https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Integrating Ground Penetrating Radar and machine learning for assessment of lake bed permeability and potential vertical-water-loss zones in shallow lake under climatic stress

Diaa Sheishah[1,2]*, Enas Abdelsamei[1,2], Ahmed Mohsen[3,4], Viktória Blanka Végi[1], Péter Kozák[5], Károly Fiala[5], Ferenc Kovács[1], Izabella Babcsányi[1], György Sipos[1]*

[1]University of Szeged, Department of Geoinformatics, Physical and Environmental Geography, 6722 Szeged, Egyetem u. 2-6., Hungary
geo_diaa@nriag.sci.eg, enas.mohammed@nriag.sci.eg, blankav@geo.u-szeged.hu , kovacsf@geo.u-szeged.hu , i.babcsanyi@gmail.com , gysipos@geo.u-szeged.hu
[2]National Research Institute of Astronomy and Geophysics, 11421, El Marsad st., Helwan, Cairo, Egypt
geo_diaa@nriag.sci.eg , enas.mohammed@nriag.sci.eg
[3]Department of Hydraulic and Water Resources Engineering, Faculty of Civil Engineering, Budapest University of Technology and Economics, Muegyetem Rakpart 3, H-1111 Budapest, Hungary : ahmed_mohsen250@f-eng.tanta.edu.eg
[4]Department of Irrigation and Hydraulics Engineering, Tanta University, Tanta 31512, Egypt
[5]Lower Tisza District Water Directorate, Stefania 4., 6720 Szeged, Hungary
kozakp@ativizig.hu , fialak@ativizig.hu

Corresponding authors: György Sipos, email: gysipos@geo.u-szeged.hu ,
Diaa Sheishah, email: geo_diaa@nriag.sci.eg

Abstract

Climate change and increasing anthropogenic pressures have intensified the vulnerability of inland water bodies, altering their hydrological balances, reducing their water levels, and degrading their water quality. One critical issue in this context is the limited understanding of lake bed hydrogeology, particularly the extent to which sediments hinder (as aquitards) or permit subsurface leakage. Although sediment sampling provides valuable point-based information, its spatial coverage is limited, emphasizing the need for high-resolution, lake-wide geophysical methods. This study determined whether the bed of Lake Vadkerti, a shallow lake experiencing persistent water level decline, facilitates vertical water loss. An integrated method combining ground-penetrating radar (GPR) and sediment sampling was used to evaluate subsurface sediment structures. A dense grid of GPR profiles was collected, enabling 2D profile interpretation and 3D time-slice visualization. Amplitude polarity, reflector geometry, and attenuation modeling were applied to identify stratified sedimentary layers. The resulting aquitard zoning map revealed heterogeneous lake bed conditions: low-permeability aquitards dominate the central and southern areas, whereas higher-permeability non-aquitards appear along the northeastern and central-western margins,

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

indicating potential zones of groundwater interaction. The performance of four machine learning models—K-nearest neighbors, random forest, extra trees, and gradient boosting—in classifying aquitard zones based on GPR amplitude features was evaluated. The extra trees model demonstrated the most balanced performance across all classes and stronger generalization, with 97% accuracy and high recall across all classes (aquitard: 100%, leaky aquitard: 86%, non-aquitard: 79%). Moreover, its spatial predictions were consistent with observed hydrostratigraphic patterns. This approach provides a comprehensive framework for understanding the hydrological functioning of lake beds and informing sustainable water management in climatically sensitive freshwater systems.

Keywords: Ground-Penetrating Radar, Aquitard characterization, Machine learning, Sediment sampling, Python scripts, Attenuation loss model

1 Introduction

Climate change, often exacerbated by human activities (e.g., increased evaporation and decreased runoff, overexploitation of surface and groundwater resources, and inappropriate drainage practices), puts significant pressure on climate-sensitive saline lakes and wetlands, leading to a decrease in the water volume of the lakes. The proper management of these highly vulnerable ecosystems requires an understanding of the lake's complex hydrological, ecological, and geochemical systems (Pitchford et al., 2012; Wurtsbaugh et al., 2017; Ladányi et al., 2023). While surface data help assess current ecological and hydrological conditions, subsurface studies are key to understanding sediment behavior, groundwater interactions, and morphological changes. Numerous studies have illustrated the impact of seepage on 1) lake water quality, such as groundwater abundant in cations (Dean et al., 2003; Cullmann et al., 2006; Klenk et al., 2025), dissolved inorganic and organic carbon (Striegl and Michmerhuizen, 1998; Staehr et al., 2010), the influx of nutrient-laden groundwater (Loeb and Goldman, 1979; Belanger et al., 1985; Ito et al., 2007), or, more broadly, alterations in lake alkalinity due to weathering processes within the watershed (Schafran and Driscoll, 1993); and 2) biological communities, including biodiversity and species distribution in seepage zones (Hagerthey and Kerfoot, 1998; Hayashi and Rosenberry, 2002; Sebestyen and Schneider, 2004). Additionally, the properties of aquitards beneath lakebeds, particularly the distribution of low-permeability materials such as clay, play a crucial role in regulating vertical exchanges between groundwater and surface water. Ground-penetrating radar (GPR) provides a high-resolution method for identifying subsurface features, including clay-rich barriers and more permeable zones, such as sand lenses or fractures.

GPR is a highly effective tool for high-resolution imaging of subsurface features in lake environments. It reveals sedimentary layers, submerged landforms, and human-induced changes to the lakebed (Neal, 2004; Bristow & Jol, 2003). Applications range from mapping bathymetry and submerged objects to reconstructing pale landscapes and identifying

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

ancient shorelines (Ayolabi & Akinwale, 2022; Corradini et al., 2020; Eli & López, 2022). While 2D profiling remains common, advances in 3D imaging and time-slice visualization have significantly enhanced spatial resolution and interpretation, thereby expanding the role of GPR in sedimentology, archaeology, and environmental studies. GPR works by sending electromagnetic waves into the ground and detecting reflections from material boundaries with differing dielectric properties. Its performance in lakes depends on factors like water depth, sediment type, and signal attenuation (Neal, 2004). Antenna frequency is key—high frequencies (250–900 MHz) offer detailed imaging, while lower ones (50–100 MHz) penetrate deeper with less resolution (Annan, 2009). Although deep water can weaken signals, GPR has proven effective in shallow lakes, where it can penetrate both water and sediment to map lakebed structures and stratigraphy.

Recent studies affirm GPR's versatility in lacustrine environments. Ayolabi and Akinwale (2022) utilized 80–160 MHz antennas to detect submerged objects in MTU Lake, identifying metallic drums and distinguishing debris types based on their impedance. While effective for object detection, the study's narrow scope suggests the value of broader, multi-technique approaches in varied conditions. Similarly, Corradini et al. (2020) combined GPR and coring to reconstruct Mesolithic settlements at Lake Duvensee, offering insight into past land use. Though impactful for cultural heritage and groundwater management, its applicability remains limited without broader validation. Jol and Smith (1992) demonstrated the superiority of GPR in resolving internal sediment structures of lake deltas, thereby improving depositional models through enhanced resolution and flexibility. Yet, many lacustrine systems still lack detailed subsurface data, which limits model development and underscores the need for high-resolution imaging.

Modeling signal attenuation is especially valuable in environmental contexts, as it aids in assessing sediment stability, planning dredging operations, and monitoring morphological changes (Grimm et al., 2007; Elsner et al., 2021; Purmalis et al., 2016). Fediuk et al. (2022) demonstrated the effectiveness of GPR for shallow-water archaeological surveys (2–4 meters), outperforming seismic and geoelectric methods in terms of resolution and target detection. However, limitations remain, including environmental constraints and insufficient evaluation of biomass and sediment effects—highlighting the need for further methodological refinement in diverse freshwater settings. Eli and López (2022) used high-resolution 3D GPR to analyze stratigraphy at the Boise Hydrogeophysical Research Site, revealing sediment heterogeneity and fluvial compartmentalization. Although effective, the study's site-specific scope and lack of supporting hydrogeological data limit broader applicability. This highlights the need for multi-scale, cross-site research to build generalizable models of sedimentary and geophysical characteristics in lake settings.

GPR signal amplitude provides key information on substrate composition, influenced by dielectric contrasts between water and sediment. Sand and clay, commonly found in lakes, exhibit distinct signal behaviors—clay's higher conductivity causes stronger attenuation, while sand allows clearer reflections (YongShuai et al., 2019; EPA, 2023; Castro et al., 2024). These

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

differences impact both geophysical interpretation and ecological conditions, including nutrient retention and habitat quality. Seasonal moisture also impacts GPR performance, with drier conditions improving signal clarity. Studies show sandier substrates yield more reliable results (Castro et al., 2024; Paredes-Palacios et al., 2022), emphasizing the need to account for sediment variability and seasonal effects in GPR-based lake assessments.

The application of machine learning is particularly crucial in geospatial analysis as it allows for the efficient processing and interpretation of complex, high-dimensional datasets, such as those derived from GPR surveys, and was used by many authors (e.g., Küçükdemirci & Sarris 2022; Dai et al. 2022; Bai et al. 2023; Mojahid et al. 2025; Shehab et al. 2025). Traditional methods for classifying aquitard zones are often limited by manual interpretation and are subject to human error and bias. Machine learning provides an objective and scalable alternative, enabling the automated classification of aquitard types and thereby enhancing the accuracy and consistency of predictions across large datasets.

Machine learning (ML) techniques have shown promising results in studying riverbed and lakebed sediments. Random forest and deep neural network models have been effective in classifying substrate types from underwater imagery (Geisz et al., 2024). ML approaches have also been successful in mapping riverbed grain-size distribution (Ren et al., 2020) and predicting sand-gravel-cobble facies patterns using lidar data (Diaz-Gomez et al., 2022). These methods can reveal sedimentological patterns in river deposits (Demyanov et al., 2019) and improve predictions of flow resistance in alluvial channels (Mir & Patel, 2023). Machine learning (ML) models have demonstrated better accuracy than traditional approaches in modeling sediment transport (Bhattacharya et al., 2007; Lund et al., 2022; Mohsen et al., 2025). However, challenges persist in striking a balance between model complexity, accuracy, and interpretability (Ho & Goethals, 2022).

Few studies have explored the integration of GPR amplitude features with ensemble learning models to delineate lakebed hydrostratigraphy and assess aquitard behavior. Moreover, the validation of ML-based classifications using ground-truth sediment sample data is often lacking. This study addresses these gaps by employing amplitude-derived features from GPR surveys, coupled with sediment samples observations, to classify aquitard, leaky aquitard, and non-aquitard zones in a shallow saline (alkaline) lake environment in Hungary (Carpathian Basin). By evaluating the performance of multiple supervised machine learning models, this work presents a novel methodological framework for subsurface lakebed characterization, contributing to the broader application of machine learning in hydrogeophysical investigations.

Shallow saline lakes in Hungary are facing increasing stress due to declining water quality and quantity, primarily caused by climate change and prolonged drought periods. These issues have lowered water levels and disrupted the ecological balance, which has been intensified by human activities. To address these challenges, understanding the hydrological, ecological, and geochemical systems of the lakes is essential. The unfavorable changes in hydrological conditions and their drivers have

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

been studied thoroughly by many authors (Rakonczai 2011; Ladányi et al. 2016, 2023; Molnár et al. 2019). However, the interactions between surface and groundwater have not been fully explored.

