Waste Management 222 (2026) 115630

Contents lists available at ScienceDirect

Waste Management

journal homepage: www.elsevier.com/locate/wasman

Research Paper

Well-constrained LDA and SVM classification of ERT/IP data for leachate mapping in urban landfills

Giorgio De Donno [a], Davide Melegari [a], Ester Piegari [b,*]

[a] Dipartimento di Ingegneria Civile Edile e Ambientale, "Sapienza" Università di Roma, Rome, Italy
[b] Dipartimento di Scienze della Terra, dell'Ambiente e delle Risorse, Università degli Studi di Napoli "Federico II", Naples, Italy

ARTICLE INFO

Keywords:
MSW landfills
Electrical resistivity tomography
Induced polarization
Linear discriminant analysis
Support vector machine
Well data

ABSTRACT

Landfill leachate represents a major environmental concern due to its potential to contaminate groundwater. Geophysical methods such as electrical resistivity tomography and induced polarization are widely used to characterize landfill interiors and infer leachate plumes, as leachate typically shows distinct resistivity and chargeability relative to surrounding waste and natural soils. However, the interpretation of geophysical data is still largely qualitative or semi-quantitative, often leading to ambiguous results. To overcome these limitations, we propose a supervised machine-learning framework based on Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM), trained using resistivity (ρ), chargeability (m), and normalized chargeability (Mn) derived from well data. The classifiers applied to the full geophysical models obtain spatially continuous maps of leachate occurrence. Benchmark models (Logistic Regression and Random Forest) were tested to assess the role of model complexity under sparse training conditions. The results highlight the complementary strengths of LDA and SVM: LDA systematically achieves perfect recall (100%), providing a more inclusive reconstruction of leachate-affected zones, whereas SVM attains perfect precision (100%), enabling a more conservative delineation of leachate accumulations. In the three-parameter space (ρ, m, Mn), selected for both physical completeness and comparability with previous studies, both methods reach 95% accuracy, with F1-scores of 89% (LDA) and 86% (SVM), while Random Forest shows results comparable with SVM, and Logistic Regression yields lower performances. Overall, the proposed approach bridges the gap between point-scale well information and volumetric geophysical imaging, providing a more objective, reproducible, and operational tool for leachate mapping in complex urban landfills.

1. Introduction

Urban landfills are widely used systems for disposing of municipal solid waste (MSW), although they pose significant environmental challenges, particularly concerning the generation of leachate. Leachate is a toxic compound formed when water percolates through waste, whose accumulation can lead to severe environmental risks, including soil and groundwater contamination. Additionally, for landfills located on slopes, excessive leachate accumulation can also trigger instability phenomena. Therefore, mapping and monitoring of leachate accumulations are crucial for mitigating these risks and ensuring compliance with environmental regulations (Yuan and Zoungrana, 2025).

Traditional methods for leachate detection often rely on point measurements performed on scattered wells, which can be inadequate for capturing the spatial variability of leachate plumes. Geophysical methods, such as electrical resistivity tomography (ERT) and induced polarization (IP), can overcome these limitations to image the internal composition of MSW landfills and to successfully map low-resistivity and high-chargeability zones associated with leachate accumulation and preferential flow paths within complex waste deposits (e.g., Martorana et al., 2016; De Donno and Cardarelli, 2017; Soupios et al., 2017; Di Maio et al., 2018; Juarez et al., 2023; Xia et al., 2022; Ma et al., 2024a, b).

On the one hand, the ERT method, which senses the variation in electrical resistivity (ρ), can provide high-resolution images that indicate the presence of leachate, typically exhibiting lower resistivity compared to the surrounding dry waste due to its high salt content (e.g. Audebert et al., 2016). On the other hand, the study of the IP response of the subsurface through chargeability (m) is a proxy for detecting the waste mass, as chargeability is linked to the biogeochemical activity of

* Corresponding author.
E-mail address: ester.piegari@unina.it (E. Piegari).

https://doi.org/10.1016/j.wasman.2026.115630
Received 14 February 2026; Received in revised form 3 May 2026; Accepted 27 May 2026

the waste mass (Flores-Orozco et al., 2020; Han et al., 2025) The ratio of the two electrical parameters yields the normalized chargeability ($Mn = m/\rho$), which is often used to better separate the surface polarization contribution from the bulk conduction (Slater and Lesmes 2002). When combined, ERT and IP methods can offer an effective non-invasive tool for mapping leachate distribution over large areas, significantly enhancing landfill monitoring capabilities (see De Donno et al. 2024a, for a recent review).

Despite the effectiveness of ERT and IP techniques, the detection of leachate-saturated areas from geophysical models is not straightforward. In landfill sites, the application of common petrophysical relationships (i.e., Archie, 1942 and its subsequent developments) that link electrical parameters (resistivity and chargeability) to petrophysical parameters (saturation and porosity) is limited, because the behaviour of the waste mass differs significantly in compositional and spatial distribution from the lithologies for which these empirical relationships were originally established. Thus, interpretation of ERT and IP models often remains qualitative or based on simple thresholding, even when borehole and/or well data can provide a direct correlation between resistivity and chargeability values and leachate saturation.

Over last decades, artificial intelligence (AI) and machine-learning (ML) algorithms have been applied in the field of waste management providing innovative solutions, which initially primarily focused on waste generation forecasting, bin level monitoring, and collection route optimization (see Abdallah et al., 2020 for an extensive review of 2004–2019 works; Guo et al., 2021) and, more recently, on satellite-based landfill detection (Fraternali et al., 2024; Sharma and Rajendra, 2025) and waste sorting/classification (Lu and Chen, 2022; Fang et al., 2023; Farghali and Osman, 2024; Nwokediegwu et al., 2024; Ahmad et al., 2025). Thus, most of the studies have regarded intelligent waste management systems that operate on the surface and much less what regards the subsurface. Recently, Deep Learning approaches have been proposed to improve the inversion of electrical resistivity tomography data (Liu et al., 2020; Luan et al., 2023), especially to optimize leachate leakage detection (Chen et al., 2026; Sun et al., 2026).

To overcome the limitations of standalone ERT interpretation, several ML approaches have been proposed (e.g., Isunza Manrique et al., 2023; Piegari et al., 2023; Zhang et al., 2024; De Donno et al., 2024b; De Donno et al., 2025; Costanzo et al., 2026; Juarez et al., 2026) that use unsupervised ML algorithms to combine electrical resistivity with different geophysical and geochemical data for better identifying zones with anomalous electrical properties potentially associated with leachate accumulation. In particular, we have proposed a ML-based approach that directly group ERT/IP inverted data with similar properties to infer different degrees of leachate contamination (Piegari et al. 2023; De Donno et al. 2025). These studies demonstrate the potential of data-driven analysis but typically rely on clustering methods that are only weakly informed by independent geological and hydro-geological information. As a result, the connection between point-scale ground truth from wells and volumetric geophysical images of leachate plume remains partial. Therefore, compared with extensive drilling campaigns, the integration of geophysical imaging and supervised classification could potentially provide a less invasive and spatially continuous alternative for delineating leachate accumulations. In fact, direct detection through drilling is often costly, operationally difficult, potentially disruptive to landfill liners, and limited to sparse point measurements. In contrast, the combined use of resistivity, chargeability, and normalized chargeability provides a non-invasive, lower-cost, and spatially continuous means of imaging subsurface conditions that can be constrained and/or validated through a sparse logging of the leachate levels.

To fill this gap, in this study we propose the application of supervised classifiers such as Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM) that allow us to constrain the interpretation of geophysical subsurface imaging by leveraging well data ground truth. In fact, once resistivity and chargeability ranges for leachate-affected and unaffected zones are established from well data, we train LDA and SVM to recognize these classes and then apply them to the entire ERT/IP models, yielding spatially continuous images of the leachate accumulation zones. LDA and linear SVM were selected because they are particularly suitable for relatively small labelled datasets, computationally efficient, and directly interpretable through explicit decision boundaries in parameter space. These properties make them especially attractive in geophysical applications, where training data are often sparse and physically meaningful interpretation is required. In contrast to more complex nonlinear models, these classifiers reduce the risk of overfitting while preserving interpretability (Bishop and Nasrabadi, 2006; Hastie et al., 2017).

We examine how different parametrizations affect the spatial reconstruction of the leachate accumulation zones, validate results against independent well data and provide an uncertainty assessment in terms of leachate probability of occurrence. Additionally, we compare the results of LDA and SVM against other supervised algorithms, such as Random Forest (RF) and Logistic Regression (LOG), in terms of well-established classification metrics.

2. Materials and methods

2.1. Workflow of the proposed procedure

The overall workflow adopted in this study is summarized in Fig. 1. The approach integrates geophysical data acquisition, well-based ground truth, supervised classification, and spatial prediction.

First, ERT and IP data are acquired along multiple survey lines and simultaneously leachate levels are logged at selected (sparse) wells. Geophysical data are then processed and inverted to obtain resistivity and chargeability models, and normalized chargeability is derived as the ratio of m and ρ. In the third phase, among the whole wells we selected some wells for ML taring and we extract the values of geophysical parameters at their location to define training labels corresponding to leachate-saturated and non-saturated conditions. Supervised classifiers (LDA and SVM) were trained on the well-constrained dataset and ML classification is validated against independent information coming from other wells. Finally, the trained models were applied to the full geophysical dataset to generate spatial classification maps and probability of occurrence maps, allowing the identification of both high-confidence leachate zones and transitional areas.

