International Journal of Sciences and Innovation Engineering
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
website: www.ijsci.com
Volume 3 Issue 9 , September 2026
E-ISSN : 3049-0251
DOI: https://doi.org/10.70849/ijsci03092641938

Provenance-Aware GPR Concrete Cover-Depth Inversion Under Electromagnetic Distribution Shift

Bharanitharan S[1], Joesonraj Santino Noel J[2], Sudharshanan S[3], Chris A[4]

[1,2,3,4] *Department of Civil Engineering, Chennai Institute of Technology, India.*

Article Info

Article History:

Published: 22 Sept 2026

Publication Issue:
Volume 3, Issue 9
September-2026

Page Number:
726-742

Corresponding Author:
Bharanitharan S

Abstract:

Reliable estimation of concrete cover using ground-penetrating radar (GPR) remains challenging because electromagnetic responses are strongly influenced by concrete dielectric properties, reinforcement geometry and acquisition conditions. This study introduces a framework that is aware of the provenance and is based on physical principles in order to assess the robustness and physical transferability of GPR inversion models that have been trained using simulations beyond the stage at which they have been calibrated. A two-dimensional finite-difference time-domain (FDTD) model was used to create 28 controlled cases of reinforced concrete covering four relative permittivities (εr = 6–9) and seven cover depths (30–150 mm). Fixed-calibration and adaptive linear depth-inversion models were developed using the rebar-echo time-of-flight (TOF) and peak amplitude. For the εr = 8 case that was not included, adaptive calibration reduced the RMSE from 6.41 to 3.80 mm and raised the R^2 from 0.9743 to 0.9910. When one relative permittivity was left out for validation, the fixed and adaptive models had mean RMSE values of 9.29 and 8.18 mm, respectively. An ablation analysis showed that TOF was the main predictor, while amplitude offered only a small extra advantage. Importantly, the experimental and field studies that had already been published were kept as an independent benchmark for evidence and were not combined with the synthetic data. The results show that models based on simulation can display robust performance across different dielectric values, while stressing the need for dielectric calibration and field validation if reliable concrete cover assessments are to be made using GPR.

Keywords: Electromagnetic Distribution Shift

1. Introduction

The concrete cover refers to the layer of concrete that lies between the reinforcing steel and the exposed surface of the concrete. Since the thickness of this cover has a strong effect on durability, it slows down the penetration of moisture, chlorides and carbon dioxide, thus influences the risk of corrosion. When it comes to reinforced-concrete structures that are already in place, non-destructive methods are useful since it is possible to evaluate the cover without having to damage the member. GPR is especially useful since the reinforcing bars which are embedded in the concrete cause electromagnetic reflections that can be interpreted in order to estimate both the position of the reinforcement and the depth of the cover [1, 5, 18, 25]. The interpretation of GPR data is still dependent on the conditions in question. The velocity of electromagnetic waves is mainly determined by the effective dielectric properties of concrete, whereas the amplitude of the echo is influenced by attenuation, coupling, propagation distance and electromagnetic contrast at the interface with the reinforcement [1, 19, 25]. It follows that a travel-time-to-depth relationship which has been calibrated for one particular type of concrete will lead to systematic errors if it is used in a different situation. Studies conducted in the field have therefore regarded the estimation and calibration of electromagnetic velocity as essential components of GPR interpretation [1, 22]. Machine learning has steadily increased its role in the interpretation of GPR data, ranging from automatic reinforcement localization [6, 24] to deep-learning based detection, multimodal fusion, diameter classification and cover-depth estimation [9, 15, 17, 19]. Physics-based simulation supplements experimental work since the electromagnetic and geometric variables can be adjusted independently. FDTD models have been used

Copyright © 2025 The Author(s) : This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

to reproduce the GPR responses of reinforced concrete [16] and simulation-to-real learning has proven to be feasible for the estimation of reinforcement diameter and other related tasks [10, 14, 20]. Yet a good performance on randomly divided synthetic observations may lead to an overly optimistic assessment of deployment reliability, since observations produced under the same electromagnetic conditions can appear in both the training and test sets, enabling the model to take advantage of relationships that are specific to those conditions. Real concrete is heterogeneous and exhibits variations in moisture, has aggregate-scale scattering, includes complex reinforcement cages, involves antenna coupling and is affected by acquisition noise all of which are not fully captured by simplified forward models. Evaluation should therefore involve testing against an electromagnetic state that has not been seen before, identify the contribution of features that have a physical meaning, quantify the level of predictive uncertainty and interpret the numerical results in light of real-world evidence without treating the two lines of evidence as being equivalent. A provenance-aware framework of this kind is developed in the study. The synthetic FDTD simulations make up the main quantitative experiment, while published laboratory and field studies provide an external empirical benchmark and are intentionally left not combined with the synthetic observations. The study poses five questions: (i) does increasing the peak amplitude improve the inversion of cover depth under an excluded dielectric condition? (ii) how stable is the adaptive model when there is a change in the relative permittivity? (iii) what does feature ablation tell us about TOF and amplitude? (iv) are the physical interpretations and the magnitudes of the errors generally in agreement with the evidence from published GPR studies? and (v) which of the conclusions are supported by the simulation and which still need to be validated through laboratory or field work?

2. Methods

The workflow contains six stages: electromagnetic simulation, signal acquisition and preprocessing, extraction of physically meaningful features, depth inversion, uncertainty estimation and provenance-aware comparison with external empirical evidence. The physics-guided character arises from the electromagnetic forward model and from restricting the learning problem to interpretable observables rather than high-dimensional image features. The six-stage workflow is summarized in Figure 1.