Although atmospheric drivers such as increasing evaporation and climate variability are identified as driving factors for the noticeable decline in the water level of the studied lake (Lake Vadkerti), there is growing speculation that the lakebed may allow downward water movement, indicating that it may not function as a fully impermeable barrier. This possibility raises concerns about previously unquantified water losses resulting from infiltration into the underlying strata. The central aim of this research is to evaluate whether the lakebed serves as an effective barrier (aquitard) to subsurface flow or if it facilitates seepage, thereby accelerating the decline in lake volume. To address this, the study integrates ground-penetrating radar (GPR) and machine learning-based classification to characterize the subsurface without the need for sediment sampling. These methods will help define the physical and hydrological nature of the lakebed and assess its influence on the lake's overall water balance. The specific objectives are:

1. To analyze lakebed composition and structure using GPR by identifying stratigraphic features, reflection amplitudes, and attenuation models to assess subsurface permeability and potential for groundwater exchange.
2. To classify lakebed permeability zones and predict soil types using GPR amplitude data through Python-based processing, spatial analysis, and machine learning models.

2 Study area

Lake Vadkerti is a saline (alkaline) lake situated on the sandy region of the Danube-Tisza Interfluve, on the Kiskunság sand ridge, approximately 3 km northwest of Soltvadkert, within the deepest part of an 8-10 km long northwest-southeast oriented depression (Fig. 1). The lake basin lies 1–2 meters lower than its immediate surroundings (at 110 m above sea level). Facilitates the flow of both surface and groundwater into the lake basin. The surface area of the Lake is 67.3 hectares (at an elevation of 108.5 m B.A.), with 43.7 hectares open water and 23.6 hectares covered by reed in the eastern part of the Lake. Characteristically, the Lake has a relatively shallow water depth, ranging from 1 to 2 meters across its entire extent.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

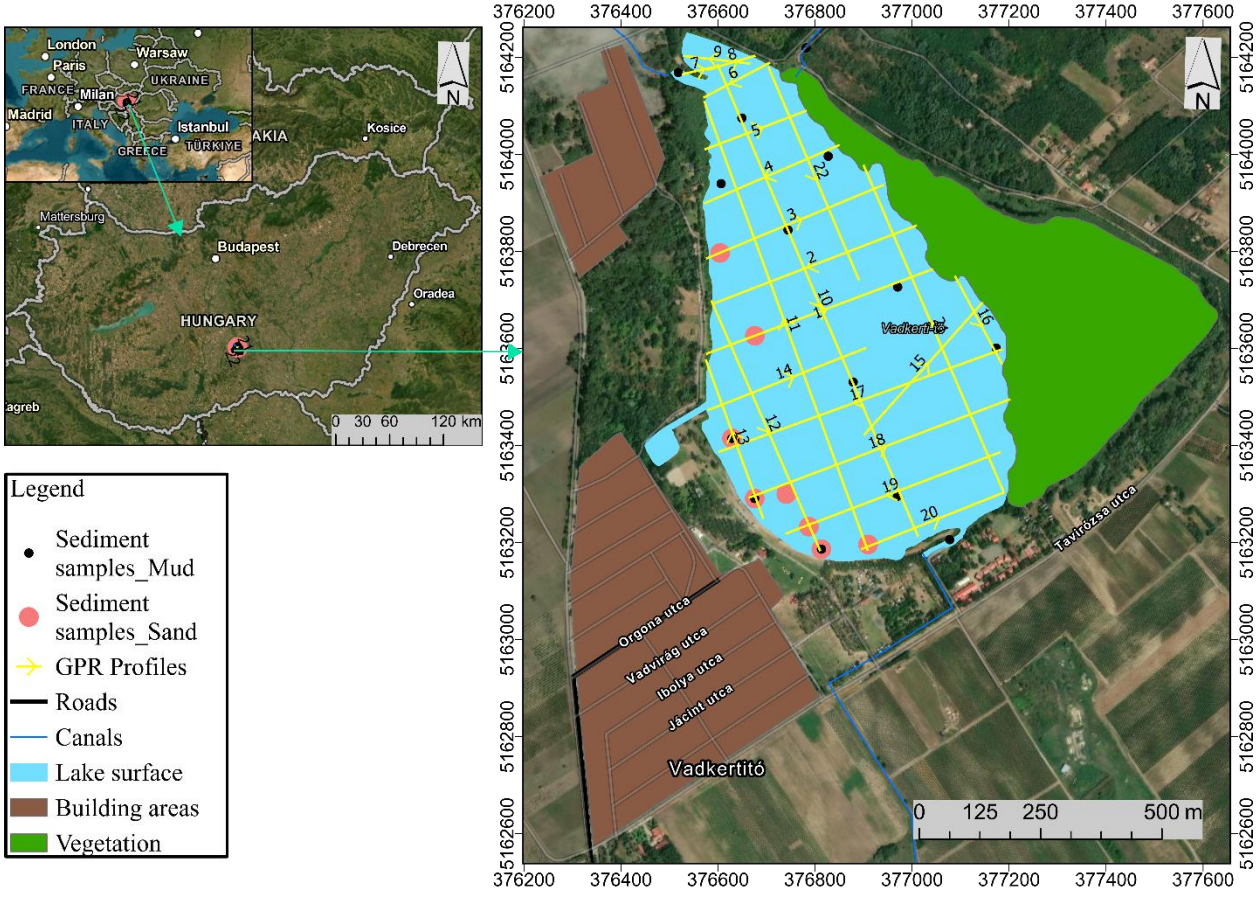


Fig 1 Map of the study area indicating the location of GPR profiles and the sediment samples. The map was created in ArcGIS Pro using Esri basemap imagery. © Esri, Maxar, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community.

The climate of the area is characterized by a yearly mean temperature of 10.2-10.5 °C and a yearly precipitation sum of 520-570 mm, with 320 mm occurring during the vegetation period (Dövényi, 2010). Regarding the lake's sustainability, it is essential to note that the annual precipitation sum is highly variable. The highest recorded yearly precipitation sum was 1012 mm (in 2010), while the lowest was 336 mm (in 2000) (data source: Lower Danube Valley Water Directorate). Furthermore, according to Baics and Centeri (2011), the annual precipitation decreased by 30.5 mm

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

between 1951 and 2009, suggesting that the conditions for water recharge from rainfall in the lake will be increasingly limited in the future.

The long-term water level fluctuation is 1.32 m. In comparison, typical intra-annual variability is approximately 0.5 m, except during hydrologically extreme events (Fig. 2). Under average conditions, based on the lake bed geometry, the lake holds an estimated volume of 914,000 m³, increasing to 1.4 million m³ at the highest recorded level.

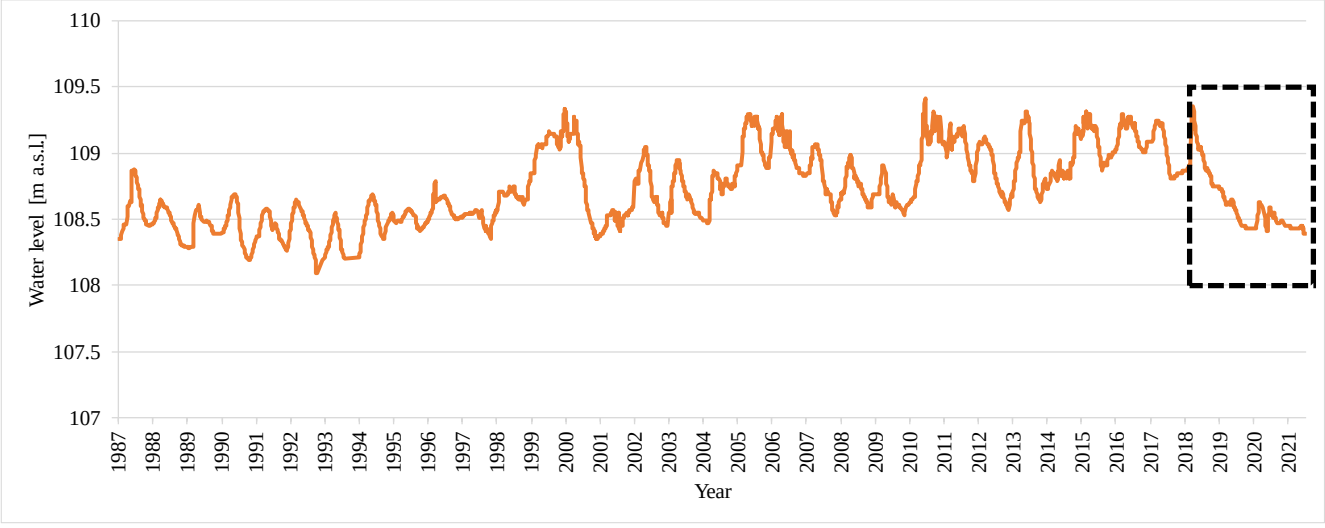


Fig. 2 Long-term evolution of Lake Vadkerti's water level from 1987 to 2021, based on hydrographic data provided by the Lower Danube Valley Water Directorate. A noticeable decline in lake water levels in recent years is highlighted by the dotted rectangle box.

The lake's hydrological balance is shaped by both natural and anthropogenic components, with clear implications for its long-term water level stability. Direct atmospheric precipitation onto the lake surface contributes approximately 350,000 m³ annually, representing a significant input with a stabilizing effect. Although this input is subject to interannual variability, it does not contribute to the decline in water levels. Surface runoff from the surrounding catchment is minimal, while subsurface inflow from groundwater reservoirs adds an estimated 245,000 m³ annually. However, this depends on the relative elevation of the lake's water surface and the surrounding aquifers. Evaporative losses constitute the most significant and persistent negative component of the lake's water budget, approximately 440,000 m³, with model projections indicating an additional 15,000–30,000 m³ increase for the 21st century (Keve & Nováky, 2011; Csáki et al., 2018). Infiltration into the

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

sandy subsurface results in additional losses, estimated at 130,000 m³ per year. Although this infiltration process is gradual, it consistently subtracts from the lake's volume. To counteract these deficits, approximately 400,000 cubic meters of groundwater are artificially introduced each year. However, this artificial water supply is insufficient to reverse the long-term downward trend. Overall, the cumulative effect of high evaporation rates, limited natural inflows, and relatively inefficient artificial supplementation has led to a persistent and measurable decline in the lake's water level despite interannual variations (Sipos et al., 2021).

3 Materials and methods

GPR measurements were conducted on the water surface of the lake using a GSSI SIR-3000 data acquisition unit in time mode. The apparatus was connected to an antenna with a central frequency of 200 MHz; the device and the antenna were placed on two tightly connected boats, with the antenna's boat being plastic to allow the EM waves to penetrate the lake freely (Fig. 3). In total, 22 profiles with varying lengths were collected, depending on the length or width of the lake, to cover the whole lake; specifically, 15 GPR profiles were measured transversely, and the seven remaining profiles were measured longitudinally. The exact location and length of each profile are shown in Fig. 1. The start and end points of each GPR profile were measured using GPS. The data were collected over a time range of 180 ns. A total of 32 scans were obtained per second, 35 scans were obtained per meter, and 1,024 samples were obtained per scan. The lake bottom and sides were clearly observed during data collection, with variations in the first arrivals of the water–soil interface from positive to negative. The dielectric permittivity (ε) of the water was 81.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Fig. 3 Field photo showing GPR data acquisition on Lake Vadkerti using a GSSI SIR-3000 system equipped with a 200 MHz center frequency antenna. The system was mounted on a boat, and a high-resolution RTK-GPS unit was used to provide precise positioning and accurate tracking of the GPR profiles.

