

Geoscanning an active construction site: Integrating handheld TEM models with geotechnical logs through machine learning

C.W. Christensen¹, G.H. Skurdal¹, F. Pacheco¹, P. Rodrigues¹, A. Pfaffhuber¹, K. Kawano²

¹ EMerald Geomodelling; ² Kajima Corporation

Summary

We present a Proof of Concept combining resistivity models from a compact time-domain electromagnetic (TEM) scanner with supervised machine-learning interface workflow, deployed at an active construction site in Yokohama, Japan. In a single field day, two operators acquired walking TEM data across six sub-areas between sheet piles, vehicles and buried utilities. The resulting quasi-3D resistivity model was interpreted with an ensemble of 20–30 artificial neural networks trained on two on-site boreholes and manually picked points from the resistivity data, predicting both the top and base of a soft volcanic clay unit together with quantified uncertainty. Resulting interfaces match the borehole lithology closely (mean depths 7.2 m and 16 m, uncertainties 2.3 m and 3.1 m), and in one case prompted a refinement of the operator's original geotechnical pick. The workflow delivers continuous, uncertainty-quantified 3D ground models from sparse drilling, makes its own limits explicit through the uncertainty maps, and points towards a scalable template for engineering geophysics on infrastructure-dense sites.

Geoscanning an active construction site: Integrating handheld TEM models with geotechnical logs through machine learning

Introduction

Geotechnical site characterisation on active construction sites is still dominated by point-based drilling: accurate at the borehole, but spatially sparse, expensive, and disruptive to ongoing works. Near-surface electromagnetic (EM) methods can fill the gaps between drillings, but historically required point-based, vehicle- or airborne platforms that are difficult to deploy on a congested live site. A new generation of compact, walking time-domain EM (TEM) scanners now enables data acquisition between machinery, stockpiles, and temporary works—provided that the surrounding workflow can both cope with cultural noise and convert sparse soundings into decision-ready ground models.

In this paper, we present results from Phase 1 of a Proof of Concept (PoC) carried out for Kajima Corporation at an active construction site in Yokohama, Japan. Data from a handheld TEM system was combined with sparse geotechnical drillings through a supervised machine-learning (ML) interface modelling workflow to deliver a 3D, uncertainty-quantified model of a soft-clay unit critical for foundation design. The objective was threefold: (1) test whether a handheld TEM scanner can be operated effectively in an infrastructure-dense environment, (2) evaluate the quality of the resulting resistivity models against borehole observations, and (3) demonstrate how an ML-based interface workflow can propagate sparse drilling information into continuous, uncertainty-aware geological surfaces.

Method

The site is characterised by soft volcanic clays and clay tuff (a unit named Dc1) overlying coarser silts, sands and gravels. The primary survey target is the delineation of the Dc1 clay unit, which controls foundation design. Four boreholes were available within the main EM survey footprint (specifically Area 1, Figure 1), though two were shallow drillings that only intersected soils that were completely excavated before the AEM survey. Additional logs were performed near but outside other survey areas.