2.2. Site description and input dataset

The input dataset derives from a geophysical investigation of an MSW landfill in central Italy. The landfill is located on a steep slope and has been active since the 1980 s, serving a medium-sized city. It is provided with a bottom high-density polyethylene (HDPE) liner preventing the leachate flow outside the landfill site. At this site, mapping the leachate accumulation zones with high-resolution was mainly required for planning future remediation activities, such as the design of a drainage network for ensuring the stability of the landfill.

To this end, ERT and IP data were acquired along four profiles (L1-L4) located on the terraces built for landfill management (Fig. 2, blue lines). Data acquisition and inversion results are described in detail in Piegari et al. (2023) and De Donno et al. (2025). Among the four ERT/IP profiles, we selected for the application of ML algorithms the L2 and L3 lines (solid blue lines in Fig. 2), since they are the only lines provided with active piezometers (P1-P5 in Fig. 2) for logging the leachate levels at the time of the geophysical measurements.

Firstly, we extracted ρ, m and Mn values from the inverted models on L2 and L3 at the same location of the wells. These datasets are shown in Fig. 3 together with the direct information coming from the piezometers, showing the leachate levels (dashed black line) and the bottom of the wells (solid black line). Although a general decrease of resistivity and increase of chargeability is generally visible passing from the

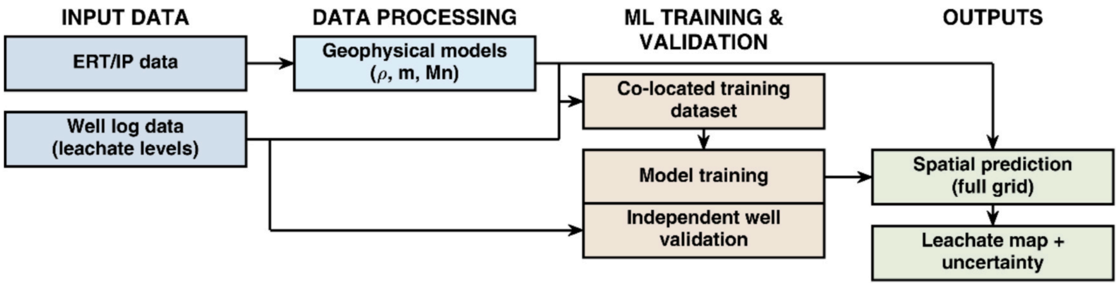

Fig. 1. Workflow of the proposed procedure for leachate mapping in urban landfills.

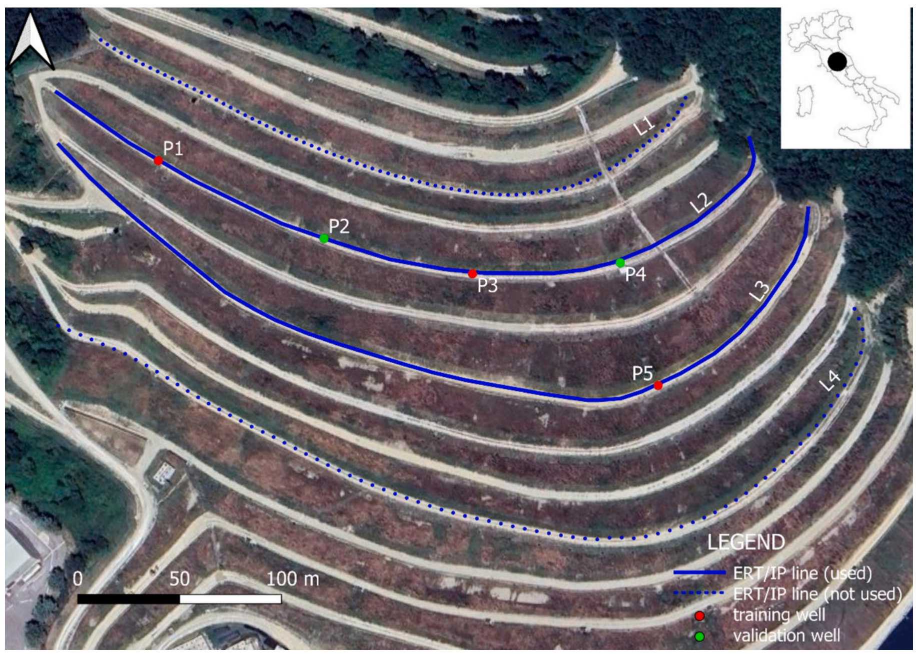

Fig. 2. Aerial view of the landfill site with location of the ERT/IP lines (blue lines). The lines L2 and L3 employed in this work are marked with blue solid lines. The location of wells used for training (piezometers P1, P3 and P5) is marked with red circles, while validation wells (P2 and P4) are marked with green circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

unsaturated to the saturated zones, these trends are not consistent among the five wells. Therefore, it is difficult to find common thresholds to be applied for the five wells, in order to extend the well information to the entire cross-sections and to different profiles, and ML techniques (LDA and SVM) help address this limitation by integrating well data with geophysical models.

2.3. Linear discriminant analysis (LDA)

Linear Discriminant Analysis (LDA) is a supervised technique used for classification and dimensionality reduction (Fisher, 1936). It works by finding a linear combination of features that maximizes the separation between multiple classes and in the field of waste management has been mainly used to improve waste material sorting and classification (Lu and Chen, 2022; Yang et al., 2023; Manakkakudy et al., 2024).

The working principle is to maximize the ratio of between-class variance and within-class variance (also known as the Fisher's ratio) ensuring that the projected data offers the best class separability.

The within-class scatter matrix $\mathbf{S_w}$ quantifies the spread of samples within each class and is written as:

$$\mathbf{S_w} = \sum_{k=1}^{K} \sum_{x \in C_k} (\mathbf{x} - \boldsymbol{\mu}_k)(\mathbf{x} - \boldsymbol{\mu}_k)^{\mathrm{T}} \tag{1}$$

where K is the number of classes ($K = 2$ in this case: dry/saturated), C_k represents class k, $\mathbf{x}$ is the parameter vector and μ_k is the mean vector of class k.

The between-class scatter matrix $\mathbf{S_b}$ quantifies the scatter of class means around the overall mean $\boldsymbol{\mu}$ and is defined as:

$$\mathbf{S_b} = \sum_{k=1}^{K} N_k (\boldsymbol{\mu}_k - \boldsymbol{\mu})(\boldsymbol{\mu}_k - \boldsymbol{\mu})^{\mathrm{T}} \tag{2}$$

where N_k is the number of samples in class k.

The LDA method finds the projection vector $\mathbf{w}$ that maximizes the following objective function

$$J(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}} \mathbf{S_b} \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \mathbf{S_w} \mathbf{w}} \tag{3}$$

The solution is the eigenvector corresponding to the largest eigenvalue in the generalized eigenvalue problem:

$$\mathbf{S_b} \mathbf{w} = \lambda \mathbf{S_w} \mathbf{w} \tag{4}$$

where λ quantifies the ratio of between-class to within-class variance along the projection direction $\mathbf{w}$: the larger is λ, better is class separation. This vector $\mathbf{w}$ is used to project the data onto a lower-dimensional space that best separates the classes and corresponds to the separation line between blue and red areas in the LDA plots presented in Section 3.

Binary classifications were complemented by posterior probability maps to quantify prediction confidence and identify transitional zones

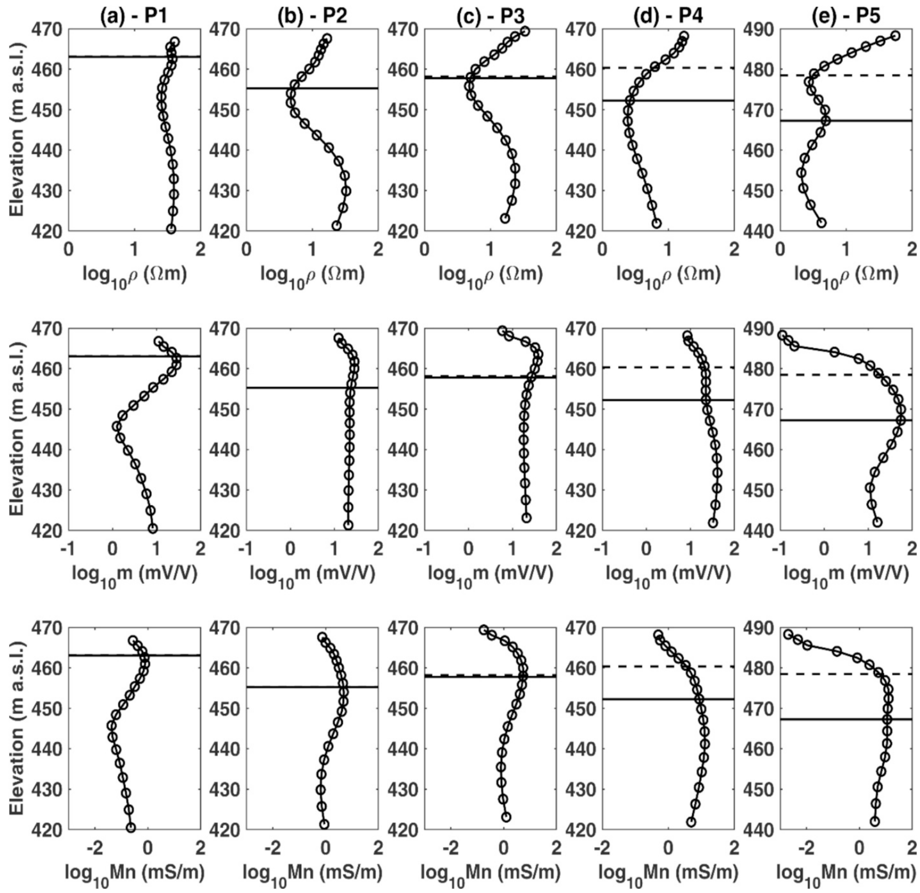

Fig. 3. Resistivity, chargeability and normalized chargeability values extracted from the inverted models in correspondence of the wells P1 (a), P2 (b), P3 (c), P4 (d) and P5 (e).