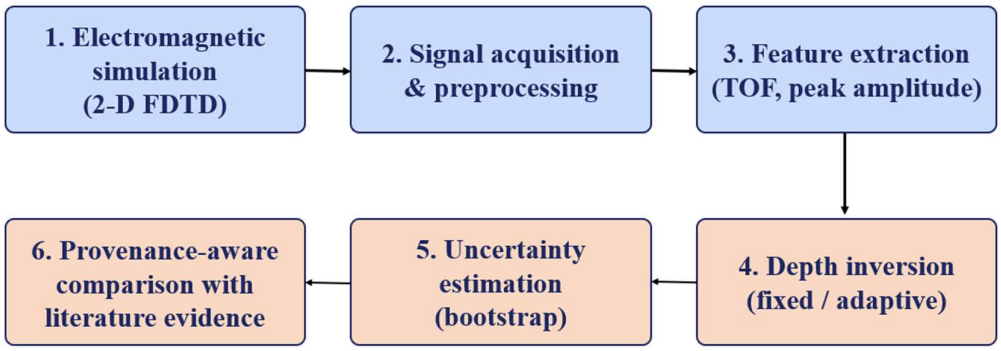

Figure 1. Six conceptual stages of the physics-guided GPR inversion framework

2.1. Synthetic dataset and parameter space

The electromagnetic response was computed using a two-dimensional transverse-magnetic (TMz) Yee FDTD method, in accordance with the standard staggered-grid form of Maxwell's equations [27]. The model under review employs a spatial step of 2 mm in each direction and consists of an 180 by 280 grid together with 15-cell padding. The domain includes 25 cells of air and 225 cells of concrete. The source and receiver are located 30 mm above the interface with the concrete. The excitation is a 1.5 GHz Ricker wavelet having a t_0 of 0.4 ns. The Courant number is 0.5, which results in a time step of $\Delta t = 2.35865$ ps and 2400-time steps lead to a total simulation time of 5.66077 ns. A quadratic lossy sponge with a maximum conductivity of 0.06 S/m is applied in order to minimize boundary reflections. The concrete is treated as homogeneous and lossless, and the reinforcement is modelled as a cylindrical perfectly conducting inclusion with a radius of 8 mm. These values establish the reproducibility boundary of the computational study. The two-dimensional simplified representation is intentional since it allows for a sensitivity analysis to be carried out in a controlled way without introducing any additional uncontrolled variables. In this model, aggregate-scale scattering, three-dimensional reinforcement cages, corrosion-dependent conductivity, antenna coupling and stochastic measurement noise have all been omitted. It must therefore be regarded as a controlled electromagnetic experiment rather than as a full replica of field acquisition.

2.2. Synthetic dataset and parameter space

The dataset contains 28 simulations formed by the Cartesian combination of four relative-permittivity values, εr = 6, 7, 8 and 9, and seven cover depths: 30, 50, 70, 90, 110, 130 and 150 mm. The reinforcement diameter is 16 mm. Each held-out permittivity condition consequently contains seven observations. The compact design makes each fold physically interpretable, although the small number of cases limits statistical power and broad generalization. The resulting parameter space is illustrated in Figure 2.

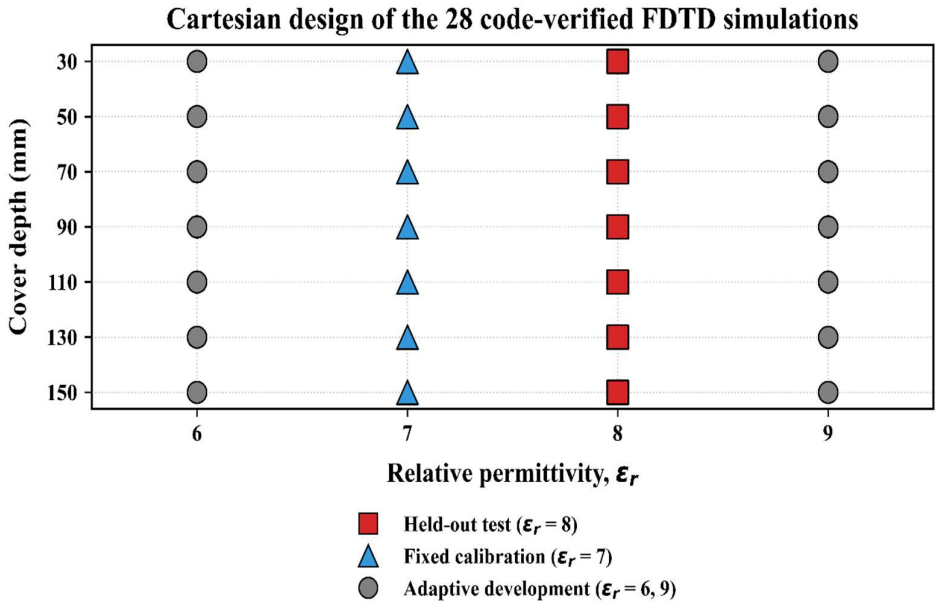

Figure 2. Cartesian design of the 28 code-verified FDTD simulations

Table 1 shows the FDTD parameterization used in the quantitative study

Table 1. FDTD parameterization used in the quantitative study

Parameter	Value
Relative permittivity, εr	6, 7, 8, 9
Cover depth	30–150 mm, 20 mm increments
Rebar diameter	16 mm (8 mm radius)
Number of simulations	28
Excitation	1.5 GHz Ricker wavelet
Spatial step	2 mm
Time step	2.35865 ps
Time steps	2400

2.3. Feature Extraction

For every simulated trace, the reinforcement response was separated out and two scalar descriptors were obtained: the rebar-echo time-of-flight (TOF) and maximum absolute peak amplitude. Since an increase in cover leads to a greater propagation distance and thus a longer arrival time, the TOF is expected to be the main depth-sensitive characteristic. The peak amplitude was kept as a secondary descriptor since both attenuation and coupling depend on the propagation conditions [3, 25]. The amplitude was not regarded as a direct measure of depth; rather, it was seen as providing information on the electromagnetic state that could be used to adjust the calibration based on TOF.

2.4. Fixed and adaptive depth inversion

Two ordinary least-squares models were evaluated. The fixed-calibration model was trained using εr = 7 and used TOF as its only predictor: $\hat{d} = a_0 + a_1 TOF$. The adaptive model used TOF and peak amplitude: $\hat{d} = a_0 + a_1 TOF + a_2 A_{peak}$. For each leave-one-permittivity-out fold, the excluded permittivity was not used to estimate regression coefficients. The models were intentionally kept low-dimensional. The purpose was to determine whether an interpretable electromagnetic-state descriptor improves robustness when deployment differs from calibration.

2.5. Uncertainty estimation

The prediction uncertainty was calculated by means of 500 bootstrap refits, using the audited random seed value of 42. In the case of the εr = 8 experiment that was reproduced, 21 of the observations made up the development set and seven were set aside. The percentile intervals were obtained from the bootstrap predictive distribution and the resulting empirical coverage was then compared with the nominal 95% interval Conformal prediction is conceptually relevant to distribution-shift problems, but no executable conformal implementation was present in the audited archive, so no conformal results are reported.

2.6. Provenance-aware literature benchmark

The GPR studies that had been published were kept in as an external layer of evidence; the original purpose of these studies, the data they were based on and the metric that had been used were maintained rather than attempting to reconstruct the unavailable raw observations. It is important to note that the published studies vary with respect to antenna frequency, acquisition geometry, condition of concrete, arrangement of reinforcement, ground-truth procedure and the definition of metric. The literature benchmark therefore provides contextual evidence about physical relevance and achievable real-data performance; it is not a pooled validation sample.