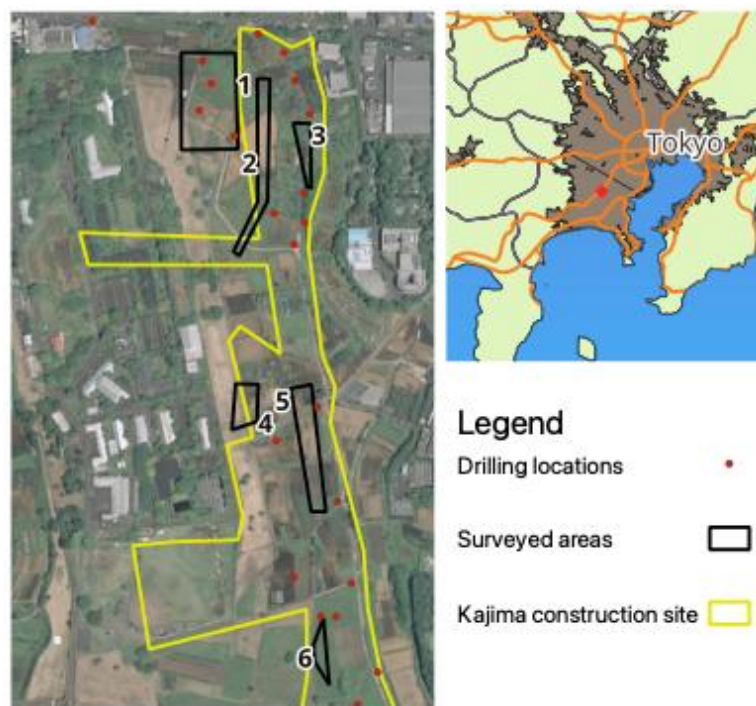


Figure 1 Survey layout at parts of the Yokohama construction site. Six sub-areas (orange polygons, five visible on this map) were acquired with the handheld TEM scanner around active works; boreholes (red dots) provide direct training data for the ANN interface workflow. Basemap imagery adapted from “Seamless Nationwide Latest Photo”, Geospatial Information Authority of Japan.

Data were acquired in a single field day with a backpack-mounted TEM system (TEMCompany's TEM2Go). The instrument consists of a wireless-coupled transmitter and receiver carried by two operators with a fixed 15 m separation, monitored in real time on a tablet app to a 10 cm tolerance. Data are streamed continuously while walking, allowing on-site quality control of decay curves and immediate replanning of survey lines when interference is detected. Six smaller sub-areas were ultimately surveyed within the construction footprint after the originally proposed larger blocks proved inaccessible due to active works, sheet piles, vehicles, fences and buried utilities. Line spacing was kept close to the nominal 20 m where space allowed.

After acquisition, the data were stacked and screened to flag most soundings affected by metallic infrastructure (so-called cultural noise). Such couplings, if retained, typically appear in inversion results as unrealistically conductive layers at depth. The remaining soundings were inverted with a laterally constrained 1D scheme to produce a quasi-3D resistivity volume for each sub-area.

The base of the Dc1 clay is interpreted using a supervised-ML interface modelling workflow. We use an ensemble of 20–30 artificial neural networks (ANNs), each trained on a bootstrap (80%) of the available training data with 20% retained as validation data. Training labels are interface depths from boreholes, supplemented where coverage is sparse by manually picked points extracted from the resistivity model and assigned a 15% depth uncertainty. Features are the per-layer resistivities of the inverted model; only the layers reaching the relevant depth range are kept. The ensemble predicts both a mean depth-to-interface and the standard deviation of the predictions, which provides a per-pixel uncertainty map propagated directly from data variability and from the spread between the trained networks (Lysdahl et al., 2022). An additional post-processing step is applied wherein local adjustments are made near training to ensure an exact match between predicted depth and drillings. The resulting interface surfaces are continuous across the survey area, and explicitly communicate where the prediction is robust and where additional ground truth would help most.

Results

In sub-areas free of significant cultural infrastructure (notably Areas 1, 4 and 6), the inverted resistivity sections show a clear three-layer signature: a moderately resistive near-surface loam, a conductive middle layer corresponding to the Dc1 clay, and a more resistive sand and gravel base (Figure 2). Two stratigraphic interfaces could be identified consistently: the top of the Dc1 clay, marking the loam–clay transition, and the base of the Dc1 clay, marking the transition to the underlying coarser deposits.