near class boundaries. For each cell of the inverted geophysical model, a probability of belonging to the leachate class was estimated. For LDA, posterior probabilities were computed using a Bayesian approach (McLachlan, 2004). Probability values close to 1 indicate high-confidence leachate occurrence, values close to 0 indicate high-confidence non-leachate conditions, whereas intermediate values identify uncertain transitional zones. These maps therefore complement the hard classifications by providing spatial confidence information.

2.4. Support vector machine (SVM)

Support Vector Machine (SVM) is a powerful classifier that performs well even with small training datasets (Cortes and Vapnik, 1995; Burges, 1998) and it is manly used in the field of waste management to waste classification (Chinassylov et al., 2022; Yang et al., 2023), waste generation prediction (Cha et al., 2021). The core concept behind SVM is to find the optimal hyperplane that best separates data points of different classes by maximizing the margin between them. Considering a binary classification problem with training data points x_i and labels $y_i \in \{-1, +1\}$, the separating hyperplane can be defined as the set of points that satisfy the following equation:

$$\mathbf{w} \bullet \mathbf{x} + b = 0 \tag{5}$$

where $\mathbf{w}$ is the weight vector perpendicular to the hyperplane, and b is the bias term.

The goal is to maximize the margin, which is the distance between the closest points (support vectors) and the hyperplane, while correctly classifying the training data. This leads to the optimization problem:

$$\frac{1}{2}\|\mathbf{w}\|^2 \tag{6}$$

subject to the constrain:

$$y_i(\mathbf{w} \bullet x_i + b) \geq 1 \forall i \tag{7}$$

To solve this, the Lagrangian dual formulation is used, introducing Lagrange multipliers α_i which are nonzero only for support vectors. The decision function $g(\mathbf{x})$ for predicting the class of a new point $\mathbf{x}$ is:

$$g(\mathbf{x}) = \mathrm{sign}\left(\sum\nolimits_i \alpha_i y_i K(x_i, \mathbf{x}) + b\right) \tag{8}$$

where $K(\bullet, \bullet)$ is a kernel function that enables the algorithm to handle nonlinear separation by implicitly mapping data to higher-dimensional feature spaces (Scholkopf and Smola, 2001). In this study, we used the linear kernel that computes the direct dot product $K(x_i, x_j) = x_i^T \bullet x_j$ finding the separating hyperplane in the original feature space. We selected the linear kernel as it is the simplest, and also provides the fastest computation and the maximum interpretability, requiring no additional parameter tuning. The straight lines separating red-blue regions in the plots of Section 3 exactly represents the linear boundary defined by eq. (5). For SVM, posterior probabilities were obtained by calibrating the decision function of eq. (8) through a sigmoid transformation (Platt, 1999).

Both the LDA and SVM algorithms were numerically implemented using MATLAB built-in functions (*fitcdiscr* and *fitcsvm*, respectively).

2.5. Benchmark classifiers and classification metrics

To assess the performance of LDA and SVM, two additional benchmark classifiers were tested: Logistic Regression (LOG) and Random Forest (RF). Logistic Regression (Hastie et al., 2017) represents a probabilistic linear classifier that provides a direct comparison with LDA and linear SVM, whereas Random Forest (Breiman, 2001) is a nonlinear ensemble method capable of modelling more complex interactions among predictors. These additional tests were not intended to replace the proposed framework, but rather to assess whether increased model complexity yields tangible improvements under sparse well-constrained training conditions.

The effectiveness of the two proposed methods (LDA, SVM) as wells as of the benchmarks (LOG and RF) were tested through the use of standard metrics, which can provide reliable benchmarks for quantifying classifier effectiveness, especially in imbalanced cases like leachate detection, beyond visual inspection (Sokolova and Lapalme, 2009; Powers, 2011). In this study, classification performance was evaluated using 4 widespread used standard metrics (Hastie et al., 2017): *accuracy*, defined as the proportion of correct predictions across all samples; *precision*, the ratio of true positive predictions to all positive predictions (minimizing false positives); *recall* (or sensitivity), the fraction of actual positives correctly identified (minimizing false negatives); and *F1-score*, the harmonic mean of precision and recall, which balances both when classes are imbalanced. These are formally expressed as:

$$\begin{aligned} accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \\ precision &= \frac{TP}{TP+FP} \\ recall &= \frac{TP}{TP+FN} \\ F1-score &= 2\frac{precision \bullet recall}{precision+recall} \end{aligned} \tag{9}$$

where *TP*, *TN*, *FP*, and *FN* denote true positives, true negatives, false positives, and false negatives, respectively.

3. Results

We present the results of our analyses performed using 2-D parameter spaces (ρ and *m*; ρ and *Mn; m and Mn*), and a 3-D parameter space (ρ, *m* and *Mn*). Although *Mn* is mathematically related to ρ and *m*, it is retained as an independent predictor in the 3D parametrization because it often enhances contamination sensitivity by combining conductive and polarization responses into a single physically meaningful descriptor, as previously experimented also for unsupervised approaches (Piegari et al., 2023; De Donno et al., 2025).

Both LDA and SVM algorithms are supervised techniques and they need a training phase using labelled (well) data, then extending the results to the whole ERT/IP dataset. We used geophysical and well data acquired in correspondence of the P1-P5 piezometers (Fig. 3), where P1, P3 and P5 (see Fig. 2 for location) are chosen for training the two algorithms, while P2 and P4 (see Fig. 2 for location) as validation wells to check a posteriori the reliability of the presented procedure. The motivation of the choice of the training wells is twice: i) they are representative of the different saturation conditions encountered in the landfill, passing from a significant leachate accumulation (P5, leachate level = 11.2 m), to a residual level of tens of cm (P3) and dry conditions (P1); ii) their location (Fig. 1) spans the whole investigated area from the eastern to the western part as well as upstream (P5) and downstream (P1 and P3). On the other hand, P2 and P4 were left for validation since they are also representative of the dry (P2) and leachate-saturated (P4, leachate level = 8.1 m) conditions.

3.1. 2-D parameter spaces

Using resistivity (ρ) and chargeability (*m*) as parameters (Fig. 4a), the two-dimensional LDA and SVM classifiers reproduce the training data from wells P1, P3 and P5 with a clear separation between dry and leachate-saturated conditions (Fig. 4b and d). Dry samples cluster around lower chargeability values, whereas saturated samples occupy a distinct region at higher *m*, and the learned decision boundaries correctly assign class 0 (dry) and class 1 (saturated) to most of the training points. When applied to the full set of geophysical data (Fig. 4c and e), both classifiers generate spatially coherent classifications, with predicted dry and saturated zones consistent with the independent information from validation wells P2 and P4, though with some exceptions, with LDA generating 1 FP (P2 – Fig. 4c) and SVM 2 FN (P4 – Fig. 4e).

A slightly different behaviour is observed when ρ and normalized chargeability (*Mn*) are used as model parameters (Fig. 4f). In the training space, defined by wells P1, P3 and P5, LDA and SVM again succeed in separating dry (low *Mn*) from saturated (high *Mn*) samples and in correctly assigning the binary labels 0 and 1 (Fig. 4g and i). However, in this case, the projection of these classifiers onto all ERT/IP data yields more robust maps of leachate occurrence, in which the predicted classes align well with the responses measured at the validation wells P2 and P4. In fact, only 1 FP is generated by both methods without FN, confirming that the ρ–*Mn* parametrization better enhance the discrimination between dry and saturated conditions using SVM (Fig. 4j), while no significant differences are seen for LDA (Fig. 4h) with respect to the ρ-*m* pair.

Mapping the results of the classification as a three-dimensional view within the landfill of the accumulation zones (Fig. 5, dark red areas), we identified a similar geometry of the saturated areas depending only mildly on the choice of classifier. Using ρ and *m*, some differences in the lateral extent of saturated pockets are clearly visible between LDA and SVM (Fig. 5a and b), due to the different performance of the two algorithms. However, the reconstruction with *Mn* replacing *m* (Fig. 5c and d), produced very similar 3-D patterns of leachate accumulations for LDA and SVM, as expected owing to the similar good performances highlighted by the two methods using the ρ-*Mn* pair.

