Introduction

The use of Ground Penetrating Radar (GPR) as an investigative method for near surface geophysical problems, and for numerous engineering applications, is well established (Daniels, 2004). However, interpretation of complex GPR data is often not a simple process, primarily due to the relevant proximity of the GPR transducers to the air-earth interface and often to the target they sense. In addition, background media inhomogeneities, as well as noise and clutter add to the problem. To aid interpretation, numerical modelling is almost the only viable option for realistic simulation of GPR signals, especially when considering challenging near surface problems. Unfortunately, if simulated responses are to be directly compared to real GPR data the full-wave 3D nature of the electromagnetic problem must be considered. The most popular numerical method for computational electrodynamics, that is regularly applied to the GPR forward problem, is the Finite-Difference Time-Domain (FDTD) (Yee, 1966; Taflove and Hagness, 2000). gprMax is open source software that utilises the FDTD method, and has been specifically developed for GPR (Warren et al., 2016). To be useful for GPR applications a GPR model needs to predict as closely as possible what a real GPR response will be. Unfortunately a common, and critical, characteristic of FDTD GPR modelling is the demand for significant computational resources. As a result, the iterative use of such GPR models either by trial and error, to assist interpretation, or as part of an automatic inversion scheme, is still a rather laborious process requiring significant time and resources. It is thus difficult to incorporate it in everyday end-user GPR applications.

Here we present some preliminary results of the application of machine learning (ML) concepts to speed up the process of forward modelling of GPR. The basic premise is to construct an accurate and very fast GPR forward solver by successfully training a neural network that can then be used to quickly provide GPR modelled responses based on parametrised inputs. It is important to recognise that such an approach cannot solve the general forward problem of GPR, but can be applied to specific subsets that can be easily parametrised and constrained in variability. Although this might appear at first to be too restrictive, in practice there are a number of geophysical and engineering applications that are suitable candidates.

The machine learning training process is computationally demanding, however the final trained model can run in practically real-time, as it is based on simple inference of the final trained neural network. To verify the suitability of such an approach the fast forward ML based GPR model is compared to real full 3D FDTD models, and is used in full waveform inversion (FWI) of real GPR data for rebar detection in a concrete slab.

Theory

We recently reported (Giannakis et al., 2018) the application of a ML approach to accurately predict the direct coupling response of a GPR antenna having variable height above a rough terrain with soil parameters that can be described by the Peplinski model (Peplinski et al., 1995). This approach worked well in predicting the direct coupling response of the GPR antenna but did not predict as accurately the complex scattering from targets in close proximity to the interface. A different approach however has been found to be promising and successful in allowing for training a ML system with good final accuracy and fidelity in its predictive capability, even considering complex interactions and near surface targets. As an example, the problem of predicting the A-Scan response of a single rebar in concrete (Fig.1) is used. The 3D FDTD simulation, using gprMax, utilises an accurate model of our own GSSI 1.5GHz antenna (Giannakis et al., under review) that has been used to obtain real GPR data from such targets to verify the model. The dielectric properties of concrete are described by an extended Debye model and the properties were calculated using experimental measurements (Bourdi et al., 2012; Soutsos et al., 2001; Shaari et al., 2004) resulting in a model for concrete with electrical properties only related to its water fraction. A spline interpolation is used to evaluate the parameters for any water fraction between 0.2%–12%. Describing the dielectric properties of concrete using only the water fraction is particularly useful for FWI since there is only one unknown parameter to be estimated instead of three.

The first part of the new approach is to parametrise the inputs for generating 3D gprMax GPR models.

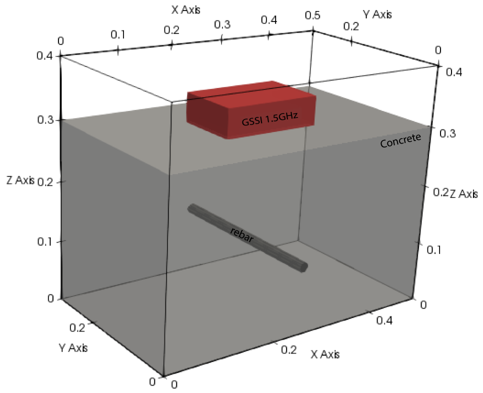

***Figure 1** 3D FDTD gprMax model used to generate GPR A-Scans for training the ML-based model.*

These were: R the radius of the rebar, D its cover depth and WC the water fraction of the concrete material used to determine its electrical properties. The training approach has two key steps: firstly, it utilizes principal component analysis (PCA); and secondly, it employs a realistic and comprehensive training dataset obtained using detailed GPR models based on our 3D FDTD gprMax solver. PCA has been used in GPR to eliminate direct antenna coupling and remove the ground reflection (Tebchrany et al., 2014; Kaczmarek and Pietrasinski, 2014). However, the main purpose of PCA is to compress and reduce the dimensionality of data (Bengio et al., 2013). We employ PCA to reduce our A-Scan size to only 40 principal components and use them in a predictive manner for training our neural network.