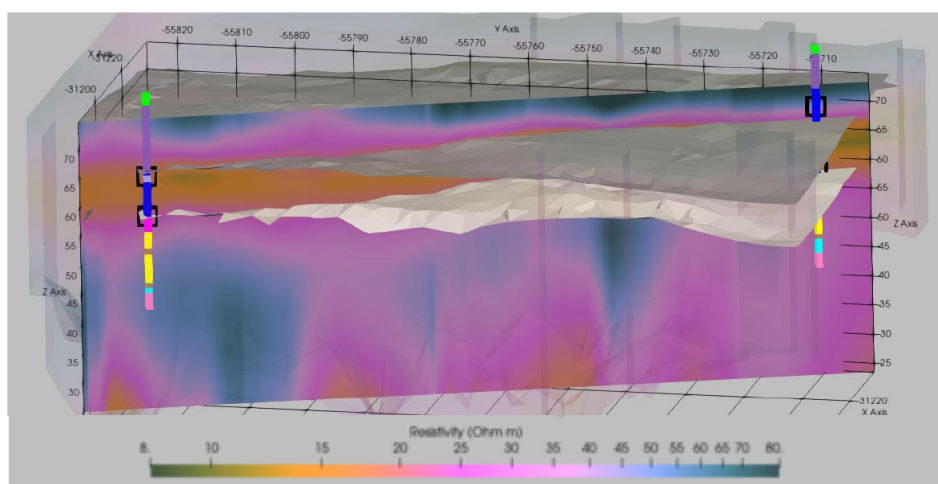


Figure 2 3D view of the inverted resistivity model at Area 1 with boreholes A-3 (left) and A-4 (right and intersected with ANN-derived top-of-clay and base-of-clay surfaces.). The conductive Dc1 clay (orange/red layer) is well resolved between resistive loam and basal sand/gravel.

Where the two boreholes A-3 (Figure 2, left) and A-4 (Figure 2, right) fall inside the survey footprint, the agreement between resistivity model and lithology is very good. At borehole A-3 the resistivity transitions track the loam-to-clay contact at 13.5–15.5 m depth and the base of clay at 21.3 m depth. Borehole A-4 was particularly instructive: the resistivity model placed the top of clay deeper than the operator's original pick. Closer inspection of the borehole log revealed coarser, more resistive material (pumice, scoria, carbonitic fragments) interbedded with the clay between 7.65 and 9 m depth, supported by elevated n-SPT values down to 12.2 m. The resistivity model thus prompted a refinement of the geotechnical interpretation rather than the other way around—a tangible illustration of the value of integrating geophysics with drilling.