To further assess the role of polarization-related parameters, an additional classification was performed using the (*m*, *Mn*) parameter space (Supplementary Fig. S1). In the training domain, both LDA (Fig. S1b) and SVM (Fig. S1d) achieve a clear separation between dry and leachate-saturated samples, with decision boundaries consistent with those observed in the other parameter spaces. When extended to the full geophysical dataset (Fig. S1c and S1d respectively for LDA and SVM) the resulting spatial classifications remain coherent and consistent with the independent validation wells P2 and P4, showing only minor discrepancies between predicted and observed conditions. In particular, LDA preserves its characteristic inclusive behaviour, identifying slightly larger saturated regions, whereas SVM produces a more conservative delineation, limiting the classification to areas with stronger polarization signatures.

3.2. 3-D parameter space

Using ρ, *m*, and *Mn* as three-dimensional model parameters (Fig. 6a), LDA and SVM classifiers effectively separate the training data from wells P1, P3, and P5 (Fig. 6b and d), with dry conditions (crosses) distinctly clustered from saturated ones (circles). The validation wells P2 (black points) and P4 (green points) align well with the predicted classes across the full geophysical dataset (Fig. 6c and e), displaying only 1 FP for LDA and 1 FN for SVM.

The three-dimensional projection of these classifications in the real-world space (Fig. 7a–b) reveals coherent leachate accumulations as dark red areas within the landfill. Both LDA and SVM using ρ, *m*, and *Mn* produce similar subsurface patterns of saturated zones, with the main

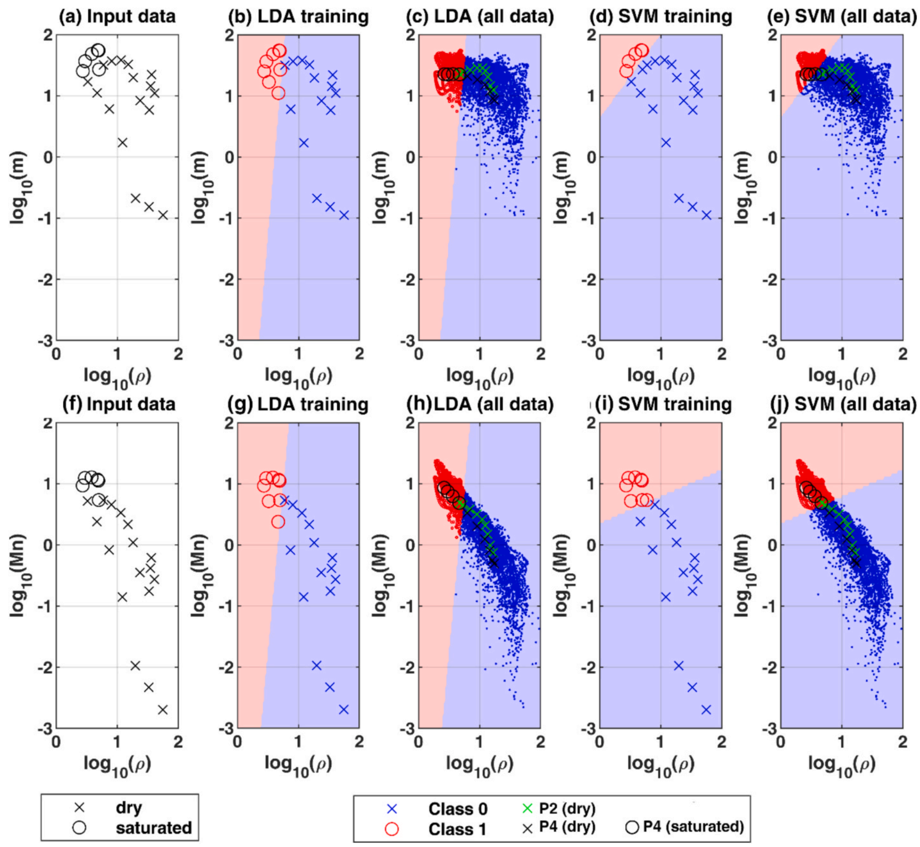

Fig. 4. Classifications achieved using ρ and m (a-e) and ρ and Mn (f-j) as model parameters: (a, f) input data for training (P1, P3, P5 wells), (b, g) LDA training results, (c, h) LDA classification for all geophysical data, (d, i) SVM training results, (e, j) SVM classification for all geophysical data. The dry and saturated values are marked with cross points and circles respectively. The predicted classes (0-dry, 1-saturated) are marked with blue and red points respectively, while the values from the validation wells are marked with black (P2) and green (P4) points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

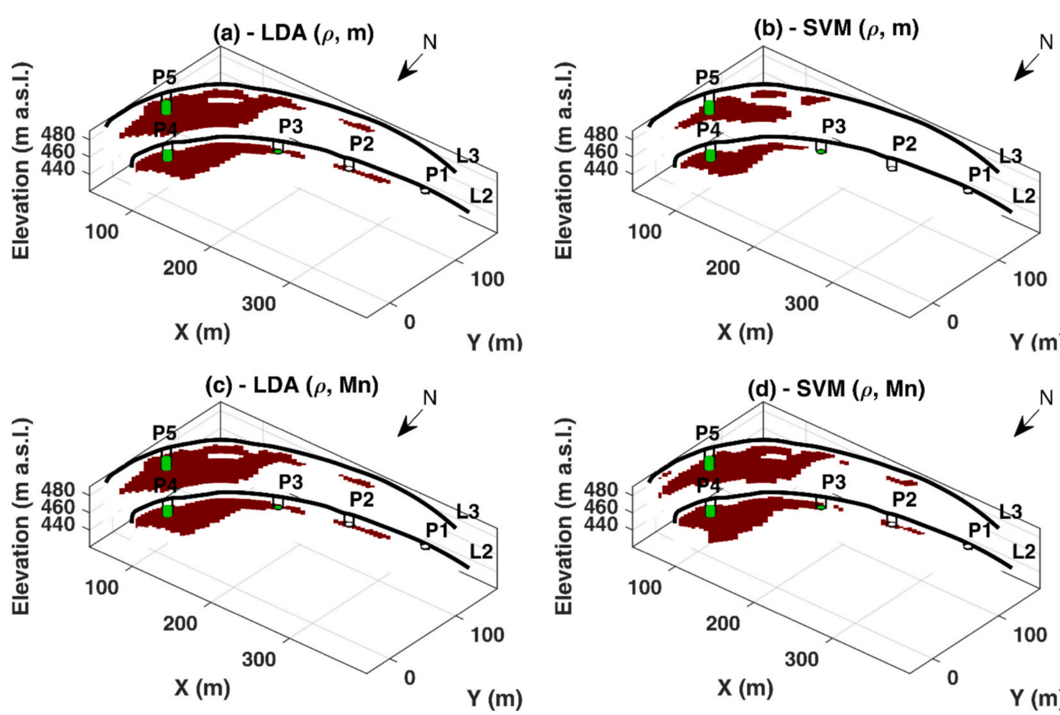

Fig. 5. 3D reconstruction of the leachate accumulations in the investigated landfill (dark red areas): (a) LDA classification using ρ and m, (b) SVM classification using ρ and m, (c) LDA classification using ρ and Mn, (d) SVM classification using ρ and Mn. Saturated wells are represented by green-filled symbols, while dry wells are represented by white-filled symbols. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

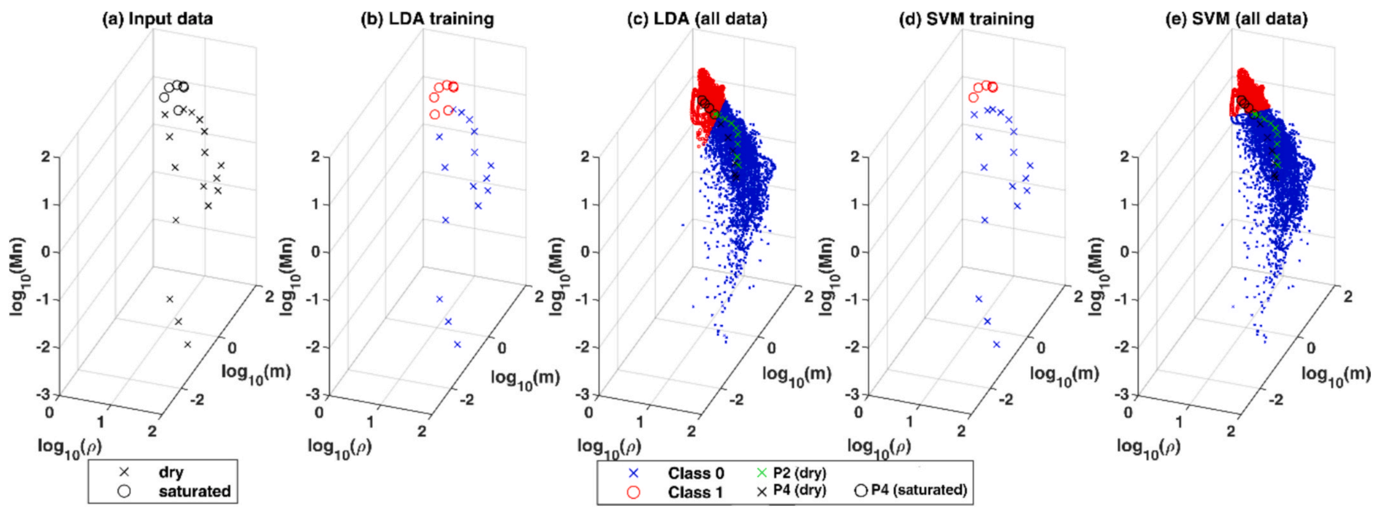

Fig. 6. Classifications achieved using ρ, m and Mn as model parameters: (a) input data for training (P1, P3, P5 wells), (b) LDA training results, (c) LDA classification for all geophysical data, (d) SVM training results, (e) SVM classification for all geophysical data. The dry and saturated values are marked with cross points and circles respectively. The predicted classes (0-dry, 1-saturated) are marked with blue and red points respectively, while the values from the validation wells are marked with black (P2) and green (P4) points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