2.7. Evaluation protocol

The principal validation method is leave-one-permittivity-out (LOO-εr) evaluation. Four dielectric conditions are taken out one at a time and the other conditions are then used to fit the model. The performance is summarized using RMSE, MAE, MAPE and R^2. RMSE places more weight on larger errors, MAE gives an interpretable measure of absolute error, MAPE shows the error relative to cover depth and R^2 gives a summary of the variance that is explained. Because the folds contain only seven test observations and the training sets overlap, the comparisons are descriptive rather than formal hypothesis tests.

3. Methods

3.1. Held-out εr = 8 condition

The clearest test of adaptive inversion is obtained by excluding εr = 8 from model fitting. The fixed model, calibrated at εr = 7, gives an RMSE of 6.41 mm, MAE of 5.69 mm, MAPE of 6.10% and R^2 of 0.9743. Adding peak amplitude to TOF reduces RMSE to 3.80 mm and MAE to 3.77 mm, lowers MAPE to 5.23% and increases R^2 to 0.9910. Thus, all four reported metrics improve on the unseen εr = 8 condition. The corresponding predicted-versus-true comparison is shown in Figure 3.

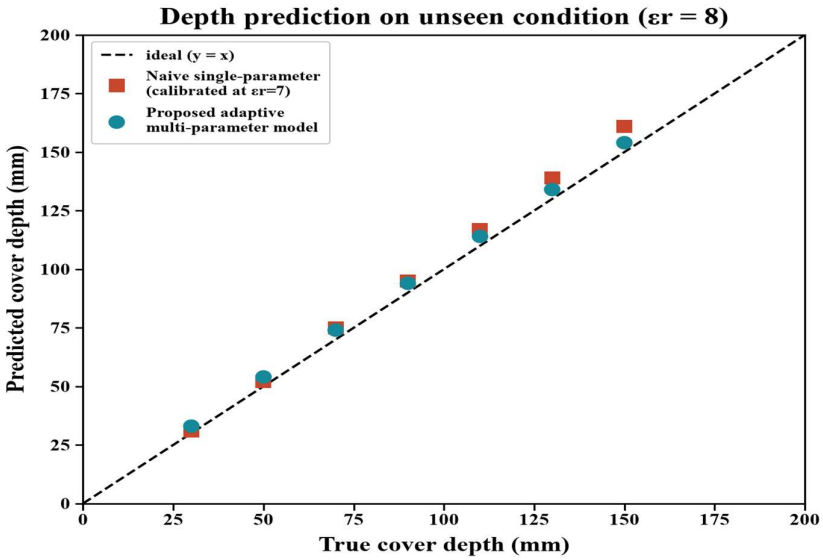

Figure 3. Predicted versus true cover depth for the fixed and adaptive models on the unseen εr = 8 condition

Table 2. Performance on the held-out εr = 8 condition

Metric	Fixed calibration	Adaptive fusion
RMSE (mm	6.41	3.80
MAE (mm)	5.69	3.77
MAPE (%)	6.10	5.23
R^2	0.9743	0.9910

Figure 4 summarizes the associated error metrics.

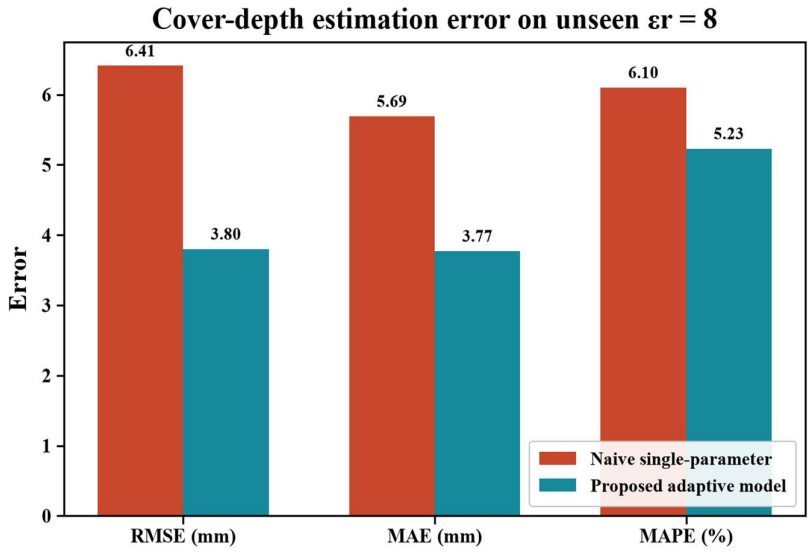

Figure 4. Error metrics for the held-out εr = 8 condition

The simulated traces show the expected physical trend: increasing cover delays the rebar reflection and reduces its peak amplitude. This behavior supports the interpretation of TOF as the principal depth-sensitive observable and amplitude as a secondary state descriptor. Figure 5 shows the simulated rebar-echo responses across the seven cover-depth conditions.

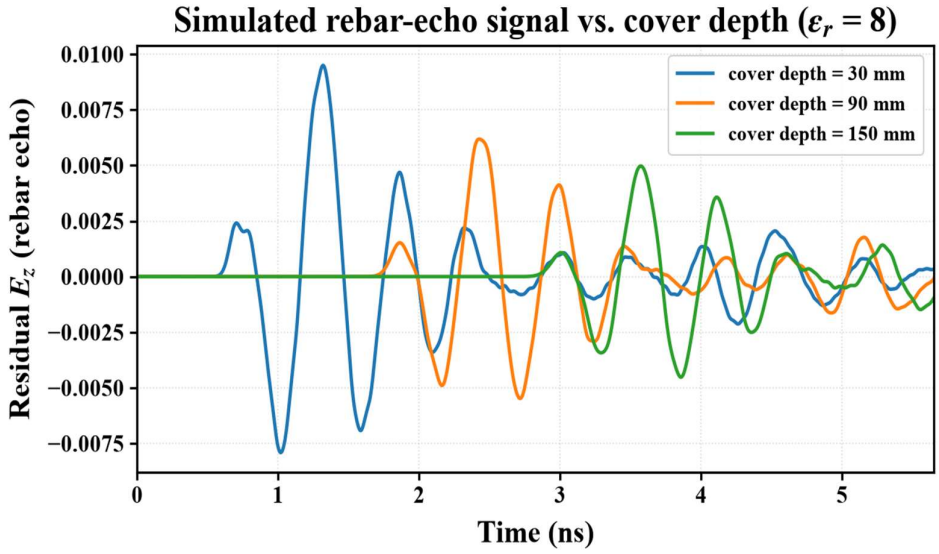

Figure 5. Simulated rebar-echo response at εr = 8 across the seven cover-depth conditions

3.2. Leave-one-permittivity-out robustness

The full LOO-εr analysis provides a stronger assessment than the single εr = 8 case. Mean RMSE decreases from 9.29 mm for the fixed calibration to 8.18 mm for adaptive fusion, corresponding to an approximately 12% reduction. The improvement is not uniform. For εr = 7, RMSE decreases from 5.97 to 4.58 mm. For εr = 8, it decreases from 6.41 to 3.80 mm. For εr = 9, it decreases from 12.62 to 12.12 mm. At εr = 6, the adaptive model changes RMSE only slightly, from 12.17 to 12.21 mm. Figure 6 summarizes the RMSE for each held-out permittivity condition.