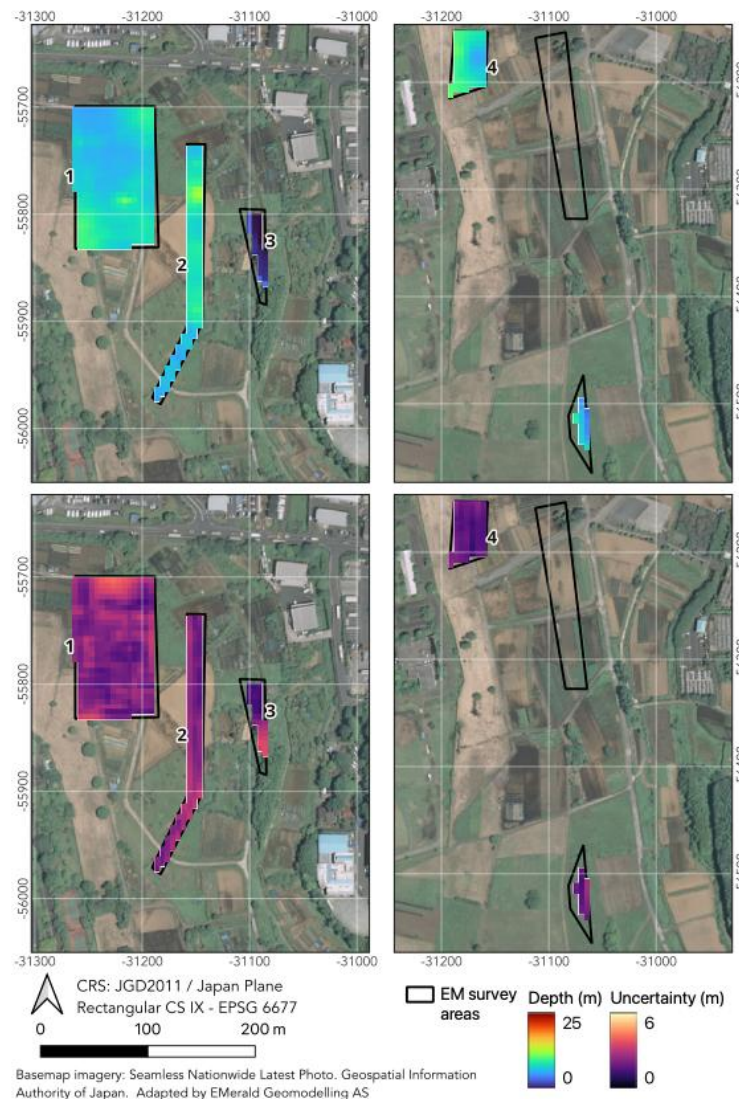


Figure 3 ANN interface model for the top of the Dc1 clay, shown for the five interpretable sub-areas. Upper maps: predicted depth below surface (mean 7.2 m). Lower maps: associated uncertainty as the ensemble standard deviation (mean 2.3 m). Basemap imagery adapted from “Seamless Nationwide Latest Photo”, Geospatial Information Authority of Japan.

The ANN ensemble was trained on these borehole picks together with manually interpreted points from Areas 1, 2, 4 and 6 (top of clay) and Areas 1, 4 and 6 (base of clay). Zero-thickness training points were imposed in Area 3, where field observations and surrounding boreholes confirm that the loam unit had been completely removed by recent excavation. The resulting interface models are stable across the surveyed sub-areas (Figure 3), with mean depths of 7.2 m for the top of clay and 16 m for the base, and

average ensemble uncertainties of 2.3 m and 3.1 m respectively. The notable exceptions are physically meaningful: the top of clay coincides with the present-day terrain in the excavated Area 3, while the base of clay is somewhat deeper in Area 4 on the embankment fill.

The survey also defined the operational envelope of the method. Where sheet piles, parked vehicles, metal fences or buried utilities sat within roughly ten metres of the survey line (Areas 2, 3 and 5), the deeper part of the resistivity response is dominated by inductive coupling to those structures, manifesting as anomalously low resistivities at depth that are not reflective of the geology. The shallow response in Areas 2 and 3 nevertheless remains usable and matches the loam thicknesses observed in surrounding boreholes, allowing the top-of-clay interface to be modelled even where the base of clay cannot. Area 5 is too strongly influenced for any reliable geological interpretation.

Discussion and Conclusions

The Phase 1 PoC demonstrates that a handheld TEM scanner, combined with a supervised ML interface workflow, can deliver continuous and uncertainty-quantified 3D ground models for an active construction site—something that point-based drilling alone cannot achieve at comparable cost or speed. With only two co-located boreholes, an interface model spanning the entire survey area was produced, validated against drilling, and used to highlight a previously unidentified deep heterogeneity. Equally important, the same workflow makes its own limitations visible: weaker resistivity contrasts or poorer resolution with increasing depth both translate into elevated ensemble uncertainty rather than into hidden bias.

The natural role of such a workflow is not to replace boreholes, but to optimise where they go. By delivering an early, low-disturbance image of the subsurface together with quantitative uncertainty, the method allows geotechnical engineers to focus drilling on the locations where it most reduces risk, and to integrate every new borehole back into the model as additional training data. Phase 2 of the project will scale acquisition coverage, integrate the volumetric ML classification workflow that estimates lithology probabilities throughout the 3D volume (Christensen et al., 2021), and feed the resulting models directly into the end user's engineering workflows. We see this combination of compact field hardware and ML-driven interpretation as a generalisable template for engineering geophysics on infrastructure-dense sites.

Acknowledgements

We thank Kajima Corporation for permission to publish, providing site access, borehole data and logistical support, and our partner TEMCompany for technical assistance during field operations. We are grateful to colleagues at EMerald Geomodelling for fruitful discussions on ML workflow design and uncertainty quantification.

References

- Christensen, C.E., Harrison, E.J., Pfaffhuber, A.A. and Lund, A.K. [2021] A machine learning-based approach to regional-scale mapping of sensitive glaciomarine clay combining airborne electromagnetics and geotechnical data. *Near Surface Geophysics*, 19(5), 523–539.
- Lysdahl, A.K., Christensen, C.W., Pfaffhuber, A.A., Vöge, M., Andresen, L., Skurdal, G.H. and Panzner, M. [2022] Integrated bedrock model combining airborne geophysics and sparse drillings based on an artificial neural network. *Engineering Geology*, 297, 106484.