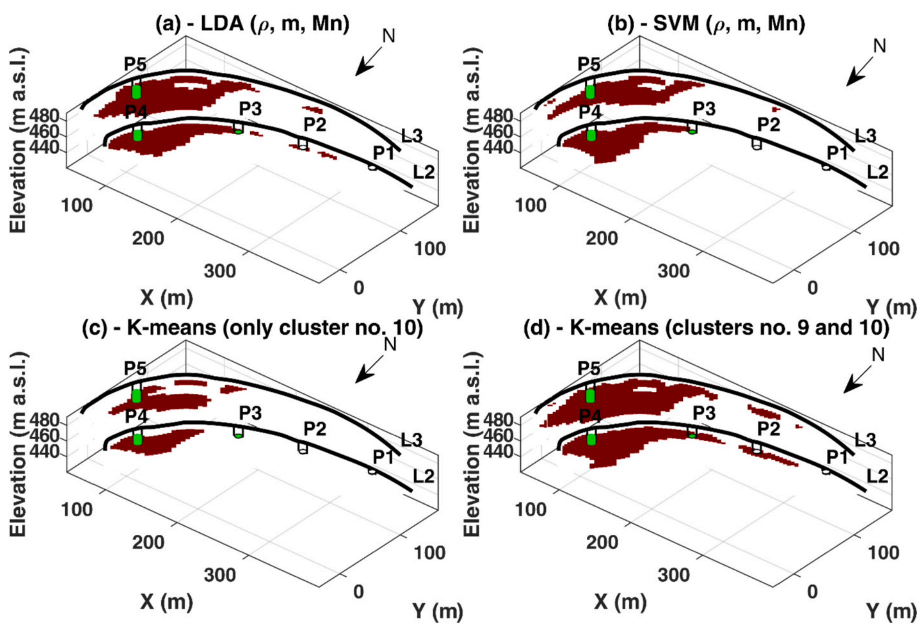

Fig. 7. 3D reconstruction of the leachate accumulations in the investigated landfill (dark red areas): (a) LDA classification using ρ and m and Mn, (b) SVM classification using ρ and m and Mn, (c) K-means clustering considering only the cluster no. 10 after Piegari et al. (2023), (d) K-means clustering considering clusters no. 9 and 10 after Piegari et al. (2023). Saturated wells are represented by green-filled areas, while dry wells are represented by white-filled areas. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

differences being that SVM does not identify the lateral zones (around P2 and P3 wells) shown by LDA, though for profile L3 a shallow accumulation zone in the upper left zone of the profile is detected contrary to LDA.

The reconstruction derived from the 3D closely resembles that obtained using the full (m, Mn) pair (Fig. S1), both in terms of geometry and spatial extent of the saturated zones. This result indicates that polarization-related attributes alone retain a significant discriminatory capability for leachate detection, even in the absence of explicit resistivity information. However, minor differences in the lateral continuity and depth extent of the mapped zones suggest that resistivity still provides complementary constraints on bulk conductive pathways.

3.3. Classification performance

For each configuration, classifier performance was evaluated in terms of accuracy, precision, recall and F1-score Eq. (9) on both the training and validation sets both for the two main algorithms (LDA and SVM) and the two benchmark methods (RF and LOG). The complete set of classification metrics achieved for all parameter spaces and classifiers (including RF and LOG) is reported in Supplementary Material (Table S1).

Using the ρ–m parameter pair, LDA achieved values on the training dataset (P1, P3 and P5 wells) indicating that it correctly retrieved all leachate samples at the cost of few false positives. In contrast, the SVM obtained higher accuracy, precision, and F1-score, but lower recall, meaning that it always correctly predicted leachate where it is present, but it missed approximately one-sixth of the leachate samples. On the corresponding validation set (P2 and P4 wells), LDA maintained similar performance, which points to good generalization and no evident overfitting. The SVM, instead, showed a stronger trade-off: in the most conservative case, it preserved perfect precision (100%) on validation, but its recall dropped to 50%.

Using the ρ–*Mn* pair, both LDA and SVM classifiers achieved identical performance metrics on the training data and validation set, indicating equivalent good effectiveness in separating leachate from dry conditions.

In contrast, for the *m-Mn* pair as well as with the full set of parameters ρ, *m* and *Mn*, the metrics diverged between LDA and SVM, mirroring patterns seen in prior analyses. Both classifiers maintained high accuracy on training data, but LDA consistently achieved perfect recall while SVM prioritized perfect precision at the expense of lower recall. On validation sets, this trade-off persisted, with LDA sustaining robust generalization across all metrics and SVM exhibiting reduced recall despite its precision advantage.

The benchmark classifiers (LOG and RF) do not provide consistent improvements over LDA and SVM, particularly under the limited size of the training dataset. This confirms that simpler, well-constrained linear classifiers are more robust and suitable in this context.

3.4. Uncertainty estimation

To complement the binary hard-classification results, posterior probabilities of leachate occurrence were computed for the LDA (Fig. 8a) and SVM (Fig. 8b) using Eqs. (10) and (11) respectively, and compared with the fuzzy membership values obtained from the previous unsupervised fuzzy C-means (FCM) clustering analysis (Fig. 8c).

Both LDA and SVM highlight the same main leachate accumulation sectors previously identified in the deterministic maps, namely the western-central portion of the landfill near wells P4–P5 and, secondarily, the central sector around P3. However, relevant differences emerge in the confidence patterns. LDA (Fig. 8a) produces smoother and more spatially continuous probability distributions, reflecting its probabilistic linear-discriminant formulation. In contrast, SVM (Fig. 8b) yields sharper high-probability cores surrounded by abrupt probability gradients, consistent with its margin-based decision mechanism.

The comparison with FCM membership values pertained to cluster no. 10 only (Fig. 8c), interpreted as the cluster most representative of leachate-affected conditions, shows a generally consistent spatial pattern, though more fragmented compared to the supervised approaches. The combined fuzzy membership of clusters no. 9 and 10 (Fig. 8d) produces a more spatially continuous pattern compared to cluster no. 10 alone; however, low membership values are observed along the transition zones between the two clusters, highlighting areas of weak affiliation and increased ambiguity in the delineation of leachate occurrence.

For the sake of completeness, posterior probability maps were also computed using the benchmark classifiers LOG and RF (Supplementary Fig. S2). The LOG-derived probability distribution (Fig. S2a) appears spatially smooth, likewise LDA (Fig. 8a) but even with a higher uncertainty and a limited contrast between high-confidence and transitional zones. In contrast, the RF model (Fig. S2b) produces high-probability patches with a reduced spatial continuity, similarly to SVM (Fig. 8a). Overall, compared to LDA and SVM (Fig. 8a,b), both benchmark classifiers do not yield significant advantages in terms of identification of coherent leachate accumulation zones or delineation of transitional areas.

