). Coupling dielectric physics with calibration models for soil moisture sensing: Progress and perspectives. Ain Shams Engineering Journal, 17(2), Article 103968. https://doi.org/10.1016/j.asej.2025.103968

[15] Jiao, T., Guo, C., Feng, X., Chen, Y., & Song, J. (2024). A comprehensive survey on deep learning multi-modal fusion: Methods, technologies and applications. Computers, Materials & Continua, 80(1), 1–35. https://doi.org/10.32604/cmc.2024.053204

[16] Rojas Vivanco, J., Breul, P., Talon, A., Benz-Navarrete, M. A., Barbier, S., & Ranvier, F. (2024). On site fouling index assessment of ballasted layers using cone dynamic penetrometer and miniature borescope image analysis. Transportation Geotechnics, 45, Article 101213. https://doi.org/10.1016/j.trgeo.2024.101213

[17] John, K., Saurette, D. D., & Heung, B. (2025). The problematic case of data leakage: A case for leave-profile-out cross-validation in 3-dimensional digital soil mapping. Geoderma, 455, Article 117223. https://doi.org/10.1016/j.geoderma.2025.117223

[18] Pathirana, S., Lambot, S., Krishnapillai, M., Smeaton, C., Cheema, M., & Galagedara, L. (2024). Potential of ground-penetrating radar to calibrate electromagnetic induction for shallow soil water content estimation. Journal of Hydrology, 633, Article 130957. https://doi.org/10.1016/j.jhydrol.2024.130957

[19] Agaian, S. S., Grigoryan, A. M., & Ayunts, H. (2026). A survey and tutorial on image quality assessment with a contrast-weighted structural similarity framework. Information, 17(7), Article 632. https://doi.org/10.3390/info17070632

[20] Arthur, R. (2025). A general method for resampling autocorrelated spatial data. Geographical Analysis, 57(2), 302–319. https://doi.org/10.1111/gean.12417

[21] Rojas-Vivanco, J., Benz-Navarrete, M., García, J., Breul, P., Talon, A., & Villavicencio, G. (2025). Assessment of railway ballast fouling using GPR and AI-based learning from LDCP and geoendoscopy data. Transportation Geotechnics, 55, Article 101701. https://doi.org/10.1016/j.trgeo.2025.101701

[22] Xu, T., Feng, D., Qin, H., Wang, X., Zhang, B., Chen, C., Li, D., Tao, X., Cai, L., Xiao, Y., & Cao, W. (2025). TGPInvNet: Deep learning-based ground-penetrating radar data inversion for tunnel geological prediction. IEEE Transactions on Geoscience and Remote Sensing, 63, Article 4511225. https://doi.org/10.1109/tgrs.2025.3641163

[23] Lones, M. A. (2024). Avoiding common machine learning pitfalls. Patterns, 5(10), Article 101046. https://doi.org/10.1016/j.patter.2024.101046

[24] Liu, D., Zhou, M., Wang, X., Shi, Z., Sacchi, M. D., Chen, W., Liu, Z., & Wang, X. (2024). Enhancing ground-penetrating radar (GPR) data resolution through weakly supervised learning. IEEE Transactions on Geoscience and Remote Sensing, 62, Article 5107513. https://doi.org/10.1109/tgrs.2024.3410184

[25] Hou, F., Qian, X., Meng, Q., Dong, J., & Lyu, F. (2025). Multi-scale GAN-driven GPR data inversion for monitoring urban road substructure. Automation in Construction, 174, Article 106140. https://doi.org/10.1016/j.autcon.2025.106140

[26] Fu, L., Lan, J., & Xu, Z. (2026). Ground-penetrating radar contamination analysis method based on time–frequency features and ballast condition. Applied Sciences, 16(6), Article 2728. https://doi.org/10.3390/app16062728

[27] Zhan, Y., Yang, R., You, J., Huang, M., Liu, W., & Liu, X. (2025). A systematic literature review on incomplete multimodal learning: Techniques and challenges. Systems Science & Control Engineering, 13(1), Article 2467083. https://doi.org/10.1080/21642583.2025.2467083

[28] Wiecha, N., Hoppin, J. A., & Reich, B. J. (2025). Two-stage estimators for spatial confounding with point-referenced data. Biometrics, 81(3), Article ujaf093. https://doi.org/10.1093/biomtc/ujaf093