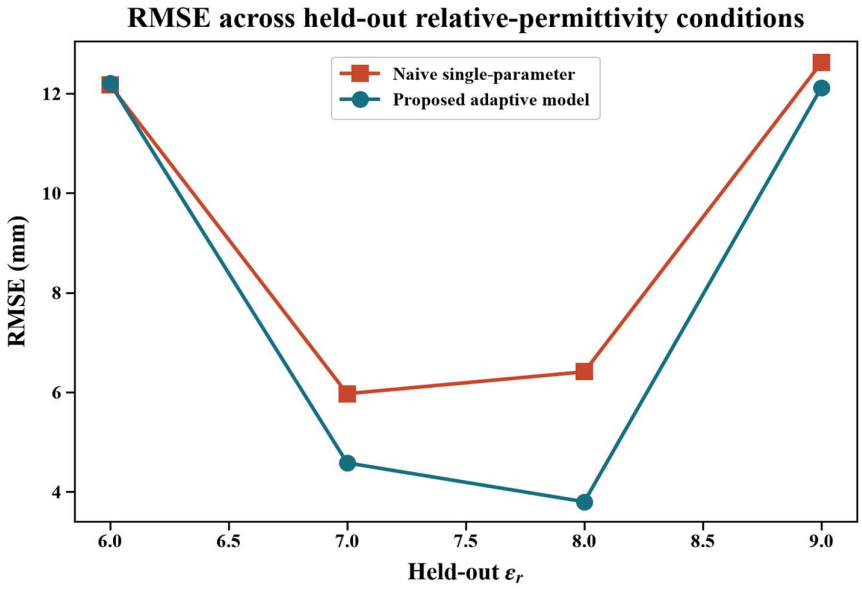

Figure 6. RMSE across held-out relative-permittivity conditions

Table 3. Leave-one-permittivity-out validation on the 28-case dataset

Held-out εr	Fixed RMSE (mm)	Adaptive RMSE (mm)	Fixed R^2	Adaptive R^2
6	12.17	12.21	0.9075	0.9068
7	5.97	4.58	0.9777	0.9869
8	6.41	3.80	0.9743	0.9910
9	12.62	12.12	0.9005	0.9082
Mean	9.29	8.18	0.9400	0.9482

The pattern shows that robustness depends on the condition. Adaptive fusion is advantageous when the excluded state is still sufficiently compatible with the electromagnetic information present in the development conditions, but it does not ensure an improvement under every boundary condition. The analysis of the errors indicates that the absolute depth error tends to increase as the cover depth increases, especially when the dielectric properties are at the values tested, which is consistent with the idea that distance from the calibration range plays a role. The overall error distributions is given in Figure 7.

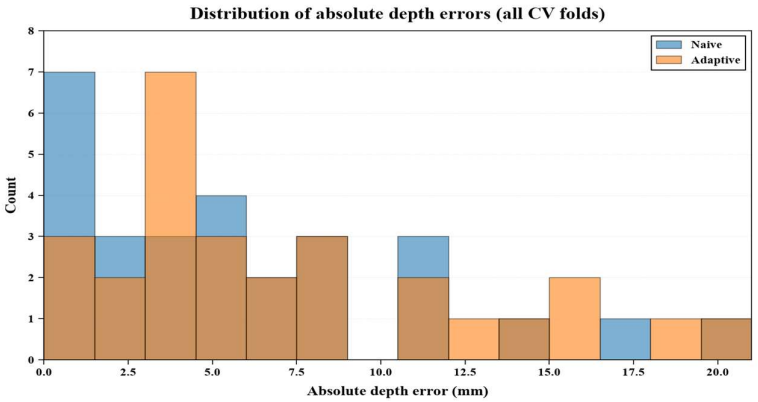

Figure 7. Distribution of absolute depth errors across the LOO-εr folds

Absolute depth error as a function of held-out relative permittivity and true cover depth is depicted in Figure 8.

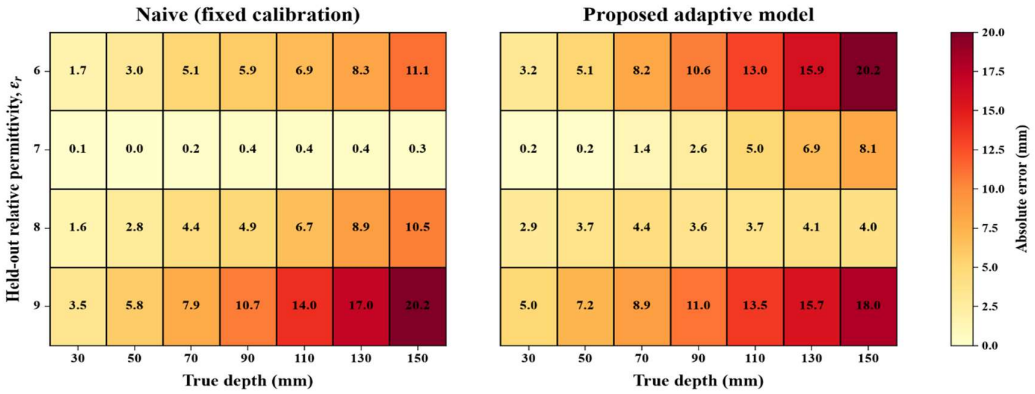

Figure 8. Absolute depth error as a function of held-out relative permittivity and true cover depth

The feature correlations show a strong relationship between TOF and true cover depth (r = 0.99), a strong inverse relationship between TOF and peak amplitude (r = −0.94), comparatively weak relationships between permittivity and the extracted features. The correlation structure is consistent with the ablation results. The corresponding Pearson correlation structure is shown in Figure 9.