4. Discussion

4.1. Comparison with other ML-based approaches

In our previous studies the unsupervised clustering of ERT/IP models using K-means (Piegari et al., 2023), and fuzzy C-means (De Donno et al., 2025) carried out on the same dataset, led to the identification

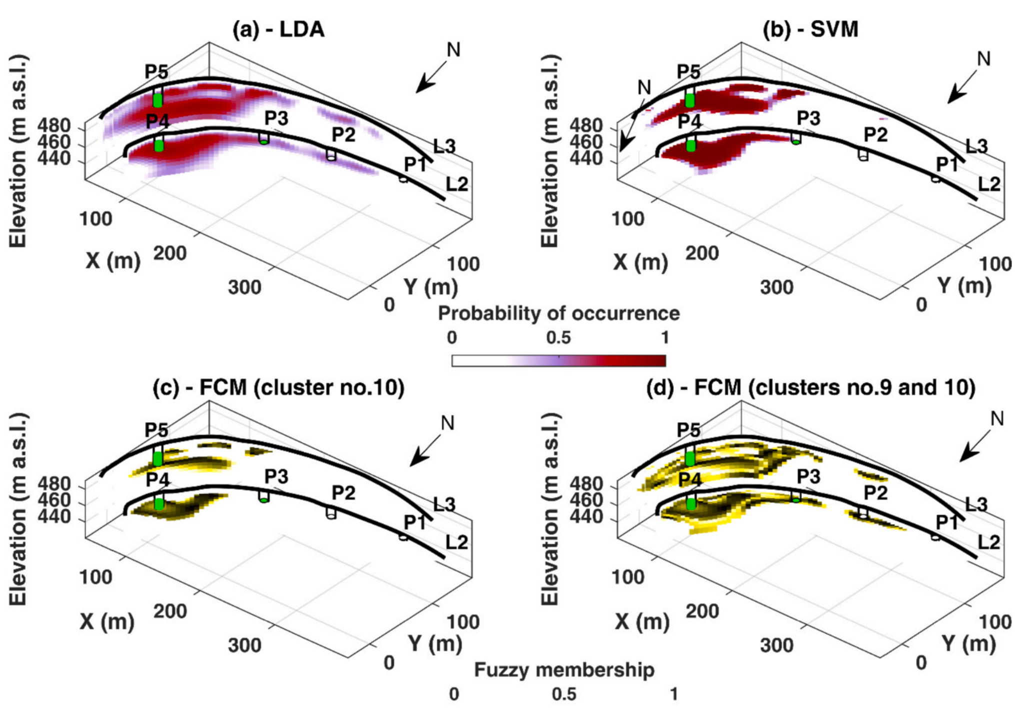

Fig. 8. Uncertainty assessment of leachate occurrence: (a) posterior probability map derived from LDA; (b) posterior probability map derived from SVM after score calibration; fuzzy membership values for (c) cluster no. 10, and (d) clusters no. 9 and 10 obtained from the previous FCM analyses (De Donno et al., 2025). Saturated wells are represented by green-filled symbols, while dry wells are represented by white-filled symbols. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of multiple classes, interpreted as different degrees of leachate saturation. Here, we compare the leachate-affected volumes for the two approaches (supervised and unsupervised) in terms of percentage ratio of pixels (PRP) belonging to the saturated area with respect to the total number of pixels (Table S2 in Supplementary Material). It is worth noting that the hard unsupervised approaches assign every point in the geophysical sections to a cluster, reproducing a full cross-section coloured by cluster indices. In contrast, the supervised approach, with its binary leachate/no-leachate classification, yields cross-sections, where only areas corresponding to leachate accumulation zones are depicted. Since two clusters were previously identified as indicative of the presence of leachate (i.e., clusters no. 9 and no. 10 after Piegari et al., 2023), two possible comparisons are made considering pixels belonging only to the cluster no. 10 and to the combination of clusters no. 9 and 10. The PRP values quantifying the relative extension of leachate accumulation zones are 20.10% for LDA and 10.35% for SVM, considering ρ and m (see Fig. 5a and b), 20.10% and 18.70% for ρ and Mn, (see Fig. 5c and d), while it is respectively to 18.29% and 14.71% for the 3D analysis (see Fig. 7a and b). Therefore, the PRP values obtained using the 3D parametrization are generally intermediate between the more inclusive LDA estimates and the conservative SVM ones derived from 2D spaces, suggesting that the combined use of ρ, m and Mn provides a more balanced reconstruction of leachate extent. From the previous unsupervised analyses, PRP is 10.68% considering only cluster no. 10, while it is significantly increased to 23.92% also considering cluster no. 9 (Fig. 7c and d). We note that the area identified by cluster no. 10 alone in the unsupervised analysis matches almost perfectly with the one resulting from the supervised SVM analysis with ρ and m (see Fig. 5b and Fig. 7c). Since SVM achieves 100% precision (Table S1)—meaning that the dark-red areas depicted in the models are truly saturated—this supports prior clustering analyses, confirming that the dark red cluster corresponds to leachate accumulation zones. Moreover, the combined area of clusters no. 9 and no.10 exceeds the less conservative LDA estimate (i.e., the largest accumulation area) for both ρ-m and ρ-Mn pairs. This result suggests that leachate accumulation lies between the two estimates, similar to LDA and SVM results in the 3D ρ-m-Mn parameter space. In terms of classification metrics, the unsupervised approaches (K-means/FCM) show a more variable behaviour compared to supervised classifiers (Table S1). In fact, cluster no. 10 yields high precision (100%) but limited recall (50%) and F1-score (66.7%), indicating a conservative identification of leachate occurrence. Conversely, the combined clusters no. 9 and 10 reach 100% recall but with lower precision (57.1%) and accuracy (84.2%), leading to an overestimation of affected areas. This trade-off contrasts with the more balanced performance of supervised methods, where LDA maximizes recall and SVM precision while maintaining consistently high accuracy (∼95%), resulting in a more reliable delineation of leachate distribution. Therefore, although the classification metrics achieved by unsupervised classification are satisfactory, the supervised approach is preferable, since it is not only able to better reproduce the levels in P4 and P5 but also in P3 (Fig. 7). Additionally, even though the subjectivity inherent in the choice of the number of clusters can be minimized using well-established indices, the choice of the optimal number of clusters remains a significant drawback for unsupervised techniques.

In terms of uncertainty representation, FCM provides numerical values (Fig. 8c,d) that remain purely data-driven and sensitive to local variability, lacking direct linkage to independent ground truth. In fact, it is important to note that fuzzy membership values represent degrees of similarity to cluster centroids rather than true probabilities of leachate occurrence (Garcia-Dias et al., 2020). Unlike supervised classifiers, where posterior probabilities are directly conditioned on labelled well data, fuzzy clustering does not incorporate ground-truth information, and therefore its membership values cannot be interpreted in a probabilistic sense (De Donno et al., 2025). In contrast, posterior probabilities derived from LDA and SVM (Fig. 8a,b) or from the benchmarks classifiers RF and LOG (Figs. S2a, b) are conditioned by well-constrained labels, thus offering a more physically grounded estimate of leachate occurrence. Although FCM patterns (Fig. 8c) are broadly consistent, they appear more fragmented, whereas LDA and LOG produce smoother probability fields while SVM and RF highlights sharper high-confidence cores. This behaviour is further clarified by the vertical profiles at the validation wells (Fig. S3), which provide a direct comparison between supervised posterior probabilities and unsupervised membership values along depth. At the dry well P2 (Fig. S3a), most supervised classifiers assign near-zero probabilities across the profile, in agreement with the absence of leachate, although LDA shows a limited false positive at depth. At the saturated well P4 (Fig. S3b), LDA captures perfectly the transition across the leachate level with a probability always higher than 50% within the saturated area, while SVM produces a sharper increase of probability but a more conservative classification, missing one saturated sample. RF exhibits a behaviour broadly comparable to SVM, whereas LOG shows more unstable trends and reduced agreement with the observed leachate level. The FCM membership profiles (grey solid lines in Figs. S3c, d) associated to cluster no. 10, appears similar to SVM and RF more distributed and less sharply constrained by the hydrogeological boundary. In contrast, the membership profiles associated with cluster no. 9 (grey dashed lines in Figs. S3c, d) display lower and more diffuse values across both dry and saturated intervals, indicating that this cluster primarily reflects transitional conditions and contributes to a less selective and more ambiguous delineation of leachate-affected zones. Therefore, supervised approaches do not replace unsupervised methods but complement them: clustering analyses can remain valuable for a preliminary assessment—especially in the absence of wells, where they provide the most powerful objective means to identify potential leachate accumulation areas— whereas supervised classifiers can then be used for operational and decision-making tasks.

A comparison between the proposed methods and the benchmark classifiers further highlights that, despite their different levels of complexity, neither LOG nor RF provides consistent improvements over LDA and SVM in terms of classification performance (Table S1) or spatial reconstruction of leachate occurrence (Fig. S2). For instance, in the 3D parametrization, LOG achieves 89% accuracy but only 50% recall on validation data, whereas RF attains 94.7% accuracy with 75% recall, both remaining inferior to LDA (100% recall), while RF performances are comparable to SVM. Similar trends are observed in the PRP values (Table S2), where RF (14.9%) is close to SVM, while LOG (10%) yields the smallest reconstructed leachate area, without improving the delineation of coherent saturated zones. Overall, these results indicate that, under sparse well-constrained training conditions, simpler linear classifiers such as LDA and SVM provide a more reliable balance between interpretability and predictive performance, without the need for increased model complexity.

4.2. Differences between LDA and SVM in the parameter spaces

The different behaviours observed for LDA and SVM across the investigated parameter spaces can be directly interpreted in terms of the geometry of their decision boundaries and their underlying assumptions. The LDA method assumes that each class follows a multivariate normal distribution with equal covariance matrices, leading to a linear decision boundary that maximizes class separation in a global sense (Hastie et al., 2017). As a consequence, LDA tends to produce smoother and more inclusive class regions in the parameter space, which explains its systematically high recall values achieved for all the parametrizations (Table 1). As a result for leachate mapping, this translates into a tendency to classify transitional or borderline electrical responses as leachate-affected points, thus minimizing FN (always zero in our validation datasets), but at the cost of few FP (only 1 in our validation datasets).