Our proposed neural architecture is divided into two sections, with each section further divided into 40 steps. In the first step of the first section the inputs of the model (R, D, WC) are used in order to predict the first principal axis. The architecture of the neural network for this step is based on a two hidden layers model using 30 and 10 neurones respectively. The activation functions of the neurones are sigmoid apart from the output which is linear. The second step uses the model's inputs plus the predicted value of the first principal components in order to predict the second one. Subsequently, the model's inputs plus the two predicted principal components are used to predict the third one. The procedure is repeated until all the principal components are predicted.

The first section provides a full set of predicted principal components. However, these contain errors due to inaccuracies in the predictions. Hence, a second section is introduced attempting to establish a causal relationship between the errors in the predicted values with respect to the actual principal axes and the parameters of the model. In the first step of the second section, the parameters of the model plus all the predicted principal components, apart from the first one, are used in order to predict a revised first principal axis. During the second step, the parameters of the model and the revised first component together with all the predicted components from the first section apart from the second one are used to predict the revised second principal component. The same procedure is repeated until all components have been successfully revised. Each step of the second section is trained individually using as inputs the parameters of the model and the outputs of the first section and the outputs of the previous steps of the second section. In each step the data set is divided in different training, validation, and testing sets to reduce the possibility of over-fitting.

The final neural network architecture including the first section and the second section with all their steps, consists of 240 layers. Although generating the training set and tuning the suggested neural network is computationally intensive, the final output is a near real-time (≈ 1 sec) estimator of the 40 principal components that can be decompressed to give the predicted time-domain modelled GPR A-Scan. Notice that the proposed neural network uses as inputs the predicted principal components and not the actual ones. The only inputs required from the user are therefore the values for the water content of the concrete, and the radius and cover depth of the rebar.

A test for the accuracy of the ML-based forward solver is presented in Fig. 2. The results are almost identical demonstrating that the proposed neural network architecture can resolve and predict the underlying complex pattern between the given inputs and their corresponding A-Scans.

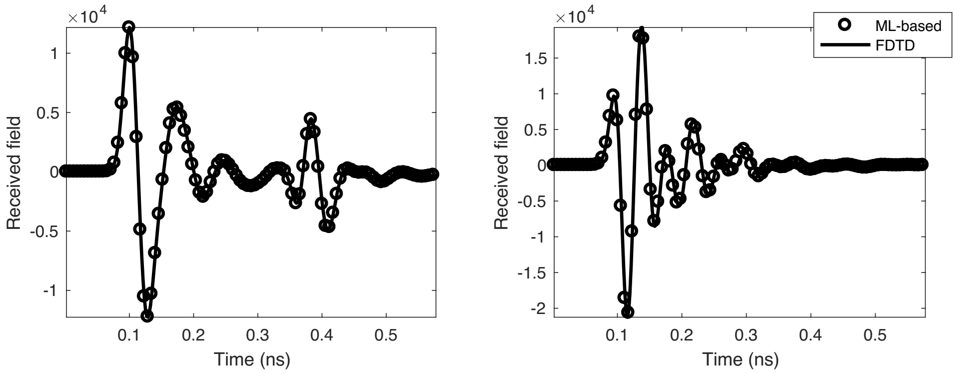

Figure 2 *Validation of the ML-based forward model. Comparison of A-Scans obtained by inference from the trained neural network and from the full 3D FDTD gprMax forward model.*

Application of the ML-based fast forward solver to FWI

The ML-based GPR forward solver has been used within a FWI process to estimate the cover depth, radius of rebar and the moisture content of the concrete. In our FWI a simulated annealing optimization (Kirkpatrick et al., 1983) is used in order to minimize the average mean absolute error between the ML-based modelled and the objective real A-Scan traces. A global optimizer can constrain the model to exist within given boundaries, which is particularly useful since our ML-based forward solver has been trained for a specific set of cases ($R = 2 - 25$ mm, $D = 0 - 300$ mm, $WC = 0.2 - 12$ %). Estimating A-Scans for values outside these predefined boundaries, will result in unreliable extrapolations.

Real A-Scans were collected on a test slab in the Non-Destructive Testing Laboratory at the School of Engineering, The University of Edinburgh. The measurements were taken using a GSSI 1.5 GHz antenna over four different steel rebars in a well-cured concrete slab ($>$ 3 years). The measurements were made directly above each rebar, with the polarization of the antenna parallel to the longitudinal rebar axis. The GPR antenna setup was according to Giannakis et al. (under review). The water fraction estimated from the FWI was approximately 11.5 %. The estimated and the actual depths and radii of the rebar are shown in Fig. 3. The predicted and the actual rebar characteristics are in very good agreement indicating the usefulness and potential of the proposed methodology for specific field applications.