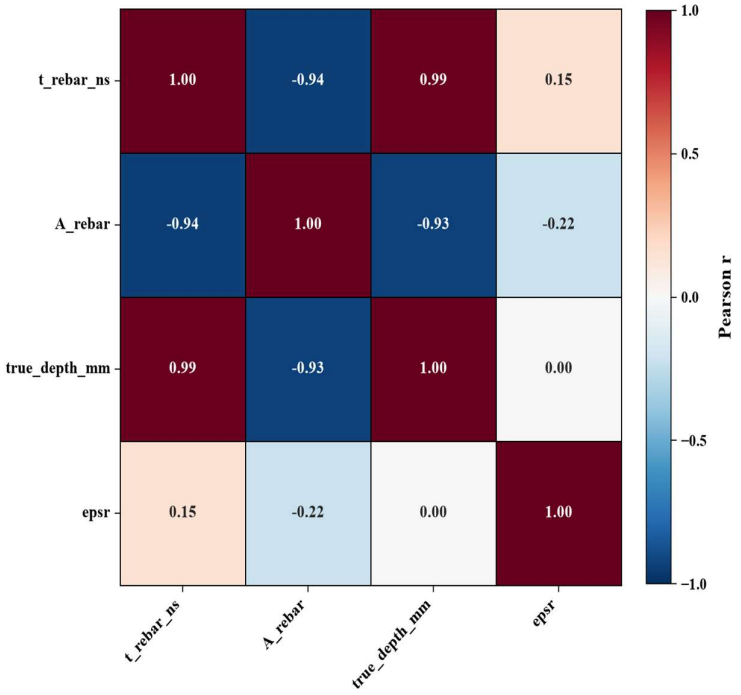

Figure 9 Pearson correlation structure among TOF, amplitude, cover depth and relative permittivity

3.3 Feature ablation

TOF alone produces a mean LOO RMSE of 8.19 mm. Peak amplitude alone produces 16.70 mm, while the combined TOF-plus-amplitude model produces 8.18 mm.

Table 4. Feature-ablation results

Feature set	Mean LOO RMSE (mm)	Interpretation
TOF only	8.19	Primary depth-sensitive observable
Amplitude only	16.70	Insufficient as a standalone depth estimator

TOF + peak amplitude	8.18	Complementary electromagnetic-state information

The tiny difference between 8.19 and 8.18 mm should not be taken as proof of a general benefit from amplitude; rather, the findings indicate that TOF already includes most of the depth information, whereas amplitude can offer useful state information in particular cases involving distribution shifts such as εr = 8. It is possible that a more comprehensive model incorporating interaction terms could make better use of amplitude, but such a model was not within the audited scope.

3.4. Uncertainty analysis

The bootstrap analysis involves 500 refits; on the seven-case condition with an εr = 8, six of the observations lie within the nominal 95% percentile prediction intervals, which corresponds to an empirical coverage of 85.7 per cent. The bootstrap prediction intervals are shown in Figure 10.

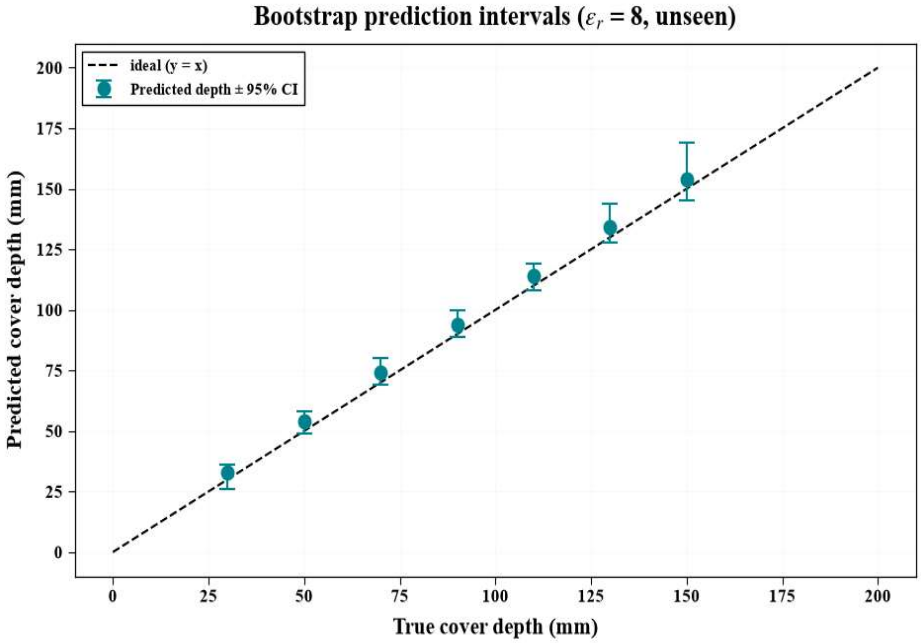

Figure 10. Bootstrap prediction intervals for the held-out εr = 8 condition

Table 5 Bootstrap uncertainty coverage

Method	Nominal level	Condition	Observed coverage
Bootstrap, 500 refits	95%	εr = 8; n = 7	85.7%

The shortfall from nominal coverage is important. It demonstrates that high point-prediction accuracy does not automatically imply calibrated uncertainty, even in a controlled numerical setting. Because only seven observations are used for this coverage assessment, the result should be regarded as an empirical diagnostic rather than a population-level calibration guarantee.

3.5. Literature-derived external evidence

The existing literature only offers contextual evidence and not pooled validation. Al-Qadi and Lahouar found an average error of about 2.6% when estimating cover depth using field GPR [1]. Liu et al. achieved a detection accuracy of 90.9% and had a laboratory depth error of less than 1.5 mm for rebar detection and localization [19]. Recent studies on real bridge decks show a mean absolute error of about 3% for automated cover-depth estimation as well as high precision and recall for rebar detection [9, 15]. Khedr et al. obtained an overall accuracy of 97.2% in the classification of rebar diameter [15]. Furthermore, simulation-to-real studies have proven successful transfer for reinforcement-diameter estimation and other GPR interpretation tasks [10, 14, 20].

The fact that these values show low error to be attainable under well-characterized acquisition conditions means that they cannot nevertheless be combined numerically with the present FDTD results since the tasks, antennas, specimens, target definitions and processing pipelines are different.

