Chemical Engineering Journal Advances 27 (2026) 101356

Contents lists available at ScienceDirect

Chemical Engineering Journal Advances

journal homepage: www.sciencedirect.com/journal/chemical-engineering-journal-advances

Synthetic noise injection asymmetrically affects machine-learning quantification of Cd^{2+} and Pb^{2+} in multi-scan cyclic voltammetry

Alya Kamilah [a], Riyanarto Sarno [a,*], Kelly Rossa Sungkono [a], Rizqy Ahsana Putri [a], Taufiq Choirul Amri [a], Fadlilatul Taufany [b], Arif Abdullah Sagran [c], Sang-Seok Lee [d]

[a] Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
[b] Department of Chemical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
[c] Faculdade de Ciências Políticas, Universidade Dili, Dili, 66000, Timor-Leste
[d] Graduate School of Engineering, Tottori University, Tottori, 680-8552, Japan

ARTICLE INFO

Dataset link: GitHub repository

Keywords:
Cyclic voltammetry
Heavy metals
Noise augmentation
Stochastic regularisation
Signal-to-noise ratio
Random Forest
Water pollution

ABSTRACT

Machine-learning quantification of toxic trace metals such as Cd^{2+} and Pb^{2+} from cyclic voltammetry (CV) reads concentration from peak features that a deterministic peak-finding step extracts from repeated scans, and its accuracy is limited by how stably those features can be recovered from noisy signals. The prevailing assumption treats acquisition noise as a nuisance to be removed before feature extraction. The present study tests the opposite hypothesis, namely that controlled noise injected before feature aggregation can improve quantification. A Random Forest regressor was applied to seventeen multi-scan peak features extracted from 1500 voltammograms spanning fifteen concentrations of each metal. Four noise distributions, namely Gaussian, Student's-t, pink $1/f^3$, and heteroscedastic, were injected at a peak-relative signal-to-noise ratio of 30 dB and evaluated over twenty seeds with exact Wilcoxon tests, Cohen's d_z, and bootstrap confidence intervals. Noise injection reduced the mid-range mean absolute percentage error (MAPE) for Cd^{2+} by 55.0 to 63.0% across all four distributions, with $p < 10^{-6}$ and d_z near -3, while the coefficient of determination remained near 0.96. Lead, whose voltammetric peaks are an order of magnitude sharper, showed no reliable benefit. Permutation feature importance, a nested-cross-validation check, a second tree-based regressor, a train-only control, and a synthetic-peak study together supported a peak-shape mechanism, in which injected noise perturbs deterministic peak finding on broad asymmetric peaks and renders across-scan dispersion features informative. Controlled acquisition noise therefore improves mid-range quantification for analytes with broad peaks, and the dissociation between a stable coefficient of determination and a large error reduction provides a diagnostic for peak-finding artefacts in related CV pipelines.

1. Introduction

Cadmium (Cd^{2+}) and lead (Pb^{2+}) contamination of surface and groundwater remains a persistent public-health concern, particularly in regions where industrial discharge, mining, and agricultural runoff intersect with limited monitoring capacity [1–4]. Both metals bioaccumulate along the food chain and exert pleiotropic toxic effects in humans: chronic Cd^{2+} exposure damages renal and skeletal systems, while Pb^{2+} disrupts neurological, hematological, and cardiovascular function, with no established safe threshold for either ion [5–10]. Recent syntheses of heavy-metal decontamination, process-safety, and environmental-management research underscore both the scale of the contamination problem and the demand for rapid, low-cost monitoring at the point of need [11–13]. Reference laboratory methods such as inductively coupled plasma mass spectrometry (ICP-MS) and atomic absorption spectroscopy (AAS) provide reliable quantification but require centralised infrastructure that limits their suitability for rapid field assessment [14,15]. The development of portable electrochemical sensors based on cyclic voltammetry (CV) and related stripping/voltammetric approaches offers a complementary route owing to low instrument cost, fast response, and compatibility with field deployment [16–20]. Working-electrode chemistry and modification strategies, particularly carbon-based and nanostructured electrodes, have been a major focus of this literature [21–23].