The GPR interpretations were validated using sediment samples collected at locations where significant lithological changes occurred within the lake bed. Twenty-three sediment samples were obtained (Fig. 1), of which 15 were composed of clay (distributed throughout the lake bed) and eight consisted of sand (primarily concentrated along the western shoreline).

The collected GPR data contained various types of noise because no filters were used in the field. This noise was eliminated by applying the following processes to the GPR cross sections using the software RADAN 7, version 7.6 (GSSI, 2018): time-zero correction to move the data to an effective time zero, infinite impulse response filtering, stacking, range gain setting, and migration. RADAN 7 was also used to identify the reflection amplitudes of the first arrivals of the water–lake bottom interface (whether they were positive or negative), which were then used to distinguish between sand and clay at the lake bottom and sides.

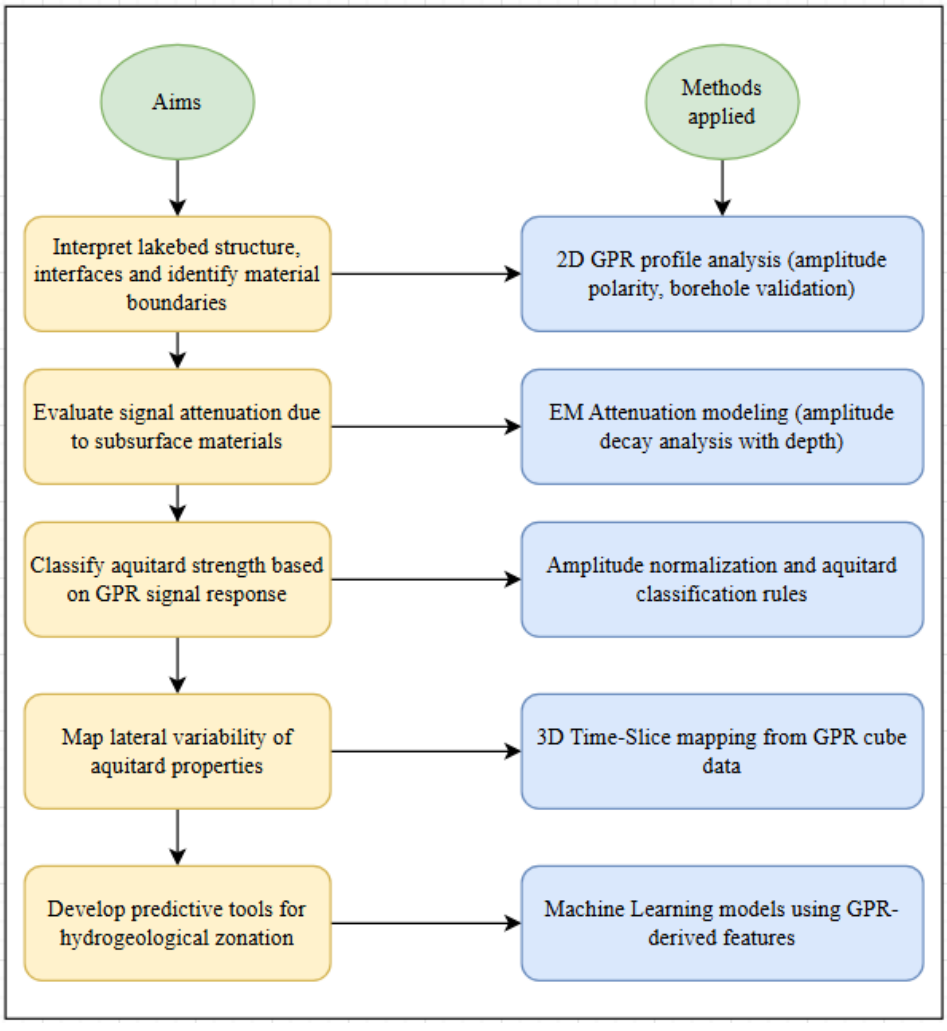
All GPR profiles were obtained in time mode and subsequently underwent distance normalization to maintain a uniform horizontal scale throughout the dataset. This procedure was essential because of small discrepancies in the antenna towing speed during data collection, which frequently arise when surveys are performed in continuous mode without survey wheels. Distance normalization rectified these discrepancies by modifying the spacing between scans by either elongating or condensing the data, ensuring a consistent number of scans per unit distance between designated markers. This method, commonly called rubber-sheeting, was implemented using marker information from the Way Points tab of the GSSI data processing software, ensuring that all profiles were spatially coherent and appropriate for precise interpretation and comparison.

After the processing of all 2D GPR profiles, the Ez Tracker tool was used to identify and trace six successive subsurface interfaces within the lake bed systematically. These interfaces were selected based on consistent reflection patterns across the profiles. The reflection amplitudes associated with these interfaces were extracted at 1 m intervals along the survey lines for a high spatial resolution. The exported amplitude data were the basis for subsequent analyses, including reflection coefficient estimation, classification of lake bed materials, and spatial interpolation for the mapping of subsurface permeability zones (Fig. 4).

3D GPR analysis was performed using the software GPR-Slice (Goodman, 2017). The coordinated GPR profiles were processed through time-zero correction, automatic gain control, band-pass frequency filtering, background removal, migration, and Hilbert transform. Then, the data were sliced at every 10 cm thickness, for a total of 27 slices. The data were

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

then gridded using a value of 80 for both the X and Y search radii, with a grid cell size of 2. Then, a 3 × 3 low-pass filter was applied. Finally, the 3D model of the lake bottom was created, and the amplitude of each slice was exported for further analysis.



220

Fig. 4 Methodological framework for characterizing aquitard properties and hydrogeological zonation using GPR

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

4 Data interpretation

4.1 EM attenuation modeling

Attenuation loss models should be developed for EM wave behavior through sand and clay in lake bed environments to obtain insights into subsurface conditions. By capturing the distinct attenuation characteristics of these materials, which are largely influenced by their moisture content and electrical conductivity, such models enable accurate differentiation between sand and clay layers. This distinction is vital for understanding lake bed stratigraphy and identifying sediment accumulation and saturation changes. Moreover, incorporating such models into GPR data interpretation helps correct for signal amplitude loss, thereby enhancing the clarity and depth of subsurface imaging and improving the detection of buried interfaces and structures.

The attenuation loss of the material (L_a) is (Daniel, 2004)

$$L_a = 8.686 \text{ x } 2 \text{ x } R \text{ x } 2\pi f \sqrt{\left(\frac{(\mu_0 \mu_r \varepsilon_0 \varepsilon_r}{2} \left(\sqrt{1 + tan^2\delta}\right) - 1\right)}, \quad (1)$$

where f is frequency (Hz), tan δ is the loss tangent of the material, ε_r is the relative permittivity of the material, ε_0 is the absolute permittivity of free space, μ_r is the relative magnetic susceptibility of the material, and μ_0 is the absolute magnetic susceptibility of free space.

The attenuation behavior of the EM waves in the lake bed sediments was analyzed using the GPR data. The dataset consisted of horizontal distances from the start to the end of the measured GPR profiles, depths to the six subsurface interfaces, EM wave amplitudes corresponding to the interfaces, and the classified soil type (sand or clay) at each measurement point (validated from the sediment sample locations). Attenuation was modeled using an exponential decay function that describes how EM wave amplitudes decrease with depth due to energy loss in the subsurface material.

The relationship between amplitude and depth was modeled using the standard attenuation equation (Daniel, 2004; GeoSci, 2025)

$$A(z) = A_0 \cdot e^{-\alpha z}, \quad (2)$$

where A(z) is the amplitude at depth z, A_0 is the initial amplitude on the surface, and α is the attenuation coefficient (Np/m). α was estimated by linearizing the equation using the natural logarithm of both sides.

$$\ln(A(z)) = \ln(A_0) - \alpha z \quad (3)$$

With this transformation, linear regression was applied to the logarithm of the measured amplitudes versus depth. For the GPR data, zero amplitude values were filtered out, and the regression was performed separately for the sand and clay samples. The slope of the resulting line corresponded to the negative of the attenuation coefficient, from which an average α value was calculated for each soil type. These coefficients were then used to plot theoretical attenuation curves, normalized

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

to an initial amplitude of 1, for a visual comparison of EM wave behavior through sand and clay. Then, a custom Python script was developed to automate EM attenuation modeling, including data filtering, regression analysis, coefficient estimation, and curve plotting.

4.2 GPR-based aquitard classification via amplitude normalization

Amplitude data from the 22 GPR profiles acquired across the lake were systematically analyzed to characterize the hydrostratigraphic properties of the lake bed. In each GPR profile, boundaries between sand and clay layers were identified based on the polarity of the reflected signals. A positive–negative–positive (P–N–P) polarity sequence was indicative of sand, whereas a negative–positive–negative (N–P–N) sequence was characteristic of clay. In cases where the polarity did not clearly conform to either pattern, a transitional zone was inferred. These interpretations were validated using sediment sample data, which provided ground-truth confirmation of the subsurface sediment composition and the reliability of the GPR-based classification.

Amplitude values were then extracted from the delineated zones in the profiles. For consistency across datasets, each amplitude was normalized by dividing it by the highest recorded value: 32,766. This value corresponded to the maximum reflectivity measured from a metal calibration plate and was typical in sand-dominated zones, which had the highest signal strength. The average normalized amplitude for each zone was subsequently calculated to guide the classification process.

Based on the normalized amplitude values, the lake bed sediments were classified into three permeability-related categories. Zones with amplitudes ≤ 0.15, corresponding to fine-grained, low-permeability materials (e.g., clay), were categorized as aquitards. Those with amplitudes between 0.15 and 0.35, likely representing heterogeneous or transitional materials (e.g., silt and weathered sediments), were classified as leaky aquitards. Zones with amplitudes >0.35, indicative of coarse or fractured sediments with high permeability, were considered non-aquitards.

The thresholds between these categories were derived from the consistent data trends. Zones interpreted as clay based on polarity signatures and sediment samples validation reliably exhibited low normalized amplitudes, whereas sand-rich areas displayed considerably higher values. This clear relationship between amplitude behavior, polarity pattern, and ground truth provided a robust scientific basis for the classification scheme. This classification also aligned with GPR-based hydrogeological interpretations in sedimentary contexts (e.g., Neal, 2004; Jol, 1992). Each profile was georeferenced using start and end coordinates, enabling the spatial interpolation of classification points across the lake bed. A gridded surface was generated through spatial interpolation and constrained to the lake's extent using a vector shoreline boundary. The results were exported as a GIS-compatible shapefile for further spatial and statistical evaluation of aquitard distribution by depth.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

4.3 Use of ML for aquitard classification

Four ML models were used in this study for aquitard classification because of their effectiveness in spatial data classification: K-nearest neighbor (KNN), RF, extra trees (ET), and gradient boosting (GB) models. KNN is a nonparametric method that classifies instances based on the majority class of the nearest neighbors; it is interpretable but sensitive to class imbalance and feature scaling. RF and ET are ensemble methods that aggregate multiple decision trees to enhance classification accuracy and robustness; they balance bias and variance using key hyperparameters such as the number of trees (n_estimators) and tree depth (max_depth) . GB is another ensemble technique where trees are built sequentially to correct errors from previous trees; model learning is controlled by hyperparameters such as the learning rate and the number of estimators. These models were cross-validated to optimize their performance across the aquitard classes in the Lake Vadkerti dataset. The specific hyperparameters are listed in Table 1.