Table 6. Literature-derived empirical evidence used as an external benchmark

Study	Evidence Basis	Primary task	Reported result	Use in present study
Al-Qadi & Lahouar [3]	Field GPR	Cover-depth estimation	≈2.6% average error	External empirical context
Liu et al. [5]	Real GPR + laboratory validation	Rebar detection and depth estimation	90.9% detection accuracy; laboratory depth error <1.5 mm	External empirical context
Faris et al. [12]	Large real bridge-deck GPR dataset	Rebar detection	0.99 precision; 0.97 recall	External empirical context
Faris et al. [10]	Real bridge-deck GPR	Automated cover-depth estimation	≈3% mean absolute error	External empirical context
Khedr et al. [11]	Experimental + site GPR	Rebar-diameter Classification	97.2% overall accuracy	External empirical context

Giannakis et al. [14]	Synthetic training + real testing	Rebar-diameter estimation	Successful real-data evaluation	Precedent for simulation-to-real transfer
Hwang et al. [15]	Simulation-driven training + real structure	Rebar clutter/defect interpretation	Qualitative field verification	Precedent for simulation-first learning
GPRTransNet [16]	Synthetic numerical data + concrete experiment	Wave-to-permittivity translation	Synthetic-to-real experimental evaluation	Recent precedent for mixed evidence

4. Discussion

4.1. Interpretation of the adaptive inversion framework

The principle finding of this study is that it is possible to improve the estimation of cover depth when using a travel-time-based inversion by including the peak rebar-echo amplitude, although the extent of this improvement varies greatly with the dielectric condition. It is important to note that the aim of the proposed framework is not to show that a two-feature regression generally outperforms a model which uses only time-of-flight (TOF). Instead, the results show that a low-dimensional, physically interpretable combination of features can offer extra information whenever the electromagnetic state of the concrete is different from the calibration condition.

The $\varepsilon r = 8$ case gives the most convincing evidence. With this condition, the adaptive TOF + amplitude model lowered the RMSE from 6.41 to 3.80 mm and the MAE from 5.69 to 3.77 mm, at the same time as raising R^2 from 0.9743 to 0.9910. The fact that both the error measures and the goodness-of-fit measures improved indicates that the advantage was not due simply to a few unusual predictions. Rather, the adaptive model better represented the relationship between the measured radar response and cover depth for this particular electromagnetic state.

The behaviour in question is physically possible. Since the electromagnetic velocity in concrete is mainly determined by the relative permittivity, the two-way travel time of a reflection from a rebar will depend on both the cover thickness and the dielectric characteristics of the medium. As a result, a model that has been calibrated under one set of electromagnetic conditions can exhibit systematic errors in depth when it is used under different conditions, even if the geometry of the reinforcement does not change. The TOF feature records the main information concerning propagation delay, while the amplitude gives an extra description of how strongly the reflected energy is received.

Although amplitude should not be seen as a direct and independent measure of depth, the amplitude of the rebar reflection is influenced by propagation distance and attenuation as well as by the electromagnetic properties of the concrete. Earlier investigations into bridge decks have shown a strong relationship between depth and amplitude and have found geometric, dielectric and conductive mechanisms to be important factors in causing signal loss. Therefore, within the present framework, amplitude is more appropriately viewed as a descriptor that depends on conditions and one which can be used to supplement the time-of-flight rather than replace it.

The ablation results provide support for this interpretation. When the mean LOO-εr RMSE was calculated, it amounted to 8.19 mm for the TOF alone, 16.70 mm for the amplitude alone and 8.18 mm for the combined model. It thus follows that the amplitude-only model was considerably weaker than the TOF model, which shows that the temporal feature contains most of the direct depth information. On the other hand, the fact that the average performance of the TOF-only and TOF + amplitude models was almost identical indicates that the amplitude should not be regarded as a universally dominant second predictor; its value only becomes important under specific distribution shifts, most clearly at $\varepsilon r = 8$.

4.2 Condition-wise validation and electromagnetic distribution shift

A key methodological consequence of the results is that GPR inversion should be assessed using condition-wise rather than purely random data splitting. This is because the synthetic dataset includes repeated combinations of the same physical variables and so a random partitioning at the observation level might cause training and testing observations that come from closely related electromagnetic states to end up in both the training and testing subsets. This kind of arrangement can lead to overly optimistic estimates of generalization and may mask the real effect of dielectric variation.

The LOO-εr evaluation is a more rigorous test since the model has to predict a complete electromagnetic condition that was excluded when the coefficients were estimated. For the four different held-out conditions, the adaptive model achieved a mean RMSE of 8.18 mm whereas the fixed model obtained 9.29 mm, representing an approximate 12% decrease. Yet the individual condition results are more revealing than the average figure alone. When $\varepsilon r = 7$, the RMSE fell from 5.97 mm to 4.58 mm; when $\varepsilon r = 8$, it dropped from 6.41 mm to 3.80 mm. By contrast, the change at $\varepsilon r = 6$ was insignificant and even slightly negative, the RMSE rising from 12.17 mm to 12.21 mm. At $\varepsilon r = 9$ the improvement was also small, falling from 12.62 mm to 12.12 mm.

It shows that the method in question does not achieve uniform improvements throughout the entire electromagnetic parameter space and the greater errors when $\varepsilon r = 6$ or 9 point to the fact that extrapolation grows more difficult as you approach the ends of the dielectric range examined. This is a more useful result than simply stating the high R^2 value obtained at $\varepsilon r = 8$, since it implies that the reliability of the model is related to how close the conditions are to the calibration domain and that a prediction made for an unfamiliar electromagnetic state should not be automatically regarded as just as reliable.

This kind of behaviour is in agreement with the known physics of GPR inspection, since the wave velocity, strength of the reflection and the attenuation all depend on the electromagnetic properties of the concrete; at the same time, actual measurements can be affected by factors such as moisture, conductivity, composition of the aggregate, antenna coupling, arrangement of the reinforcement and other structural factors. The current results thus underline the importance of distinguishing between interpolation within a calibrated electromagnetic range and extrapolation beyond that range.

The fact that TOF shows a strong correlation with true depth ($r = 0.99$) once again supports the physical interpretation of the model. The negative correlation between TOF and peak amplitude ($r = -0.94$) shows that greater depth of reinforcement is linked to weaker reflections in the simulated system. Yet correlation by itself does not prove causal independence or ensure predictive value. The ablation analysis properly shows that amplitude does provide only limited independent depth information on average, even though it is useful in some shifted situations.

4.3 Relationship to experimental and field evidence

The numerical results obtained in this study should be regarded as supplementary to, not as a substitute for, experimental and field evidence. Earlier investigations have shown that GPR can give useful estimates of the position of the reinforcement and of the concrete cover, surveying studies having reported cover-depth errors of a few percent when appropriate acquisition and calibration conditions are met. Other processing methods have combined wave velocity, the amplitude of direct coupling and the depth-corrected rebar amplitude in order to improve the interpretation of concrete bridge-deck GPR data. The findings from these studies support the general idea that several GPR observables can contain complementary information.