A growing body of work has applied machine learning to CV and related voltammetric signals to improve quantification beyond traditional univariate calibration [24–33]. Approaches range from feature-based

* Corresponding author.
E-mail address: riyanarto@its.ac.id (R. Sarno).

https://doi.org/10.1016/j.ceja.2026.101356
Received 30 April 2026; Received in revised form 7 July 2026; Accepted 8 July 2026
Available online 14 July 2026

regressors such as support vector regression and random forests [26,32,34] to end-to-end deep learning architectures including convolutional networks [29,31] and attention-based fusion models tailored to electrochemical signals [24,27]. A recurring pattern in these studies is the use of *single-scan* features even though several consecutive scans are typically acquired per measurement. The information contained in inter-scan variability is therefore underexploited as input to a regression model.

Electrochemical-noise studies have separately documented that the noise in voltammetric measurements departs from the ideal Gaussian assumption and includes $1/f^{\beta}$ spectral content, drift, and concentration-dependent variance [35–39]. Mitigation has traditionally focused on signal-side denoising via Savitzky–Golay smoothing [40,41], wavelet shrinkage [42,43], and adaptive filtering [44]. A complementary strategy from the machine-learning literature treats stochastic perturbations not as a nuisance to be removed but as a source of regularisation: training with noise is provably equivalent to Tikhonov regularisation [45], and explicit noise injection underpins dropout [46], time-series augmentation [47,48], adversarial training [49,50], and recent generalisations that mix noisy feature representations [51–53]. Whether such augmentation strategies transfer to feature-based regression on multi-scan voltammetric data — and how they interact with the underlying noise structure of CV measurements — has not, to the authors' knowledge, been systematically investigated.

The present work was initially framed around an evaluation of a multi-scan feature aggregation pipeline on a 1500-voltammogram Cd^{2+}/Pb^{2+} dataset, with the working hypothesis that aggregating features across scans would systematically reduce prediction error relative to single-scan extraction. An unexpected pattern emerged during the assessment of this pipeline's robustness to acquisition noise: noise injection prior to feature aggregation *improved* Cd^{2+} quantification accuracy substantially under multi-scan extraction, whereas Pb^{2+} accuracy was essentially unchanged.

The present paper reports this effect, validates it statistically, and offers a tentative interpretation. The main contributions are as follows.

- Characterisation of CV measurement noise that motivates the augmentation design. This study shows that the replicate noise in the dataset is non-Gaussian and coloured, namely it carries a $1/f^{\beta}$ power spectrum with β near 3 and a noise level that grows with concentration. This motivates testing coloured and amplitude-dependent noise alongside Gaussian rather than treating the choice of distributions as arbitrary (Section 3.1).
- Demonstration of an asymmetric noise-augmentation effect. This study shows that controlled noise injection prior to feature aggregation produces a large and statistically significant improvement in Cd^{2+} mid-range quantification accuracy across four distinct noise distributions, while producing no reliable benefit for Pb^{2+}, whose response is an order of magnitude weaker and survives neither multiple-comparison correction nor hyperparameter tuning. This finding is the central empirical contribution of the paper (Section 3.5).
- Sensitivity analysis across noise distributions and SNR levels. The present study provides a systematic SNR sweep, namely four noise distributions across six SNR levels and three random seeds, that maps the operating envelope of the augmentation effect and identifies pink $1/f^{3}$ noise as the dominant failure mode at low SNR (Section 3.4).
- Cross-model confirmation. The present study shows that the augmentation effect generalises beyond Random Forest to a second tree-based regressor and is weaker or absent in non-tree alternatives, a pattern consistent with the proposed mechanism (Section 3.10).
- A diagnostic dissociation between aggregate-fit and per-sample error metrics. The present study identifies a quantitative signature, namely R^2 remaining stable while MAPE drops markedly under noise injection, that may serve as a diagnostic for similar peak-finding artefacts in related CV-ML pipelines (Section 4.4).
- A mechanistic interpretation with empirical support. This study proposes a peak-shape-based interpretation linking the asymmetry to differences in deterministic peak finding for broader, less-Gaussian peaks, supported by a permutation feature-importance analysis before and after augmentation, a nested-cross-validation tuning check, a train-only control that localises the effect to acquisition time, and a controlled synthetic-peak study that reproduces the direction of the asymmetry from peak shape alone (Sections 3.6, 3.9, 3.11, and 4.1).