The SVM algorithm, by contrast, focuses on maximizing the margin between classes and relies primarily on support vectors located near the class boundary (Hastie et al., 2017). Even when a linear kernel is used, the resulting decision boundary is controlled by a limited subset of training points, making the classifier less sensitive to the overall distribution of samples. This property explains the higher precision observed for SVM with respect to LDA and its more conservative identification of leachate zones. In practical terms, SVM preferentially classifies only those pixels whose electrical properties closely resemble the leachate signatures observed in wells, while excluding marginal zones characterized by intermediate resistivity or chargeability values.

From a geophysical perspective, the decision boundaries identified by LDA and SVM in the different parameter spaces represent quantitative combinations of conductivity and polarization responses that discriminate between leachate-affected and non-affected materials. Leachate occurrence is typically associated with lower resistivity values, reflecting increased pore-fluid conductivity, and moderate to high chargeability responses, related to enhanced interfacial polarization processes within the waste mass (Flores-Orozco et al., 2020). Accordingly, the classifier boundaries can be interpreted as data-driven thresholds that jointly capture these contrasting electrical signatures, rather than relying on arbitrary single-parameter cutoff values.

The inclusion of normalized chargeability (*Mn*) plays a key role in reducing discrepancies between the LDA and SVM classifiers in waste characterization. By partially decoupling surface polarization effects from bulk conduction (Slater and Lesmes, 2002), *Mn* improves the separability of waste classes and reduces overlap in the parameter space. This is clearly proven using the ρ–*Mn* pair, where LDA and SVM converge toward nearly identical performance metrics and spatial reconstructions.

An additional sensitivity test using the (m, *Mn*) pair space further supports this interpretation (Table S1). In fact, despite excluding resistivity, the SVM classifier achieved validation scores comparable to the full three-parameter model, indicating that polarization-related attributes alone retain substantial discriminatory power for leachate detection. This suggests that chargeability responses are strongly linked to waste saturation and electrochemical activity, whereas resistivity mainly provides complementary information on bulk conductive pathways. In the full three-parameter space (ρ, m, *Mn*), which is selected also for enabling the comparison with the unsupervised approaches, the increased dimensionality slightly amplifies the intrinsic differences between the classifiers, with LDA remaining more inclusive and SVM maintaining its conservative nature.

Overall, these results indicate that parameter selection may be as important as classifier selection in landfill geophysical applications.

4.3. Practical implications for MSW landfills

The classification results carry direct implications for practical applications in landfill management, where the need for accurate leachate detection can guide the monitoring activities and the retrofitting interventions, particularly for landfills located on slopes. Beyond deterministic delineation, the uncertainty maps depicted in Fig. 8 provide an additional decision-support layer, highlighting zones of high confidence and transitional sectors where further investigations may be prioritized. Although the results of the 3D analyses (Fig. 7 and Table 1) show that LDA performs slightly better than SVM, which is more suitable for a conservative assessment of leachate accumulation, both approaches enable a spatially continuous imaging of leachate accumulation, overcoming the limitations of scattered well data, which may not capture the extreme spatial variability of leachate accumulations. This aspect is particularly relevant in heterogeneous landfill environments, where localized drilling alone may miss preferential leachate pathways or perched saturated zones. However, a sparse well network can be used as an input for running an ML supervised algorithm, whose results can then support the optimization of the monitoring network by suggesting the location of additional piezometers. For example, areas characterized by intermediate occurrence probabilities in Fig. 8a–b may represent the most informative targets for new piezometers or confirmation boreholes.

From a more geotechnical point of view, leachate accumulation strongly influences pore pressure and effective stress within the waste mass (Koerner and Soong, 2000). However, a crucial aspect that is often disregarded is the fact that the mechanisms and patterns of leachate flow in MSW landfills could be extremely complex (Zeiss and Uguccioni, 1994). Therefore, for slope landfills, the 3D reconstructions of leachate distribution achieved though the ML supervised methods can be directly incorporated into slope stability analyses as heterogeneous saturation inputs, unlike the commonly used method based only on interpolation of levels logged in wells, which fails to image the actual leachate distribution. A limitation of the present study is the relatively small number of wells available for supervised training. Therefore, extrapolation to other landfill sectors or different sites should be undertaken with caution, particularly in highly heterogeneous environments where spatial variability may affect model transferability. Although the validation results on the L2 and L3 profiles indicate robust performance within these sectors, the applicability of the proposed approach to independent profiles will be assessed in future work, including the untrained lines (L1 and L4 in this case), and/or incorporating additional wells, where available. Finally, although our models can only be rated as 2.5D, these results can pave the way for a future full three-dimensional campaign at the site, in order to provide an estimation of the leachate volumes, which is beyond the scope of this work.

5. Conclusions

This study successfully demonstrates the potential of supervised ML techniques (LDA and SVM) for effective leachate mapping in MSW landfills using geophysical ERT and IP data. By leveraging well data, both techniques provide a more accurate image of the complex leachate distribution within the landfill compared to classical standalone geophysical models or clustering methods. In particular, LDA excels in providing leachate accumulation areas with high recall, effectively identifying the full extent of leachate-affected zones, while SVM provides a more conservative definition of the leachate accumulation zones, prioritizing precision over recall. Both methods show strong consistency between training and validation sets, confirming that the supervised classification, established at the well points, can be extended to the whole geophysical dataset.

Our supervised classification strengthens the role of geophysical monitoring during the post-closure phase in a MSW landfill, particularly for slope landfills. In fact, the combination of geophysical and ML supervised techniques can be an effective tool for time-lapse monitoring, where precise and accurate leachate mapping is crucial for a real-time evaluation of accumulation zones, as well as for improving the quality of the monitoring network and planning remediation activities.

CRediT authorship contribution statement

Giorgio De Donno: Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Data curation. **Davide Melegari:** Writing – review & editing, Visualization, Formal analysis, Data curation. **Ester Piegari:** Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wasman.2026.115630.

Data availability

Data will be made available on request.

References

Abdallah, M., Talib, M.A., Feroz, S., Nasir, Q., Abdalla, H., Mahfood, B., 2020. Artificial intelligence applications in solid waste management: a systematic research review. Waste Manag. 109, 231–246. https://doi.org/10.1016/j.wasman.2020.04.057.

Ahmad, G., Aleem, F.M., Alyas, T., et al., 2025. Intelligent waste sorting for urban sustainability using deep learning. Sci. Rep. 15, 27078. https://doi.org/10.1038/s41598-025-08461-w.

Archie, G.E., 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. Trans 146, 54–62. https://doi.org/10.2118/942054-G.

Audebert, M., Clément, R., Moreau, S., Duquennoi, C., Loisel, S., Touze-Foltz, N., 2016. Understanding leachate flow in municipal solid waste landfills by combining time-lapse ERT and subsurface flow modelling – Part I: analysis of infiltration shape on two different waste deposit cells. Waste Manag. 55, 165–175. https://doi.org/10.1016/j.wasman.2016.04.006.

Bishop, C.M., Nasrabadi, N.M., 2006. In: Pattern Recognition and Machine Learning. Springer, New York, p. 738.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. 2, 121–167. https://doi.org/10.1023/A:1009715923555.

Cha, G., Moon, H., Kim, J., 2021. A method to improve the performance of support vector machine regression models for predicting demolition waste generation using categorical principal components analysis. Int. J. Sustain. Build. Technol. Urban Dev. 12 (3), 282–294. https://doi.org/10.22712/susb.20210023.

Chen, F., Zhou, A., Ye, L., Li, W., Sun, X., Zhang, J., Zhao, Y., Huang, Q., Xu, Y., 2026. A new paradigm for electrical modeling in site leakage diagnostics: a label-scarce three-dimensional simulation surrogate using physics-informed neural networks. Water Res., 125456 https://doi.org/10.1016/j.watres.2026.125456.

Chinassylov, D., Kozhamseitova, A., Kalen, M., Omirgaliyev, R., 2022. Application of information modeling and machine learning algorithms for waste classification using support vector machines. Sci. J. Astana IT Univ. 45–53. https://doi.org/10.37943/AITU.2021.52.74.005.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018.

Costanzo, V.L., Izarra-Jacome, A., Costanzo-Alvarez, V., Jacome, M., Trigo-Ferre, R., Jeffrey, M., Cheng, Y., Milagrosa, A., Amon, C.H., 2026. Machine-learning–derived hydrogeochemical facies to support geoelectrical interpretation of contaminant plumes. Environ. Process. 13, 3. https://doi.org/10.1007/s40710-025-00811-0.

De Donno, G., Cardarelli, E., 2017. Tomographic inversion of time-domain resistivity and chargeability data for the investigation of landfills using a priori information. Waste Manag. 59, 302–315. https://doi.org/10.1016/j.wasman.2016.11.020.

De Donno, G., Melegari, D., Paoletti, V., Piegari, E., 2024a. Electrical and electromagnetic prospecting for the characterization of municipal waste landfills: A review. In: Anouzla, A., Souab, S. (Eds.), Technical Landfills and Waste Management. Springer. https://doi.org/10.1007/978-3-031-52633-6_1.

De Donno, G., Cercato, M., Melegari, D., Paoletti, V., Penta de Peppo, G., Piegari, E., 2024. Fuzzy mapping of integrated multi-parameter tomographic data to detect and localize geophysical targets. IEEE Trans. Geosci. Remote Sens. 62, 5924511. https://doi.org/10.1109/TGRS.2024.3445462.