Although this study is fundamentally different from the various experimental investigations, the dataset proposed here comprises controlled 2D FDTD simulations rather than measurements obtained from actual concrete specimens. As a result, the errors cited relate to the performance based on the electromagnetic assumptions contained in the numerical model and must not be regarded as direct measures of field accuracy.

It is especially important in the case of amplitude-based features since, in real reinforced concrete, amplitude can be influenced by moisture, chloride-related conductivity, aggregate-scale heterogeneity, cracks, the diameter and orientation of the reinforcement, antenna coupling, surface condition and the presence of neighbouring reinforcement. Studies involving depth and amplitude have demonstrated that the amplitude of the reflection is greatly dependent on depth and that several attenuation mechanisms account for the signal observed. The simulations carried out here deliberately omit many of these effects in order to isolate the impact of the shift in dielectric distribution. Although this controlled approach is useful for identifying the various mechanisms, it at the same time hampers the direct applicability in the field.

The external literature benchmark is employed merely as contextual evidence. Experimental and field studies that have already been published show that it is practically possible to estimate cover on the basis of GPR, but since there are differences in antenna characteristics, method of data acquisition, composition of the concrete, way of reinforcement arrangement, ground-truth measurements and the definitions of error, their results cannot be combined in a quantitative way with the present simulations. Such a separation of provenance is essential since agreement between the synthetic and experimental error values alone would not be sufficient to establish the validity of the synthetic results with respect to reality.

4.4. Uncertainty and reliability of the inversion

The analysis of uncertainty offers another view of the practical reliability of the proposed framework. In the case of the held-out $\varepsilon r = 8$ condition, the 500 bootstrap refits yielded prediction intervals that were nominal 95% and included six out of the seven observations, giving an empirical coverage of 85.7%. Even though this result is about as close as one might expect to the nominal target, it cannot be taken as proof of statistically valid 95% calibration since the evaluation set consists of only seven observations.

Above all, the shortfall shows that accuracy alone in terms of point estimates is not enough when it comes to engineering applications. A model may have a high R^2 value even if its uncertainty is poorly calibrated under a shifted electromagnetic condition. This is especially pertinent to GPR since the same measured signal characteristics can be affected by a number of physical mechanisms. It follows that a sound inspection framework must be able to tell the difference between the estimated cover depth and the level of confidence that should be attached to that estimate.

The findings indicate a practical order of priority when it comes to interpretation. In the case of an electromagnetic regime which is well represented by the calibration data, the predictions can be regarded as more reliable, while those made in relation to boundary or unfamiliar conditions should lead to further verification. Such verification might

consist of carrying out a local calibration with known cover measurements, comparing the results with those obtained by another non-destructive method or making direct measurements selectively when the impact of an incorrect cover estimate is high.

4.5. Engineering implications

From an engineering point of view, the main advantage of the proposed framework is therefore methodological not merely numerical. It offers a structured method in which the GPR features are first understood in terms of their physical functions, then verified on a condition-by-condition basis and an uncertainty evaluation is carried out. This approach may be seen as a calibration-and-screening procedure rather than as a full substitute for conventional inspection.

The TOF measurement should continue to be the main depth-sensitive parameter since it has the closest physical and statistical relationship to cover depth; amplitude can then be used as a supplementary feature in cases where the electromagnetic state differs from the nominal calibration condition. Nevertheless, the LOO-εr results show that the resulting estimate must be accompanied by a validity assessment and should not be presented as being universally transferable.

This kind of framework is especially applicable to the inspection of infrastructure, since the cost of making a wrong decision can be greater than the computational cost of carrying out further calibration. Although GPR is already widely used for mapping reinforcement and for assessing cover, interpretation is still sensitive to the materials involved and to the conditions under which the data were acquired. A provenance-aware model can thus help to ensure a safer deployment by clearly indicating the kind of evidence that supports a prediction and the conditions under which the model has been evaluated.

4.6. Limitations and threats to validity

It is necessary to admit several limitations. The first of these is that the dataset includes only 28 controlled simulations, each dealing with four different values of relative permittivity and seven cover depths. While this approach is adequate for the purpose of a controlled proof-of-concept study, the size of the dataset is not sufficient to allow for broad statistical generalization.

Second, the simulations employ a two-dimensional, homogeneous and lossless model of concrete. Actual concrete shows aggregate-scale heterogeneity, variations in moisture content and frequency-dependent electromagnetic properties. These factors can affect both the travel time and the signal amplitude. Third, the reinforcement is shown as a perfectly conducting cylindrical bar having a diameter of 16 mm. The study did not look into variations in bar diameter, corrosion condition, orientation or reinforcement density.

Fourth, the model does not explicitly include antenna coupling, instrumental noise, complex reinforcement cages, surface roughness or stochastic material variability. The reason for omitting these factors is that the study aims to isolate the electromagnetic distribution shift, but as a result the interpretation of the numerical accuracy as an estimate of field performance is limited.

The assessment of uncertainty relies on just seven observations that were not included in the training and the comparison with the existing literature remains of a qualitative nature. Although these limitations do not render the proposed methodology invalid, they do set the limits to the claims that can be backed up. The next important stage therefore involves carrying out experimental validation with GPR profiles that have been obtained independently from concrete specimens whose cover depths have been measured and whose electromagnetic conditions have been documented. Ideally, this validation should maintain the principle of evaluation according to condition that has been applied here in order that the performance in laboratory can be tested under actually never-before-seen concrete conditions.

5. Conclusion

A provenance-aware, physics-guided approach was presented in this study for assessing GPR concrete cover-depth inversion in the presence of an electromagnetic distribution shift. The quantitative experiment involved 28 two-dimensional FDTD simulations, these covering four relative-permittivity states and seven cover depths. When tested on the previously unseen condition of $\varepsilon r = 8$, adaptive TOF-plus-amplitude fusion lowered the RMSE from 6.41 to 3.80 mm and raised the R^2 from 0.9743 to 0.9910. For all the leave-one-permittivity-out cross-validation folds, the mean RMSE decreased from 9.29 to 8.18 mm. The results of the feature-ablation study indicate that TOF is the main depth-sensitive observable, whereas amplitude gives additional information regarding the electromagnetic state but is not adequate on its own as an estimator; the advantage of amplitude is thus dependent on the circumstances rather than being generally applicable. Moreover, the bootstrap analysis shows that accurate point estimates should not be interpreted without taking uncertainty into account, since the nominal 95% confidence intervals achieved 85.7% empirical coverage on the seven-case $\varepsilon r = 8$ test set. The literature benchmark shows that the scale of the synthetic results is physically reasonable when compared with published experimental and field studies, even though it does not serve as validation for the current model. The main contribution lies in the evaluation protocol, which includes separate provenance for the synthetic and empirical evidence, validation under dielectric shift on a condition-by-condition basis, feature ablation and the explicit reporting of uncertainty. The following important step will be to carry out paired laboratory and field validation using independently measured GPR traces and the same evaluation criteria.