The remainder of the paper is organised as follows. Section 2 describes the dataset, feature extraction, noise injection scheme, prediction pipeline, and evaluation protocol. Section 3 reports the empirical noise characterisation (3.1), the SNR sensitivity analysis (3.4), the asymmetric augmentation finding with its statistical validation (3.5), and the cross-model confirmation (3.10). Section 4 discusses the proposed mechanism, the R^2/MAPE dissociation, practical implications, and limitations. Section 5 concludes.

2. Materials and methods

2.1. Reagents and electrochemical setup

$Cd(NO_3)_2$ and $Pb(NO_3)_2$ standard solutions were prepared in 0.1 M KCl supporting electrolyte at fifteen concentration levels (2, 4, 6, 8, 10, 20, 40, 60, 80, 100, 200, 400, 600, 800, and 1000 ppm). All reagents were of analytical grade, and deionised water with resistivity >18 MΩ cm was used throughout. The three-electrode configuration consisted of a 3 mm glassy carbon working electrode (GCE), a platinum counter electrode, and an Ag/AgCl reference electrode in saturated KCl. The GCE was polished with 0.05 μm alumina slurry prior to the measurement campaign. Between measurements, the electrode was cleaned by sequential rinsing with 0.1 M HCl and 0.1 M NaOH to remove residual analyte and adsorbed organic matter, followed by a final rinse with deionised water.

2.2. Cyclic voltammetric acquisition

Voltammograms were recorded using a PalmSens EmStat potentiostat (PalmSens BV, Houten, The Netherlands) controlled via PSTrace software. Scans were performed at a scan rate of 0.1 V/s over potential windows of −1.4 to −0.5 V vs. Ag/AgCl for Cd^{2+} and −1.0 to −0.1 V vs. Ag/AgCl for Pb^{2+}. No deposition or preconcentration step was applied prior to the cyclic potential sweep. Each measurement file contained ten consecutive CV scans of 180 data points each (after potential-window normalisation). Fifty replicate measurement files were acquired per concentration level, yielding 750 voltammograms per metal and 1500 in total. The first scan in each file was excluded from analysis owing to conditioning artefacts substantially larger than the steady-state response, leaving nine scans per file for feature extraction. The same dataset was previously analysed by [27,54].

2.3. Empirical noise characterisation

The selection of synthetic noise distributions for the SNR analysis was informed by the characterisation of measurement noise using replicate residuals, in accordance with established protocols from the electrochemical noise literature [35–37]. Distribution shape was assessed via excess kurtosis, quantile–quantile comparison against Gaussian and Student's-t references, and maximum-likelihood Student's-t fitting. Spectral content was estimated by Welch's method [55] on detrended residual series and fitted to a $1/f^{\beta}$ power-law model consistent with prior reports of $1/f$-type noise in electrochemical systems [38,39]. Heteroscedasticity was evaluated by linear regression in log–log space of the within-replicate noise standard deviation against concentration.