The custom Python script was used for a comprehensive geospatial classification of subsurface aquitard conditions using the GPR amplitude data. The amplitude data and profile coordinates were processed into a long format, which was subsequently pivoted to a wide format suitable for ML. The amplitude values of the six subsurface layers were normalized, and a classification scheme based on the sixth layer's amplitude was applied to categorize points as aquitard, leaky aquitard, or non-aquitard. The coordinates were interpolated along each profile to generate spatial positions. The four supervised learning models were trained to predict aquitard classes, and their predictions were exported as shapefiles. These shapefiles were then imported into ArcGIS Pro, where inverse distance weighting interpolation was applied to generate continuous spatial representations of the subsurface classifications, enabling detailed visual and analytical assessments of the model predictions across the lake bed.

Table 1 Summary of Hyperparameters Configured for the Machine Learning Models Used in Aquitard Classification

Model	Hyperparameter	Value
K-Nearest Neighbors	n_neighbors	5
	algorithm	auto
	metric	minkowski
Random Forest	n_estimators	100
	random_state	42
	criterion	gini
	max_depth	None
Extra Trees	n_estimators	100
	random_state	42
	criterion	gini
	max_depth	None

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Gradient Boosting	n_estimators	100
	learning_rate	0.1
	random_state	42
	max_depth	3

5 Results

5.1. Analysis of 2D GPR profiles and subsurface boundary interpretation

Together with the sediment samples data, the amplitude polarity variations in the GPR profiles were used to distinguish between sand and clay across the lake. These polarity changes reflected contrasts in the dielectric properties and electrical conductivity between the two materials, which influence radar wave reflection and transmission (Figs. 5a and 5b).

The amplitude polarity on the left slope (Fig. 5c), where sand was located, had a P–N–P sequence. This pattern generally indicates a highly pronounced disparity in dielectric characteristics between water and sand. Sand, having lower conductivity and greater porosity than clay, resulted in considerable reflections at its interface with water. GPR waves reaching this interface were substantially reflected due to the significant disparity in relative permittivity (dielectric constant) between water and sand. The P–N–P sequence was elucidated by phase shifts as the radar waves engaged with these interfaces, resulting in pronounced reflections; these were particularly due to the significant contrast between water and moist sand, which demonstrated high permeability.

On the right slope (Fig. 5d), which was dominated by clay, the amplitude polarity exhibited an N–P–N sequence. Clay has substantially higher electrical conductivity than sand, mostly due to the presence of water-bound ions, which decreases its resistance. Because of this high conductivity, the reflections from the radar waves striking the water–clay interface were diminished, resulting in a lower amplitude. The N–P–N pattern signified phase reversal caused by the interaction of the radar waves with the high conductivity of clay, where the diminished contrast at the water–clay interface caused weaker, more intricate reflections.

In the middle of the lake (Fig. 5d), where clay dominated at the bottom, the amplitude polarity had an N–P–N sequence. This pattern was less pronounced than that on the slopes because of clay's high water content and homogeneous composition. Clay typically exhibits low permeability and a high dielectric constant, thus slowing radar wave propagation and diminishing the reflections at interfaces. Consequently, the reflection amplitudes were small, and the amplitude polarity pattern was less pronounced. The diminished reaction may also indicate a uniform clay layer—i.e., without the pronounced differences in dielectric characteristics typical of the sand–clay interfaces on the slopes.

Fine-grained layers composed of materials such as silt, clay, and gyttja were observed beneath some parts of the sandy and clayey lake bed sediments, with thicknesses of 25–65 cm. These layers may function as aquitards, potentially influencing

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

groundwater–surface water interactions, such as seepage and infiltration. GPR profile 13, acquired along the western part of the lake (Fig. 5c), revealed a lake bed predominantly composed of a sandy unit approximately 40 cm thick. Two distinct fine-grained layers were identified beneath the sand: the upper layer (fine-grained layer 1) was narrower (~130 m) and thinner (~35 cm) than the lower layer (fine-grained layer 2), which was broader horizontally (~350 m) and vertically (~65 cm). Localized zones of weak radar reflections at the base of fine-grained layer 2, highlighted by the blue circles in Fig. 5c, indicated potential structural discontinuities or heterogeneities in the lake bed composition that may serve as preferential pathways for vertical seepage. GPR data from profile 1, acquired transversely at the central part of the lake (Fig. 5d), revealed the presence of a fine-grained layer not only beneath the sandy deposits but also beneath the clayey lake bed sediments. This fine-grained layer, measuring approximately 250 m laterally and about 50 cm vertically, was at the central bottom portion of the lake and overlain by clayey materials. The upper and lower boundaries of this layer exhibited weak radar reflections and local discontinuities in some parts, as marked by the blue circles. These zones may have increased permeability, potentially serving as vertical seepage pathways. The fine-grained layer narrowed toward the lake margins, suggesting additional potential conduits for downward seepage. The aquitard characteristics of the lake bed constituents are assessed in detail in the subsequent sections.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

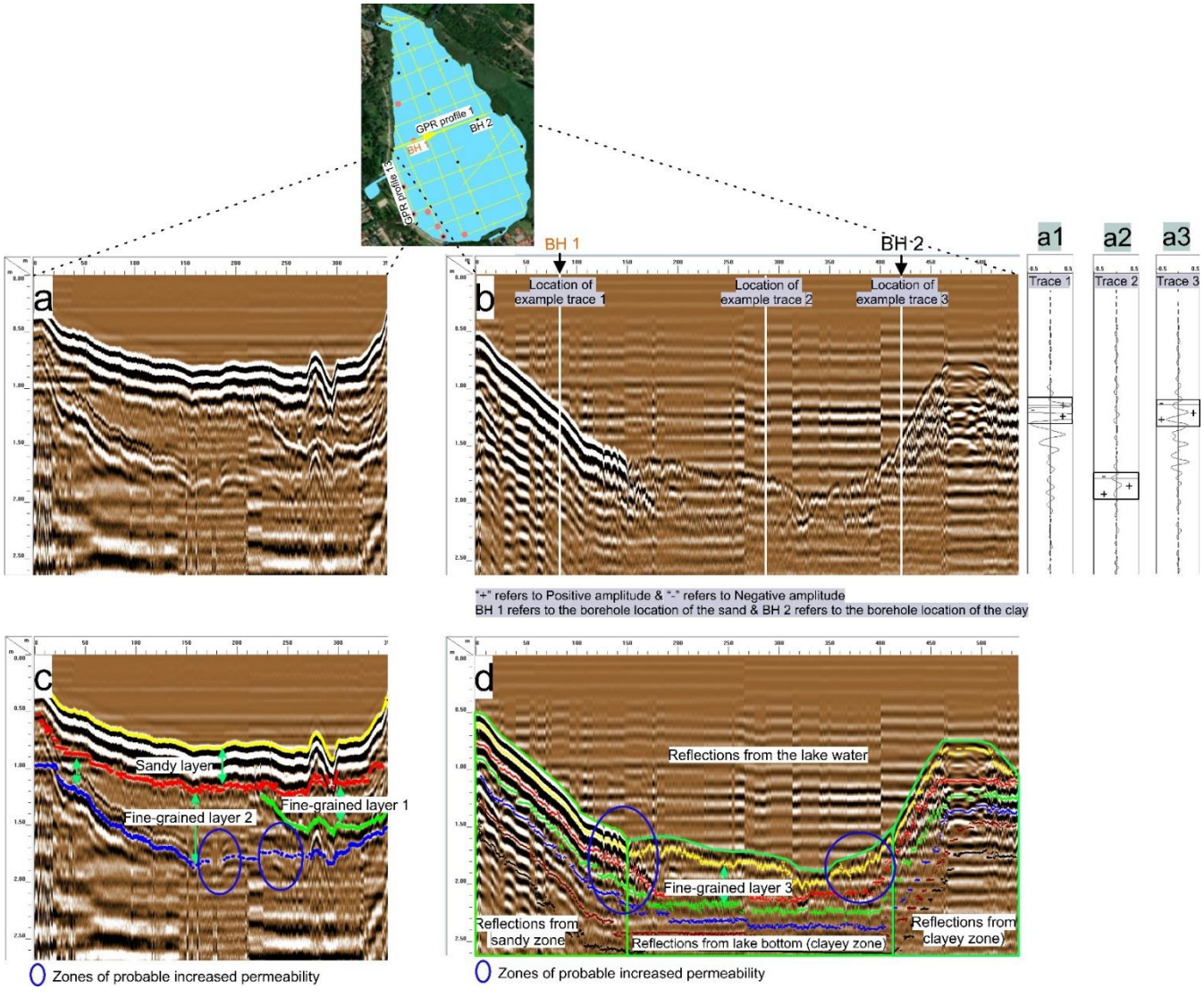


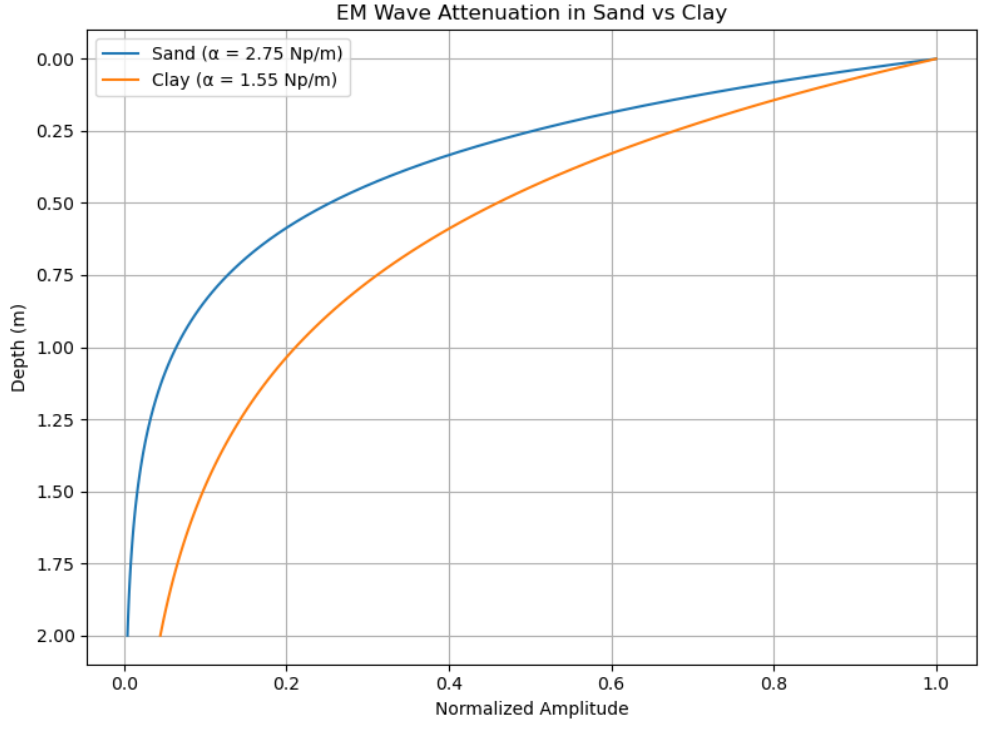
Fig. 5. 2D GPR profiles collected from the western and central parts of Lake Vadkerti. Panels (a) and (b) show Profiles 13 and 1, respectively, prior to interpretation, illustrating the accumulation places of sand and clay in the lakebed. Panels (c) and (d) present the same profiles after interpretation, highlighting amplitude polarity variations that differentiate sand from clay deposits. The interpreted profiles also reveal the presence of a fine-grained-rich layer beneath the sandy and clayey surfaces, which may function as a semi-confining aquitard influencing seepage processes and potential groundwater–lake interactions.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Blue circles indicate zones of potential vulnerability in the lakebed composition or discontinuities in the fine-grained layer that may facilitate downward seepage.