Declaration of Conflict of Interests

The authors declare that there is no conflict of interest. They have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-Assisted Technologies in Manuscript Preparation

During manuscript preparation, AI tools were used for language editing, structural refinement, literature-organization assistance and document-formatting checks. The authors reviewed and edited the output and take responsibility for the accuracy, integrity, originality and final content of the manuscript.

Funding

No external funding was received for this study.

Figures

All figures in this manuscript are original author-generated diagrams and are not generative-AI images or third-party visual material.

References

[1] ASTM International (2019) ASTM D6432-19: Standard Guide for Using the Surface Ground Penetrating Radar Method for Subsurface Investigation. ASTM International https://store.astm.org/d6432-19.html.

[2] Tosti, F. and Ferrante, C. (2019) 'Using ground penetrating radar methods to investigate reinforced concrete structures', Surveys in Geophysics, 41, pp. 485–530. https://doi.org/10.1007/s10712-019-09565-5

[3] Al-Qadi, I.L. and Lahouar, S. (2005) 'Measuring rebar cover depth in rigid pavements with ground penetrating radar', Transportation Research Record, 1907, pp. 81–85. https://doi.org/10.1177/0361198105190700109.

[4] Lim, H.J. (2024) 'Application of Ground Penetrating Radars (GPR) in Concrete Bridges: A Review', Journal of the Korea Concrete Institute, 36, pp. 329–336. https://doi.org/10.4334/JKCI.2024.36.4.329.

[5] Liu, H., Lin, C., Cui, J., Fan, L., Xie, X. and Spencer, B.F. (2020) 'Detection and localization of rebar in concrete by deep learning using ground penetrating radar', Automation in Construction, 118, 103279. https://doi.org/10.1016/j.autcon.2020.103279.

[6] Rathod, H., Debeck, S., Gupta, R. and Chow, B. (2019) 'Applicability of GPR and a rebar detector to obtain rebar information of existing concrete structures', Case Studies in Construction Materials, 11, e00240. https://doi.org/10.1016/j.cscm.2019.e00240.

[7] Dinh, K., Gucunski, N. and Duong, T.H. (2018) 'An algorithm for automatic localization and detection of rebars from GPR data of concrete bridge decks', Automation in Construction, 89, pp. 292–298. https://doi.org/10.1016/j.autcon.2018.02.017.

[8] Shaw, M.R., Millard, S.G., Molyneaux, T.C.K., Taylor, M.J. and Bungey, J.H. (2005) 'Location of steel reinforcement in concrete using ground penetrating radar and neural networks', NDT & E International, 38(3), pp. 203–212. https://doi.org/10.1016/j.ndteint.2004.06.011.

[9] Li, X., Liu, H., Zhou, F., Chen, Z., Giannakis, I. and Slob, E. (2022) 'Deep learning–based nondestructive evaluation of reinforcement bars using ground-penetrating radar and electromagnetic induction data', Computer-Aided Civil and Infrastructure Engineering, 37, pp. 1834–1853. https://doi.org/10.1111/mice.12798.

[10] Faris, N., Zayed, T., Fares, A., Ebrahim, K.M.P. and Abdelkhalek, S. (2026a) 'Deep Learning-Based Automated Cover Depth Estimation for Concrete Bridge Decks Using Ground-Penetrating Radar', Transportation Research Record, 2680(6), pp. 539–556. https://doi.org/10.1177/03611981251387128.

[11] Khedr, M., Metawie, M. and Marzouk, M. (2025) 'Integrated ground penetrating radar and deep learning approach for rebar diameter classification in concrete elements', Frontiers of Structural and Civil Engineering, 19, pp. 524–540. 10.1007/s11709-025-1177-4

[12] Faris, N., Yussif, A.-M., Zayed, T., Fares, A. and Abdelkader, E.M. (2026b) 'Deep learning-based rebar detection for automated evaluation of concrete bridge decks using ground penetrating radar', Engineering Applications of Artificial Intelligence, 165, 113459. https://doi.org/10.1016/j.engappai.2025.113459.

[13] Lachowicz, J. and Rucka, M. (2018) '3-D finite-difference time-domain modelling of ground-penetrating radar for identification of rebars in complex reinforced concrete structures', Archives of Civil and Mechanical Engineering, 18, pp. 1228–1240. https://doi.org/10.1016/j.acme.2018.01.010.

[14] Giannakis, I., Giannopoulos, A. and Warren, C. (2021) 'A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar', IEEE Geoscience and Remote Sensing Letters, 18, pp. 461–465. https://doi.org/10.1109/LGRS.2020.2977505.

[15] Hwang, C., Yang, S., Kim, Y., Park, S. and Choi, H. (2025) 'Simulation-driven deep learning for rebar clutter elimination in ground- penetrating radar images to detect backfill grout defects in segment linings', Computer-Aided Civil and Infrastructure Engineering. https://doi.org/10.1111/mice.70142.

[16] Lu, G., Kou, L., Niu, P., Lv, G., Zhang, X., Liu, J. and Xie, Q. (2025) 'GPRTransNet: A deep learning-based ground-penetrating radar translation network', Tunnelling and Underground Space Technology, 161, 106557. https://doi.org/10.1016/j.tust.2025.106557.

[17] Yee, K.S. (1966) 'Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media', IEEE Transactions on Antennas and Propagation, 14, pp. 302–307. https://doi.org/10.1109/TAP.1966.1138693.

[18] Asadi, P., Gindy, M. and Alvarez, M. (2019) 'A machine learning based approach for automatic rebar detection and quantification of deterioration in concrete bridge deck ground penetrating radar B-scan images', KSCE Journal of Civil Engineering, 23, pp. 2618–2627. https://doi.org/10.1007/s12205-019-2012-z.

[19] Efron, B. and Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Chapman & Hall. https://doi.org/10.1201/9780429246593.

[20] Angelopoulos, A.N. and Bates, S. (2023) 'Conformal Prediction: A Gentle Introduction', Foundations and Trends in Machine Learning, 16, pp. 494–591. https://doi.org/10.1561/2200000101.

[21] Ovadia, Y., Fertig, E., Ren, J., Nado, Z., Sculley, D., Nowozin, S., Dillon, J.V., Lakshminarayanan, B. and Snoek, J. (2019) 'Can You Trust Your Model's Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift', Advances in Neural Information Processing Systems, 32, pp. 13991–14002 10.48550/arXiv.1906.02530.