De Donno, G., Melegari, D., Paoletti, V., Piegari, E., 2025. An integrated study of hard and soft cluster analyses for detecting leachate in a MSW landfill site using geoelectrical data. Waste Manag. 195, 22–31. https://doi.org/10.1016/j.wasman.2025.01.034.

Di Maio, R., Fais, S., Ligas, P., Piegari, E., Raga, R., Cossu, R., 2018. 3D geophysical imaging for site-specific characterization planning of an old landfill. Waste Manag. 76, 629–642. https://doi.org/10.1016/j.wasman.2018.03.004.

Fang, B., Yu, J., Chen, Z., Osman, A.I., Farghali, M., Ihara, I., Hamza, E.H., Rooney, D.W., Yap, P.S., 2023. Artificial intelligence for waste management in smart cities: a review. Environ. Chem. Lett. 21, 1959–1989. https://doi.org/10.1007/s10311-023-01543-6.

Farghali, M., Osman, A.I., 2024. Revolutionizing waste management: Unleashing the power of artificial intelligence and machine learning. In: Vambol, V., Vambol, S., Khan, N.A., Mozaffari, N., Mozaffari, N. (Eds.), Advances in Energy from Waste. Woodhead Publishing, pp. 225–279. https://doi.org/10.1016/B978-0-443-13847-8.00007-5.

Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. Ann. Eugen. 7, 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

Flores-Orozco, A., Gallistl, J., Steiner, M., Brandstätter, C., Fellner, J., 2020. Mapping biogeochemically active zones in landfills with induced polarization imaging: the Heferlbach landfill. Waste Manag. 107, 121–132. https://doi.org/10.1016/j.wasman.2020.03.033.

Fraternali, P., Morandini, L., Gonzalez, S.L.H., 2024. Solid waste detection in remote sensing images: a survey. Waste Manag. 189, 88–102. https://doi.org/10.1016/j.wasman.2024.01.018.

Garcia-Dias, R., Vieira, S., Pinaya, W.H.L., Mechelli, A., 2020. Clustering analysis. In: Machine Learning. Academic Press, pp. 227–247. https://doi.org/10.1016/B978-0-12-815739-8.00013-4.

Guo, H., Wu, S., Tian, Y., Zhang, J., Liu, H., 2021. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: a review. Bioresour. Technol. 319, 124114. https://doi.org/10.1016/j.biortech.2020.124114.

Han, C., Zhao, R., Li, W., Liu, S., Meng, J., Li, J., Mao, D., 2025. A case study of chromium-containing solid waste deposit quantitation through time domain induced polarization signatures in a non-engineered waste disposal site. J. Appl. Geophys. 241, 105820.

Hastie, T., Tibshirani, R., Friedman, J., 2017. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer. DOI not available.

Isunza Manrique, I., Caterina, D., Nguyen, F., Hermans, T., 2023. Quantitative interpretation of geoelectric inverted data with a robust probabilistic approach. Geophysics 88 (3), B151–B165. https://doi.org/10.1190/geo2022-0326.1.

Juarez, M.B., Mondelli, G., Giacheti, H.L., 2023. An overview of in situ testing and geophysical methods to investigate municipal solid waste landfills. Environ. Sci. Pollut. Res. 30 (9), 24779–24789.

Juarez, M.B., Giacheti, H.L., do Nascimento, A.P., Barros, M.F.S., Elis, V.R., Mondelli, G., 2026. Geoelectrical imaging and cluster analysis for leachate mapping in a municipal solid waste landfill: A case study in Brazil. Waste Manag. 211, 115289. https://doi.org/10.1016/j.wasman.2025.115289.

Koerner, R.M., Soong, T.Y., 2000. Leachate in landfills: the stability issues. Geotext. Geomembr. 18 (5), 293–309. https://doi.org/10.1016/S0266-1144(00)00011-7.

Liu, B., et al., 2020. Deep learning inversion of electrical resistivity data. IEEE Trans. Geosci. Remote Sens. 58 (8), 5715–5728. https://doi.org/10.1109/TGRS.2020.2969040.

Lu, W., Chen, J., 2022. Computer vision for solid waste sorting: a critical review of academic research. Waste Manag. 142, 29–43. https://doi.org/10.1016/j.wasman.2022.01.020.

Luan, Z., Nai, C., Sun, C., Jiao, L., 2023. Monitoring leachate dispersion in landfills using electrical resistivity tomography inversion based on deep learning and data assimilation. In: Proc. 5th Int. Conf. Electronics and Communication, Network and Computer Technology (ECNCT), pp. 192–196. https://doi.org/10.1109/ECNCT59757.2023.10280964. Guangzhou, China.

Ma, X., Zhang, J., Schwartz, N., Li, J., Chao, C., Meng, J., Mao, D., 2024a. Characterizing landfill extent, composition, and biogeochemical activity using electrical resistivity tomography and induced polarization under varying geomembrane coverage. Geophysics 89 (4), 151–164.

Ma, X., Zhou, J., Li, J., Zhang, J., Han, C., Guo, L., Li, S., Mao, D., 2024b. Image-guided structure-constrained inversion of electrical resistivity data for improving contaminations characterization. Comput. Geosci. 185, 105545.

Manakkakudy, A., De Iacovo, A., Maiorana, E., Mitri, F., Colace, L., 2024. Waste material classification: a short-wave infrared discrete-light-source approach based on light-emitting diodes. Sensors 24 (3), 809. https://doi.org/10.3390/s24030809.

Martorana, R., Capizzi, P., D'Alessandro, A., Luzio, D., 2016. Electrical resistivity and induced polarization tomographies to test the efficiency and safety of the new landfill of Bellolampo (Palermo, Italy). Boll. Geofis. Teor. Appl. 57, 313–327.

McLachlan, G.J., 2004. Discriminant Analysis and Statistical Pattern Recognition. John Wiley & Sons.

Nwokediegwu, Z.Q.S., Ugwuanyi, E.D., Dada, E.A., Majemite, M.T., Obaigbena, A., 2024. AI-driven waste management systems: a comparative review of innovations in the USA and Africa. Eng. Sci. Technol. J. 5 (2), 507–516. https://doi.org/10.51594/estj/v5i2.828.

Piegari, E., De Donno, G., Melegari, D., Paoletti, V., 2023. A machine learning-based approach for mapping leachate contamination using geoelectrical methods. Waste Manag. 157, 121–129. https://doi.org/10.1016/j.wasman.2022.12.015.

Platt, J., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv. Large Margin Class. 10 (3), 61–74.

Powers, D.M.W., 2011. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *J. Mach. Learn. Technol.* 2(1), 37–63. DOI not available.

Scholkopf, B., Smola, A.J., 2001. In: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. The MIT Press, Cambridge, MA. https://doi.org/10.7551/mitpress/4175.001.0001.

Sharma, K., Rajendra, A., 2025. Remote-sensing applications for monitoring and analysis of waste disposal sites using satellite images: a review. Trans. GIS 29 (8), e70175. https://doi.org/10.1111/tgis.70175.

Slater, L.D., Lesmes, D., 2002. IP interpretation in environmental investigations. Geophysics 67 (1), 77–88. https://doi.org/10.1190/1.1451353.

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. 45 (4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002.

Soupios, P., Ntarlagiannis, D., Sengupta, D., Agrahari, S., 2017. Characterization and monitoring of solid waste disposal sites using geophysical methods: current applications and novel trends. In: Sengupta, D., Agrahari, S. (Eds.), Modelling Trends in Solid and Hazardous Waste Management. Springer, pp. 75–103.

Sun, X., Xu, Y., Nai, C., Liu, J., 2026. Three-dimensional inversion method based on multi-source fused physical information networks for leachate distribution in landfills. Comput. Geosci. 207, 106077. https://doi.org/10.1016/j.cageo.2025.106077.

Xia, T., Meng, J., Ding, B., Chen, Z., Liu, S., Titov, K., Mao, D., 2022. Integration of hydrochemical and induced polarization analysis for leachate localization in a municipal landfill. Waste Manag. 157, 130–140.

Yang, L., Xiang, Y., Li, Y., Bao, W., Ji, F., Dong, J., Chen, J., Xu, M., Lu, R., 2023. Identification and classification of recyclable waste using laser-induced breakdown spectroscopy technology. AIP Adv. 13, 075024. https://doi.org/10.1063/5.0149329.

Yuan, Q., Zoungrana, A., 2025. Systematic review of environmental and human health risk assessments in municipal solid waste management. Discover Sustain. 6 (1), 930. https://doi.org/10.1007/s43621-025-00930-1.

Zeiss, C., Uguccioni, M., 1994. Mechanisms and patterns of leachate flow in municipal solid waste landfills. *J. Environ. Syst.* 23(3), 247–270. DOI not available.

Zhang, Z., Yuan, Z., Hu, L., Coulon, F., Liu, H., Cheng, Z., Fei, X., 2024. Comprehensive geophysical, geotechnical, and geochemical assessments of an offshore landfill in Singapore. J. Hazard. Mater. 480, 135908. https://doi.org/10.1016/j.jhazmat.2024.135908.