5.2 Attenuation-based differentiation of sand and clay layers in lake bed sediments

The attenuation curves derived from the GPR data showed differences in how EM waves propagated through the sand and clay within the lake bed, reflecting their contrasting hydrogeological properties (Fig. 6). The average attenuation coefficient (α) was 2.75 dB/m for sand and 1.55 dB/m for clay, indicating greater energy loss in sand. Therefore, the clay layers, with their low attenuation, enabled deep EM wave penetration and behaved as aquitards due to their fine-grained structure and high moisture retention, which limited water movement but helped sustain EM wave transmission. By contrast, sand, characterized by high attenuation, acted as a permeable non-aquitard medium. Its increased porosity and potential variability in moisture content caused stronger signal decay. Therefore, attenuation modeling effectively differentiated between the aquitard clay layers and non-aquitard sand layers in the lake bed, serving as a noninvasive tool for characterizing sediment permeability and interpreting groundwater flow.



https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Fig 6 Exponential attenuation of EM wave amplitude with depth in sand and clay, highlighting stronger signal loss in sand compared to clay due to differences in sediment properties.

5.3 Lake bed aquitard zonation and hydraulic insights

The resulting aquitard zoning map (Fig. 7) showed a heterogeneous distribution of subsurface materials beneath the lake. Zones classified as non-aquitards, characterized by high normalized amplitude values, were concentrated along the northeastern and central-western margins of the lake. These areas were likely coarse, permeable sediments that could facilitate vertical or lateral groundwater exchange. By contrast, the aquitard zones, which had low amplitudes, were mainly distributed in the central and southern parts of the lake, suggesting the presence of fine-grained, low-permeability sediments acting as hydraulic barriers. Intermediate zones (leaky aquitards) formed transitional bands between these end members, potentially functioning as semiconfining layers. The spatial correspondence of the non-aquitard zones with known or inferred hydrological pathways suggested that these areas may be preferential zones for groundwater–lake interaction, potentially influencing lake recharge or discharge dynamics.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

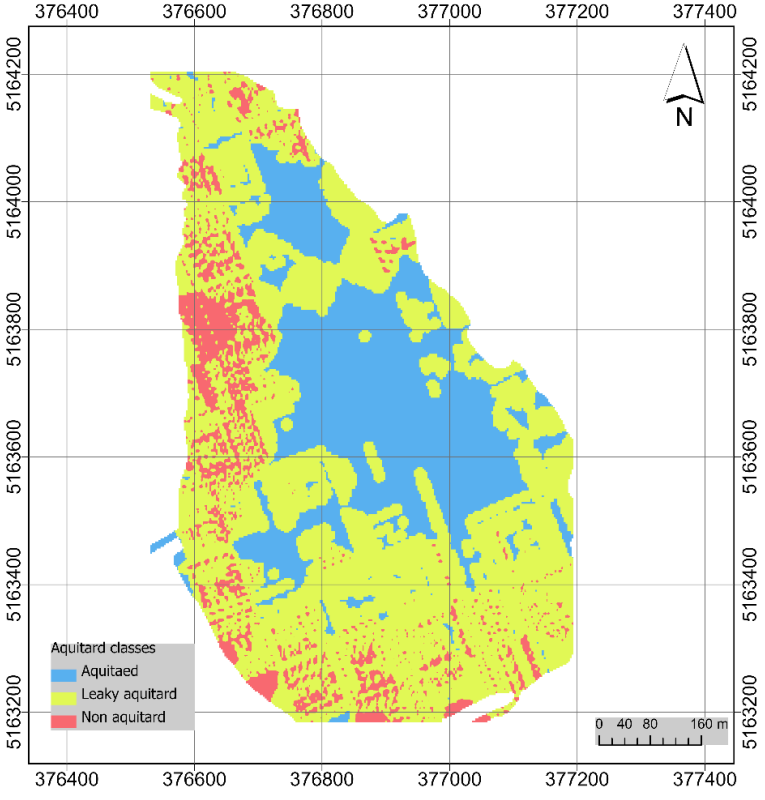


Fig. 7. Spatial distribution of lakebed classes based on aquitard characteristics, categorized as aquitard, leaky aquitard, and non-aquitard.

5.4 Spatial distribution of features from 3D time-slice mapping

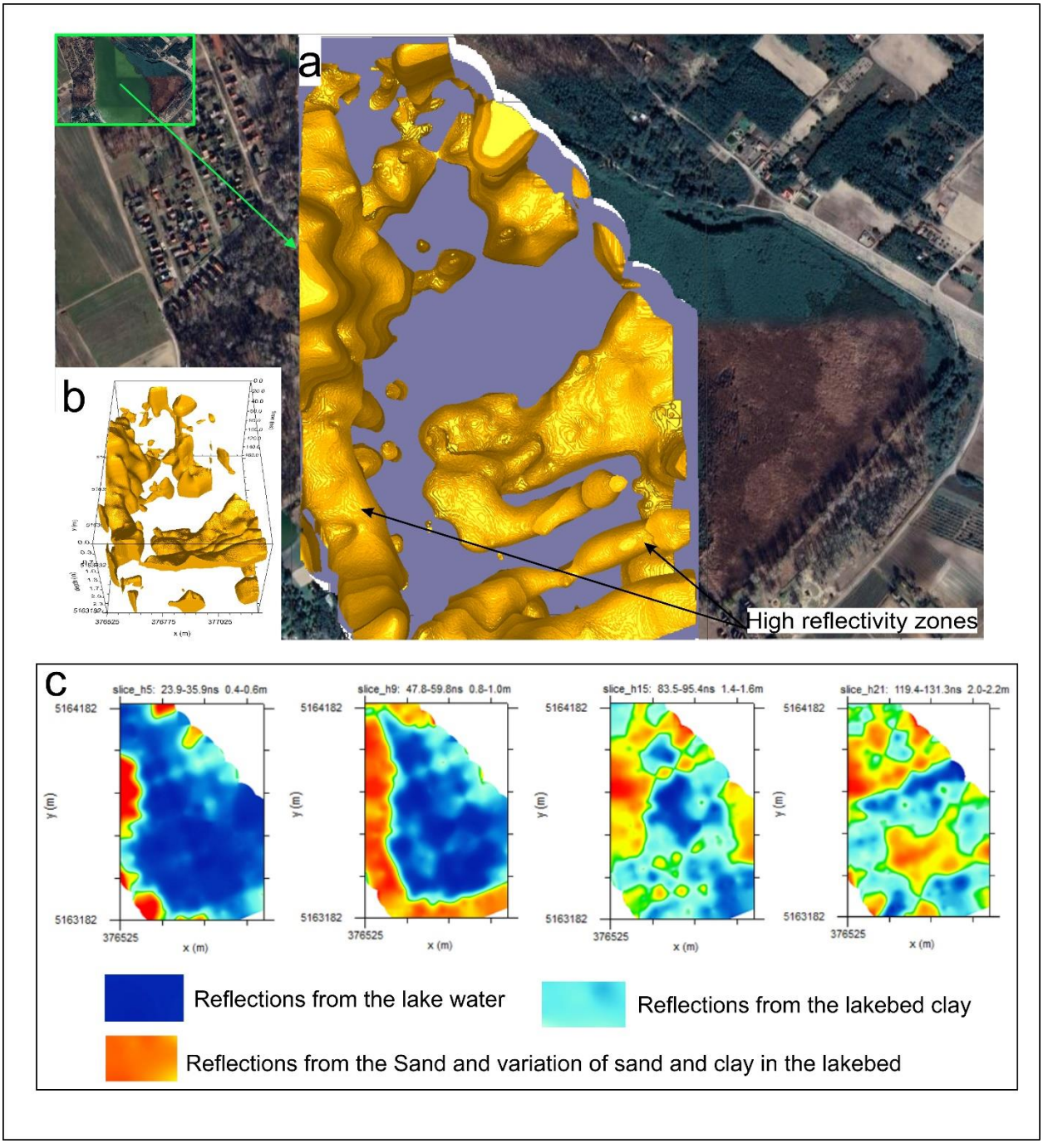
375 The GPR-Slice analysis of the GPR data enabled a detailed assessment of the EM wave reflection amplitudes, supporting the interpretation of subsurface material properties. The hydrostratigraphic variations within and beneath the lake bed were delineated, providing insights into sediment composition, layering, and potential hydraulic boundaries. With the amplitude analysis results, the two primary soil types (sand and clay) were effectively differentiated based on their distinct EM wave reflectivity, as shown in the 3D model (Figs. 8a and 8b). Time-slice amplitude maps were generated at 20 cm vertical
380 intervals (four slices are shown as examples in Fig. 8c), enabling the detailed tracking of lithological variations within the lake bed and through successive subsurface layers down to a depth of 2.6 m. The lowest amplitude values were associated

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

with the lake water, whose high electrical conductivity and dielectric constant resulted in strong signal attenuation and minimal reflection at the air–water interface. Upon reaching the lake bed, the EM waves exhibited markedly different reflection behaviors depending on the underlying material. The sand deposits produced high-amplitude reflections with limited signal attenuation, attributed to their low electrical conductivity and dielectric permittivity. This contrast resulted in a high reflection coefficient (R) at the water–sand interface, which strengthened signals. By contrast, the clay-rich zones displayed low-amplitude reflections with significant energy attenuation, primarily due to clay's high electrical conductivity and high dielectric permittivity, which increased signal absorption and weakened reflections. The theoretical reflection coefficient (R) at normal incidence supported these observations. The dielectric constant of water (~80) (Kearey et al., 2013) transitioning into saturated sand (~ 25) (GSSI, 2023) yielded a stronger reflection than the water–clay transition (~ 40) (Daniel 2004), which was more gradual and resulted in weaker signals. These amplitude-based interpretations were validated using the 23 samples. The high-reflection-amplitude zones corresponded to sandy deposits, whereas low-reflection-amplitude zones correlated with clay-rich sediments.

Although amplitude analysis effectively distinguishes major sediment types, interpreting the data can be challenging where signal differences are subtle or materials transition gradually. The situation highlights the need for machine learning, which can better capture complex patterns and improve the accuracy and detail of lakebed classification beyond what traditional methods can achieve. This approach will be discussed in the following section.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025



https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Fig. 8 3D model of the lakebed illustrating a) places of sand and mud deposits, background imagery © Esri, Maxar, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community, b) the 3D model in an inclined position to show thickness of the sand and variation of sand and mud deposits in the lakebed, c) Time slice maps exported from a depth of 0.5 m, 1 m, 1.5 m and 2.1 m illustrating the amplitude variations of water, clay and sand.