Conclusions

A ML-based forward solver is capable of predicting practically in real-time the full GPR A-Scan response for a suitably constrained GPR scenario. Realistic full-wave 3D modelling is required to create suitable training data as creating comprehensive data sets using real data is not as practical to ensure suitable training data variability. Having a very fast and accurate forward GPR model is a significant step in the process of applying FWI to GPR data in practical time-frames that are useful to GPR end-users.

References

Bengio, Y., Courville, A. and Vincent, P. [2013] Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**(8), 1798–1828.

Bourdi, T., Rhazi, J.E., Boone, F. and Ballivy, G. [2012] Modelling dielectric-constant values of concrete: an aid to shielding effectiveness prediction and ground-penetrating radar wave technique interpretation. *Journal of Physics D: Applied Physics*, **45**, 1–12.

Daniels, D.J. [2004] *Ground Penetrating Radar*. Institute of Engineering and Technology, 2nd edn.

Giannakis, I., Giannopoulos, A. and Warren, C. [2018] A machine learning approach for simulating

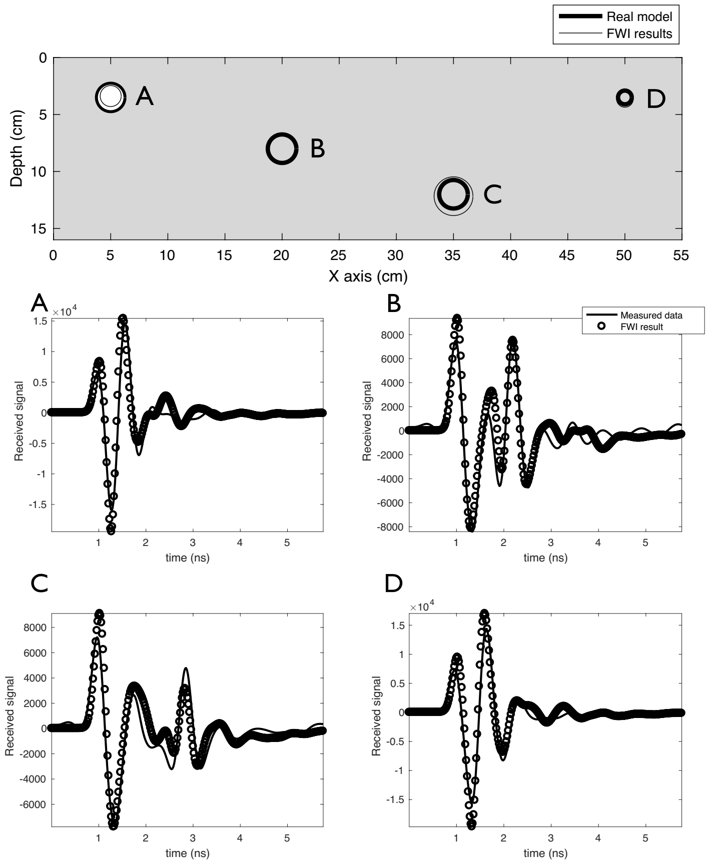

Figure 3 *FWI results for real measurements over four rebar with different radii and depths. The recovered water content of the concrete is approximately* 11.5%.

ground penetrating radar. In: *Proceedings of the 17th International Conference on Ground Penetrating Radar*.

Giannakis, I., Giannopoulos, A. and Warren, C. [under review] Realistic FDTD GPR antenna models optimised using a novel linear/non-linear full waveform inversion. *IEEE Transaction on Geoscience and Remote Sensing*.

Kaczmarek, P. and Pietrasinski, J. [2014] Principal component analysis in interpretation of A-Scan measurements in GPR system. In: *Proceedings of the 15th International Radar Symposium (IRS)*. 1–5.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. [1983] Optimization by simulated annealing. *Science*, **220**, 671–680.

Peplinski, N.R., Ulaby, F.T. and Dobson, M.C. [1995] Corrections to dielectric properties of soils in the 0.3âĂŞ1.3-GHz range. *IEEE Transactions on Geoscience and Remote Sensing*, **33**(6), 1340.

Shaari, A., Millard, S.G. and Bungey, J.H. [2004] Modelling the propagation of a radar signal through concrete as a low-pass filter. *NDT&E International*, **37**, 237–242.

Soutsos, M.N., Bungey, J.H., Millard, S.G., Shaw, M.R. and Patterson, A. [2001] Dielectric properties of concrete and their influence on radar testing. *NDT&E International*, **37**, 237–242.

Taflove, A. and Hagness, S.C. [2000] *Computational Electrodynamics, the Finite-Difference Time-Domain Method*. Artech House, 2nd edn.

Tebchrany, E., Sagnard, F., Baltazart, V., Tarel, J.P. and Derobert, X. [2014] Assessment of statistical-based clutter reduction techniques on ground-coupled GPR data for the detection of buried objects in soils. In: *Proceedings of the 15th International Conference on Ground Penetrating Radar*. 604 – 609.

Warren, C., Giannopoulos, A. and Giannakis, I. [2016] gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. *Computer Physics Communications*, **209**, 163 – 170.

Yee, K. [1966] Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, **14**(3), 302–307.