5.5 Predictive ML models

To enhance the classification of aquitard zones beyond traditional amplitude interpretation, the performance of the four ML models (KNN, RF, ET, and GB) in classifying aquitard zones based on GPR amplitude features was evaluated (Fig. 9). The KNN model achieved an accuracy of 83% on the test set, particularly for the aquitard class (precision: 84%, recall: 98%). However, it yielded significantly lower recall for the leaky aquitard (37%) and non-aquitard (36%) classes, reflecting challenges in detecting these less frequent classes. By contrast, the RF and GB models demonstrated perfect classification accuracy (100%) with precision, recall, and F1-scores of 1.00 across all classes. However, these results indicated potential overfitting; the models may have excessively memorized the training data and may thus generalize poorly to new, unseen data.

The ET model exhibited a more balanced performance, achieving an accuracy of 97% on the test set. It demonstrated high precision and recall for the aquitard class (precision: 96%, recall: 100%) and strong performance for the leaky aquitard (precision: 99%, recall: 86%) and non-aquitard (precision: 100%, recall: 79%) classes, with F1-scores of 0.92 and 0.88, respectively. Although ET did not achieve the perfect scores of the RF and GB models, it offered superior generalization across all classes without exhibiting signs of overfitting. In summary, although RF and GB excelled in test set accuracy, ET was more robust, with a better balance between model complexity and generalization.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

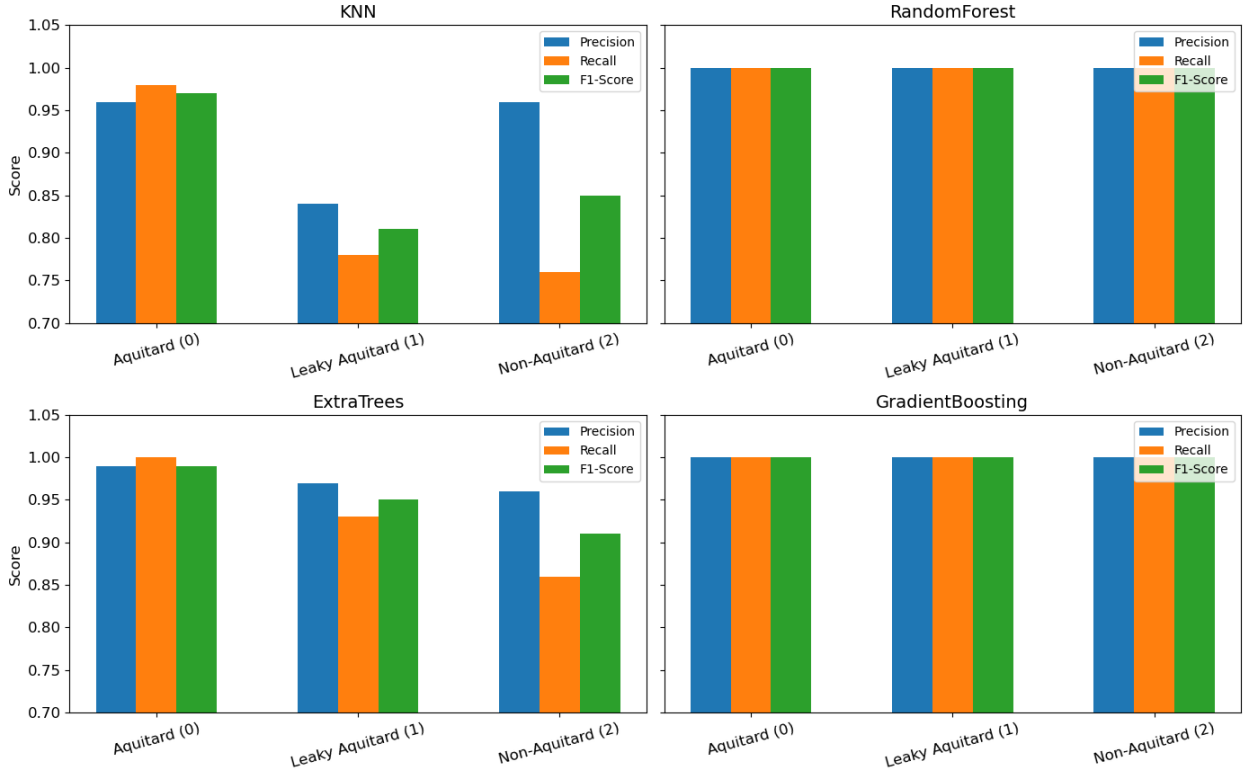


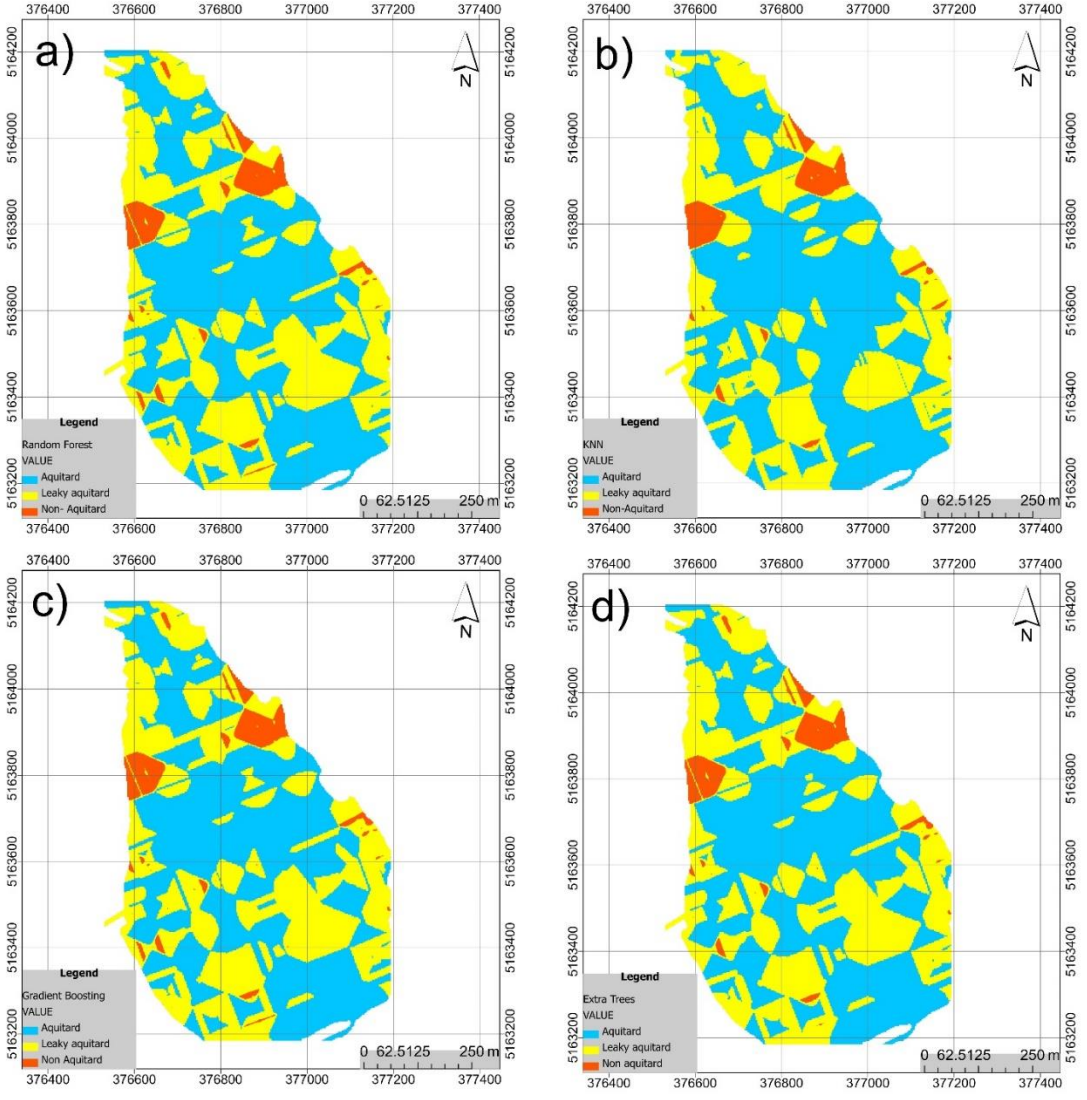
Fig. 9 Comparative analysis of model performance in terms of precision, recall, and F1-score for the four machine learning algorithms employed in this study—K-Nearest Neighbors (KNN), Random Forest, Extra Trees, and Gradient Boosting.

The classification results of the four ML models demonstrated a consistent spatial distribution of aquitard classes beneath the lake bed (Fig. 10). For all models, the non-aquitard zones were predominantly mapped along the northeastern and central-western edges of the lake, whereas aquitards were concentrated in the central and southern regions. These patterns aligned with the underlying GPR amplitude data, where higher amplitudes typically indicated coarser, more permeable sediments (non-aquitards), and lower amplitudes suggested finer-grained, less permeable materials (aquitards). Leaky aquitards appeared primarily in transitional areas between these classes, potentially representing intermediate-permeability zones that may act as semiconfining layers.

Strong spatial agreement was achieved between the ML model results (Fig. 10) and the resulting aquitard zoning map (Fig. 7), especially for the aquitard and leaky aquitard zones. The integrated map reinforced the dominant trends in the individual model outputs: non-aquitards (red) clustered around the lake margins, aquitards (blue) dominated the southern and central

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

portions, and leaky aquitards (yellow) were distributed as buffer zones between them. This close correspondence emphasized the reliability of the models and the strength of the data-driven approach in capturing the hydrostratigraphic complexity of the lake bed. These results also showed the potential of using ML and GPR data to inform groundwater–lake interaction studies with high spatial resolutions.



https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

Fig. 10 Spatial distribution of aquitard classification results using four machine learning models: a) RF, b) KNN, c) GB, and d) ET. Each map represents the predicted subsurface classification, with different aquitard types identified across the study area.

6 Discussion

The results of this study highlight the critical influence of aquitard properties identified through GPR surveys on the vertical exchange processes between groundwater and lake water. Low-permeability layers rich in fine-grained materials act as barriers at some locations below the lake bed, whereas more permeable or discontinuous zones, such as sand zones, likely facilitate seepage.

The amplitude polarity patterns identified in this study over sandy substrates and clay-dominated zones are consistent with previous research on radar wave behavior in saturated sediments. These previous works include those of Neal (2004) and Jol (2009), who demonstrated that sand typically produces stronger reflections with distinct polarity reversals due to its lower conductivity and higher porosity, whereas clay exhibits weaker, phase-reversed signals due to its higher ionic conductivity and lower dielectric contrast with water. Our results also match those of Beres and Haeni (1991), who observed diminished radar amplitudes and polarity inversions over clay-rich layers in fluvial settings, underscoring the effect of material composition on GPR signal behavior.

Our results reveal the presence of discrete fine-grained layers composed of silt, or clay beneath sandy and clayey lake bed sediments, with thicknesses of 25–65 cm. These findings are consistent with those of Kidmose et al. (2011), who identified a fine-grained organic layer (interpreted from GPR data) across Lake Hampen, although their reported layer generally covered a broader spatial extent and was interpreted in the context of peat or organic matter accumulation. Our identified fine-grained units, albeit thinner and more heterogeneous, are spatially confined and occur at varying depths beneath both sandy and clayey substrates. The lateral continuity of these layers (up to ~350 m) aligns with Kidmose et al.'s interpretation of persistent fine-grained strata across their target catchment, but the presence of localized discontinuities and weak radar reflections in the current study suggests greater structural complexity. Such heterogeneities may influence the hydraulic connectivity between the lake and the underlying aquifers by acting as potential vertical seepage pathways. Additionally, the identification of multiple fine-grained layers stacked beneath the sandy deposits in the western part of the lake extends the conceptual understanding of lake bed stratigraphy. The implied aquitard function is thus more variable than previously reported by Kidmose et al. (2011) for the Lake Hampen catchment.

The measured attenuation coefficients in this study (2.75 Np/m for sand and 1.55 Np/m for clay) are broadly consistent with those reported by Fediuk et al. (2022), who observed values ranging from 1.75 Np/m to 3.71 Np/m in shallow lacustrine

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

environments using a 400 MHz GPR antenna. Although the current study uses a lower frequency (200 MHz), the attenuation values fall within a similar range, suggesting that intrinsic material properties, such as porosity, mineralogy, and moisture content, have a greater influence on signal loss than antenna frequency alone. The higher attenuation observed in sand may appear unexpected, as sand is typically more conducive to GPR transmission due to its coarseness and low electrical conductivity (Utsi, 2017). However, the elevated porosity and moisture content in the sand layers likely contribute to high dielectric losses. Furthermore, attenuation in clay can vary significantly, depending on its mineralogical composition and saturation state, as different clay types exhibit distinct dielectric properties (Conyers, 2004).

Consistent with Geisz et al. (2024), who reported high accuracy (~95%–96%) for RF and DNN in lake bed mapping, the current study confirms that ensemble models, such as RF and ET, effectively capture hydrostratigraphic variability with a strong spatial resolution. Although Geisz et al. emphasized the efficiency of DNN, the ET model in the present work balances accuracy and generalization without complex preprocessing. Aligning with Ho and Goethals (2022), this study applies a structured ML workflow and shows that ensemble tree algorithms outperform traditional linear models in classifying aquitard zones from GPR data. Similar to Bhattacharya et al. (2007) who found ML models superior for analyzing sediment transport, the current study highlights ML's adaptability and predictive power in the geosciences. Consistent with Díaz-Gómez et al. (2022), who achieved 86% accuracy using RF and lidar, our research demonstrates the effectiveness of RF and ET in subsurface classification using GPR amplitudes. Reflecting the findings of Ren et al. (2020), who mapped riverbed grain sizes with moderate accuracy, the current work attains higher class-specific accuracy, improving aquitard delineation. Therefore, this study extends ML to the noninvasive, highly detailed spatial mapping of aquitard types, aligning with the work of Demyanov et al. (2019), who classified sediment cores to reveal depositional patterns. Mir and Patel (2023) showed that ensemble models can predict Manning's roughness ($R^2 \geq 0.99$), stating that these algorithms can robustly capture complex GPR data relationships. Finally, similar to the findings of Bhattacharya et al. (2007), who improved sediment transport modeling using ML, our results reinforce the broad applicability and accuracy of ensemble tree methods for hydrostratigraphic classification and subsurface heterogeneity mapping.

7 Conclusion

This study demonstrates that amplitude polarity variations in GPR profiles, when integrated with sediment samples data, are effective tools for distinguishing between sand and clay within the lake bed owing to these materials' contrasting dielectric and conductive properties. The polarity sequence over sandy and clay-dominated areas reflects consistent differences in radar wave behavior. Additionally, the identification of fine-grained layers beneath both sandy and clayey sediments—some exhibiting structural discontinuities and weak reflections—suggests the presence of potential vertical seepage pathways

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

through otherwise low-permeability strata. Through attenuation analysis, permeable sand and aquitard clay layers are distinguished effectively, with sand's higher attenuation indicating greater porosity and moisture variability and clay's lower attenuation reflecting its fine texture and water retention. Moreover, the aquitard zoning map reveals distinct subsurface zones beneath the lake, with permeable non-aquitard areas concentrated along the northeastern and central-western margins, likely facilitating groundwater exchange. Low-permeability aquitard zones dominate the central and southern regions, acting as barriers. Transitional leaky aquitard zones exhibit semiconfining behavior. Such mapping is supported by the GPR-Slice amplitude analysis of the GPR data, where the sand and clay within the lake bed are successfully differentiated: high-amplitude reflections correspond to sandy deposits, and low-amplitude reflections indicate clay-rich sediments. This interpretation, supported by theoretical reflection coefficients and validated using sediment sampling data, reveals the varying reflection behaviors of subsurface materials and offer valuable insights into the lake bed's lithological structure.

Overall, this research confirms that while amplitude variations in ground-penetrating radar profiles—validated by drilling and sediment analysis—offer extensive spatial coverage and meaningful insights into lake bed structure, manual interpretation of these data to delineate hydrostratigraphic zones remains inherently interpretive and methodologically demanding. By applying machine learning to amplitude features, we enabled automated, objective, and high-resolution classification of aquitard conditions across the entire lake bed. Among the tested models, the Extra Trees algorithm achieved the most balanced performance, offering high accuracy and strong generalization across all classes without signs of overfitting. Although Random Forest and Gradient Boosting models showed perfect accuracy on the test set, their results suggest potential overfitting and reduced robustness. All models, however, produced consistent spatial patterns that aligned well with the integrated aquitard zoning map, reinforcing their validity. Unlike conventional approaches that rely heavily on expert judgment or invasive sampling, the machine learning framework presented here enables scalable and reproducible prediction of subsurface permeability zones using only geophysical data.

These findings highlight the potential of combining GPR data with ML as a scalable, noninvasive approach to mapping and predicting subsurface conditions in large lake systems worldwide. By eliminating the need for traditional methods, such as sediment sampling, this approach is an efficient alternative for understanding groundwater–lake interactions in complex, extensive lake bed environments.

Author contribution

D.S. contributed to the conceptualisation, software, formal analysis, and writing of the original draft. **E.A.** participated in data curation, formal analysis and validation. **A.M.** was involved in investigations, software, data curation, and visualisation. **V.B.** contributed to methodology, and writing – review and editing. **P.K.** provided resources and contributed to the

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

investigation. **K.F.** supported the project through resource provision and project administration. **F.K.** was responsible for conceptualization and investigation. **I.B.** contributed to visualization, resources and validation. **G.S.** led supervision, funding acquisition, conceptualisation, and writing – review and editing.

Acknowledgements

This research was financially supported by the Interreg IPA Hungary-Serbia Programme (HUSRB/23R/11/006 – ADAPTisa). We gratefully acknowledge the Lower Tisza Region Water Directorate for providing a boat used in the GPR measurements, and the Lower Danube Valley Water Directorate for supplying water level data. We also thank the anonymous reviewers for their valuable comments on the manuscript. This manuscript has been edited by professional human editors at Enago to ensure language and grammar accuracy while preserving the author's intent. The authors thank Enago (www.enago.com) for the professional language editing of this manuscript.

Competing interests

The authors declare that they have no conflict of interest.

References

1. Annan, A. P. (2009). Electromagnetic principles of ground penetrating radar. In *Ground Penetrating Radar: Theory and Applications* (pp. 3-40). Elsevier. https://doi.org/10.1016/B978-0-444-53348-7.00001-6
2. Ayolabi, Elijah, and Rotimi Akinwale (2022). "Bathymetric and environmental deductions from GPR mapping of MTU Lake." Paper presented at the SEG/AAPG International Meeting for Applied Geoscience & Energy, Houston, Texas, USA, August 2022. doi: https://doi.org/10.1190/image2022-3751790.1
3. Bai, X., Yang, Y., Wei, S., Chen, G., Li, H., Li, Y., ... & Cui, H. (2023). A comprehensive review of conventional and deep learning approaches for ground-penetrating radar detection of raw data. Applied Sciences, 13(13), 7992.
4. Baics T., Centeri Cs. 2011: Kiskunhalas környéki szikes tavak vízgazdálkodási problémái, azok okai, lehetséges megoldások a Sós-tó vízutánpótlására. Tájökológiai Lapok 9/1, 53–71
5. Belanger, T.V., D.F. Mikutel, and P.A. Churchill. 1985. Groundwater seepage nutrient loading in a Florida lake. Water Res. 19:773–781.
6. Beres Jr, M., & Haeni, F. P. (1991). Application of ground-penetrating-radar Methods in Hydrogeologie Studies. *Groundwater*, *29*(3), 375-386.
7. Bhattacharya, B., Price, R. K., & Solomatine, D. P. (2007). Machine learning approach to modeling sediment transport. Journal of Hydraulic Engineering, 133(4), 440-450.
8. Bristow, C. S., & Jol, H. M. (2003). An introduction to ground penetrating radar (GPR) in sediments. https://doi.org/10.1144/GSL.SP.2001.211.01.01

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

9. Corradini, E., Wilken, D., Zanon, M., Groß, D., Lübke, H., Panning, D., Dörfler, W., Rusch, K., Mecking, R., Erkul, E. and Pickartz, N., (2020). Reconstructing the palaeoenvironment at the early Mesolithic site of Lake Duvensee: Ground-penetrating radar and geoarchaeology for 3D facies mapping. The Holocene, 30(6), pp.820-833.
10. Cullmann, J., W.J. Junk, G. Weber, and G.H. Schmitz. 2006. The impact of seepage influx on cation content of a Central Amazonian floodplain lake. J. Hydrol. 328:297–305.
11. Castro, K. C. P. L. D., Cunha, L. S. D., Sousa, A. C. D. A., Nogueira, P. V., & Borges, W. R. (2024). Influence of the sand-clay ratio of the burial material of forensic targets on ground-penetrating radar (GPR) responses–comparison of dry and rainy season data. Frontiers in Earth Science, 12, 1305496.
12. Conyers, L.B., 2004. Moisture and soil differences as related to the spatial accuracy of gpr amplitude maps at two archeological test sites. In: Slob, E., Yarovoy, A., Rhebergen, J. (Eds.), Proceedings of the Tenth International Conference on Ground Penetrating Radar, vol. 2, 435_438.
13. Csáki, P., Szinetár, M.M., Herceg, A., Kalicz, P., Gribovszki, Z. (2018) Climate change impacts on the water balance - case studies in Hungarian watersheds. Időjárás 122(1), 81-99. DOI:10.28974/idojaras.2018.1.6
14. Daniel, D. J. (2004). *Ground Penetrating Radar* (2nd ed.). London: The Institution of Engineering and Technology.
15. Dai, Q., Lee, Y. H., Sun, H. H., Qian, J., Ow, G., Yusof, M. L. M., & Yucel, A. C. (2022). A deep learning-based GPR forward solver for predicting B-scans of subsurface objects. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.
16. Dean, W.E., B.P. Neff, D.O. Rosenberry, T.C. Winter, and R. Parkhurst. 2003. The Signifi cance of ground water to the accumulation of iron and manganese in the sediments of two hydrologically distinct lakes in north-central Minnesota: A geological perspective. Ground Water 41:951–963.
17. Demyanov, V., Reesink, A. J. H., & Arnold, D. P. (2019). Can machine learning reveal sedimentological patterns in river deposits?
18. Díaz Gómez, R., Pasternack, G. B., Guillon, H., Byrne, C. F., Schwindt, S., Larrieu, K. G., & Solis, S. S. (2022). Mapping subaerial sand-gravel-cobble fluvial sediment facies using airborne lidar and machine learning. Geomorphology, 401, 108106. https://doi.org/10.1016/j.geomorph.2021.108106
19. Dövényi, Z. (ed) 2010. Magyarország kistájainak katasztere. MTA Földrajztudományi Kutatóintézet, Budapest, 70-73.
20. Eli, & López (2022). An Analog for Large-Scale Lacustrine Deposits: 3D Characterization of a Pleistocene Lake Bonneville Spit.
21. EPA United States Environmental Protection Agency. (2023, August 15). *Waterborne Ground Penetrating Radar*. https://www.epa.gov/environmental-geophysics/waterborne-ground-penetrating-radar
22. Elsner, C., John, F., & Hellbrück, H. (2021). Gigahertz Ground Penetrating Radar (GPR) for Sediment Exploration. Technische Hochschule Lübeck, Department of Electrical Engineering and Computer Science. Retrieved from https://www.th-luebeck.de/fileadmin/media_cosa/Dateien/Veroeffentlichungen/ScientificWorking2021/elsner_gigahertz_gpr_for_sediment_exploration.pdf
23. Fediuk, A., Wunderlich, T., Wilken, D., & Rabbel, W. (2022). Ground Penetrating Radar Measurements in Shallow Water Environments—A Case Study. *Remote Sensing*, *14*(15), 3659. https://doi.org/10.3390/rs14153659
24. GeoSci.xyz. (n.d.). *Basic Principles - GPG 0.0.1 documentation*. Retrieved April 9, 2025, from https://gpg.geosci.xyz/content/GPR/GPR_fundamental_principles.html
25. Geisz, J. K., Wernette, P. A., & Esselman, P. C. (2024). Classification of lakebed geologic substrate in autonomously collected benthic imagery using machine learning. Remote Sensing, 16(7), 1264.
26. Gómez, R. D., Pasternack, G. B., Guillon, H., Byrne, C. F., Schwindt, S., Larrieu, K. G., & Solis, S. S. (2022). Mapping subaerial sand-gravel-cobble fluvial sediment facies using airborne lidar and machine learning. Geomorphology, 401, 108106.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

27. Goodman, D., Piro, S., & Nishimura, Y. (2017). *GPR Remote Sensing in Archaeology*. Springer.
28. GSSI, Geophysical Survey Systems Incorporation. RADAN 7 Manual (2023); Geophysical Survey Systems, Inc.: Nashua, NH, USA, 2017. Available online: https://www.geophysical.com/wp-content/uploads/2017/10/GSSI-RADAN-7-Manual.pdf (accessed on 20 February 2025).
29. Goodman, D. (2017). *GPR-SLICE Software*. Retrieved from https://gpr-survey.com/
30. Grimm Robert, Cynthia Dinwiddie, David Stillman, and Ronald McGinnis, (2007). Frequency Dependence of Attenuation in Ground-Penetrating Radar: Insights from the Bishop Tuff. *Geophysics*, 72(2), J7–J16.
31. Hagerthey, S.E., and W.C. Kerfoot. 1998. Groundwater flow influences the biomass and nutrient ratios of epibenthic algae in a north temperate seepage lake. Limnol. Oceanogr. 43:1227–1242.
32. Hayashi, M., and D.O. Rosenberry. 2002. Eff ects of ground water exchange on the hydrology and ecology of surface water. Ground Water 40:309–316.
33. Ho, L., & Goethals, P. (2022). Machine learning applications in river research: Trends, opportunities and challenges. Methods in ecology and evolution, 13(11), 2603-2621.
34. Ito, M., M.J. Mitchell, C.T. Driscoll, R.M. Newton, C.E. Johnson, and K.M. Roy. 2007. Controls on surface water chemistry in two lake-watersheds in the Adirondack region of New York: Diff erences in nitrogen solute sources and sinks. Hydrol. Processes 21:1249–1264.
35. Jol, H. M., & Smith, D. G. (1992, June). Geometry and structure of deltas in large lakes: a ground penetrating radar overview. In *Fourth International Conference on Ground Penetrating Radar* (pp. cp-303). European Association of Geoscientists & Engineers.
36. Kidmose, J., Engesgaard, P., Nilsson, B., Laier, T., & Looms, M. C. (2011). Spatial Distribution of Seepage at a Flow-Through Lake: Lake Hampen, Western Denmark. *Vadose Zone Journal*, *10*(1), 110-124.
37. Kearey, P., Brooks, M., & Hill, I. (2013). An introduction to geophysical exploration. John Wiley & Sons.
38. Keve, G; Nováky, B (2011) Investigation of climate change effect on Bácsbokodi-Kígyó Basin with Budyko's model. In: Bálint, G; Domonkos, M (eds) XXVth Conference of the Danubian Countries on the Hydrological Forecasting and Hydrological Bases of Water Management. Budapest, Hungary, VITUKI p. 85.
39. Klenk, P., Jaumann, S., and Roth, K.: Monitoring infiltration processes with high-resolution surface-based Ground-Penetrating Radar, Hydrol. Earth Syst. Sci. Discuss., 12, 12215–12246, https://doi.org/10.5194/hessd-12-12215-2015, 2015.
40. Küçükdemirci, M., & Sarris, A. (2022). GPR data processing and interpretation based on artificial intelligence approaches: Future perspectives for archaeological prospection. Remote Sensing, 14(14), 3377.
41. Ladányi Zs, Blanka V, Deák ÁJ, Rakonczai J, Gábor Mezősi G (2016) Assessment of soil and vegetation changes due to hydrologically driven desalinization process in an alkaline wetland, Hungary. Ecological Complexity, 25, 1-10, https://doi.org/10.1016/j.ecocom.2015.11.002.
42. Ladányi Zs, Balog K, Tóth T, Barna Gy (2023): Longerterm monitoring of a degrading sodic lake: landscape level impacts of hydrological regime changes and restoration interventions (SE Hungary), Arid Land Research and Management, DOI: 10.1080/15324982.2022.2161970
43. Loeb, S.L., and C.R. Goldman. 1979. Water and nutrient transport via groundwater from Ward Valley into Lake Tahoe. Limnol. Oceanogr. 24:1146–1154.
44. Lower Danube Valley Water Directorate (LDVWD). *Water level data*. Baja, Hungary. https://www.aduvizig.hu/en
45. Lund, J. W., Groten, J. T., Karwan, D. L., & Babcock, C. (2022). Using machine learning to improve predictions and provide insight into fluvial sediment transport. Hydrological Processes, 36(8), e14648.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

46. Mir, A. A., & Patel, M. (2024). Machine learning approaches for adequate prediction of flow resistance in alluvial channels with bedforms. Water Science & Technology, 89(2), 290-318.
47. Mojahid, A., El Ouai, D., El Amraoui, K., El-Hami, K., & Aitbenamer, H. (2025). Intelligent recognition of subsurface utilities and voids: A ground penetrating radar dataset for deep learning applications. Data in Brief, 59, 111338.
48. Mohsen, A., Kovács, F., Baranya, S., Károlyi, C., Sheishah, D., Kiss, T., 2025. Insights into suspended sediment and microplastic budget of a lowland river: integrating in-situ measurements, Sentinel-2 imagery, and machine learning. Science of The Total Environment 984, 179716. DOI:https://doi.org/10.1016/j.scitotenv.2025.179716
49. Molnár, S., Z. Bakacsi, K. Balog, B. Bolla, and T. Tóth. 2019. Evolution of a salt-affected lake under changing environmental conditions in Danube–Tisza Interfluve. Carpathian Journal of Earth and Environmental Sciences 14 (1):77–82. doi:10.26471/Cjees/2019/014/060.
50. Neal, A. (2004). Ground-penetrating radar and its use in sedimentology: principles, problems, and progress. *Earth-Science Reviews, 66*(3-4), 261-330. https://doi.org/10.1016/j.earscirev.2004.01.004
51. Paredes-Palacios, D., Mota-Toledo, F., Biosca, B., Arévalo-Lomas, L., & Díaz-Curiel, J. (2022). Optimization of Dominant Frequency and Bandwidth Analysis in Multi-Frequency 3D GPR Signals to Identify Contaminated Areas. Sensors, 22(24), 9851.
52. Pitchford, J. L., Wu, C. Lin, L. S. Petty, J. T. Thomas, R. Veselka, W. E. Welsch, D. Zegre, N., and Anderson, J. T. 2012. Climate change effects on hydrology and ecology of wetlands in the Mid-Atlantic Highlands. Wetlands 32 (1):21–33. doi:10.1007/s13157-011-0259-3.
53. Purmalis, O., Alksnis, A., Taškovs, J., & Burlakovs, J. (2016). Ground-Penetrating Radar (GPR) Screening in Shallow Engure and Pape Lagoon Lakes. *Research for Rural Development 2016, Volume 1*, 228–235. Retrieved from https://llufb.llu.lv/conference/Research-for-Rural-Development/2016/LatviaResRuralDev_22nd_vol1-228-235.pdf
54. Rakonczai, J. 2011. Effects and consequences of global climate change in the Carpathian Basin. In Climate change – geophysical foundations and ecological effects, ed. J. Blanco and H. Kheradmand, 297–322. Rijeka: InTech
55. Ren H, Hou Z, Duan Z, Song X, Perkins WA, Richmond MC, Arntzen EV and Scheibe TD (2020) Spatial Mapping of Riverbed Grain-Size Distribution Using Machine Learning. Front. Water 2:551627. doi: 10.3389/frwa.2020.551627
56. Schafran, G.C., and C.T. Driscoll. 1993. Flow path composition relationships for groundwater entering an acidic lake. Water Resour. Res. 29:145–154.
57. Sebestyen, S.D., and R.L. Schneider. 2004. Seepage patterns, pore water, and aquatic plants: Hydrological and biogeochemical relationships in lakes. Biogeochemistry 68:383–409.
58. Shehab, M., Al-Kaltakchi, M. T., Dukhan, A., & Woo, W. L. (2025). Enhancing Ground Penetrating Radar (GPR) Data Analysis Utilizing Machine Learning. Journal of Engineering and Sustainable Development, 29(3), 321-330.
59. Sipos, Gy., Kiss, T., Farsang, A., Csáki, B., Murányi, G., Kiss, S. (2021) Soltvadkert, Vadkerti-tó állapotfelmérése és állapotmegőrzésre tett javaslatok kidolgozása további turisztikai fejlesztések céljából (in Hungarian), scientific report, p. 134.
60. Staehr, P.A., K. Sand-Jensen, A.L. Raun, B. Nilsson, and J. Kidmose. 2010. Drivers of metabolism and net heterotrophy in contrasting lakes. Limnol. Oceanogr. 55:817–830.
61. Striegl, R.G., and C.M. Michmerhuizen. 1998. Hydrologic infl uence on methane and carbon dioxide dynamics at two north-central Minnesota lakes. Limnol. Oceanogr. 43:1519–1529.
62. Utsi, E. C. (2017). Ground penetrating radar: theory and practice. Butterworth-Heinemann.
63. van Heteren, S., Fitzgerald, D. M., & Buynevich, I. V. (2018). Geophysical imaging of subsurface coastal and lacustrine environments. *Journal of Coastal Research, 34*(5), 1085-1098.

https://doi.org/10.5194/egusphere-2025-2904
Preprint. Discussion started: 5 August 2025

64. Wurtsbaugh, W. A., C. Miller, S. E. Null, R. J. DeRose, P. Wilcock, M. Hahnenberger, F. Howe, and J. Moore. 2017. Decline of the world's saline lakes. Nature Geoscience 10 (11):816–21. doi:10.1038/ngeo3052.
685 65. YongShuai, Y., Yajing, Y., & Guizhang, Z. (2019). Estimation of sand water content using GPR combined time-frequency analysis in the Ordos Basin, China. Open Physics, 17(1), 999-1007.