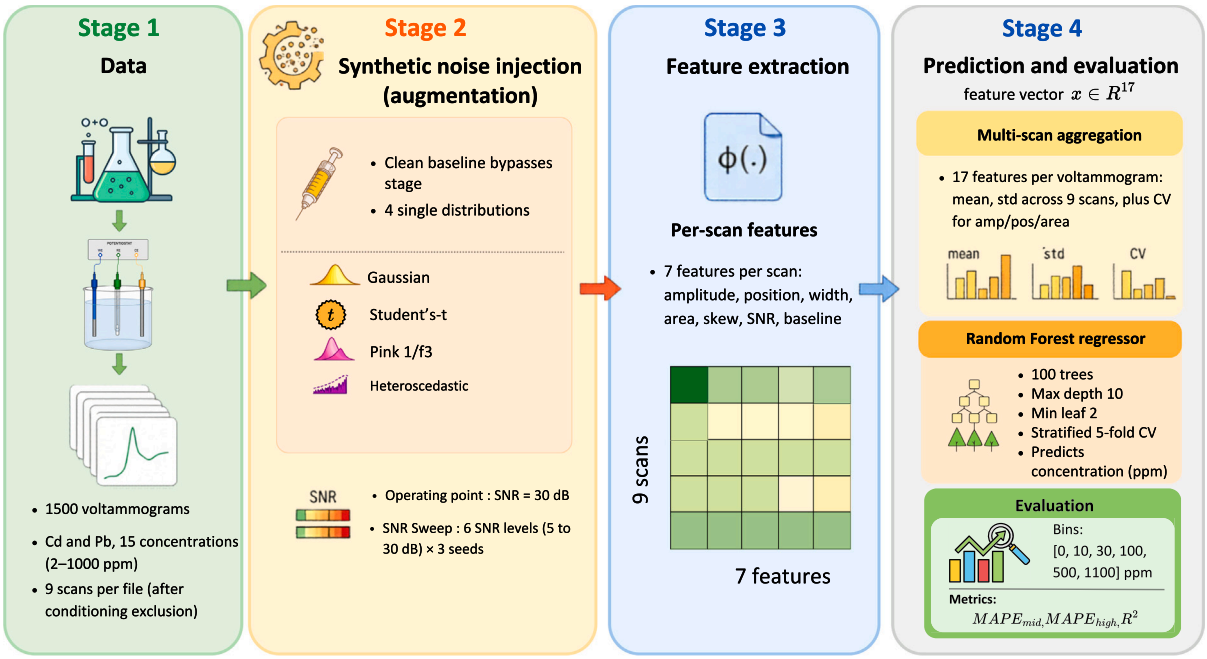

Fig. 1. Overview of the analysis pipeline in four stages. Stage 1 acquires the cyclic-voltammetry data, namely 1500 voltammograms for Cd^{2+} and Pb^{2+} across fifteen concentrations from 2 to 1000 ppm, with nine non-conditioning scans per measurement. Stage 2 optionally injects synthetic noise into the raw scans for augmentation, spanning four single distributions, namely Gaussian, Student's-t, pink $1/f^3$, and heteroscedastic, and their combinations, with the clean baseline bypassing this stage. Stage 3 extracts seven peak features for each scan. Stage 4 aggregates the per-scan features into a seventeen-dimensional vector, fits a Random Forest regressor under stratified five-fold cross-validation, and reports the coefficient of determination R^2 and the mean absolute percentage error (MAPE) per concentration band. Each stage is described in detail in the text.

2.4. Analysis pipeline overview

The complete analysis pipeline is summarised in Fig. 1 and proceeds in four stages. Stage 1 (data) acquires the cyclic voltammograms with the three-electrode cell and retains nine non-conditioning scans per measurement. Stage 2 (synthetic noise injection) optionally injects one of four single noise distributions, namely Gaussian, pink $1/f^3$, Student's-t, and heteroscedastic, or a combination of them, into the raw scans before feature extraction, and the clean baseline bypasses this stage. Stage 3 (feature extraction) computes seven peak features for each scan. Stage 4 (prediction and evaluation) aggregates the per-scan features into a seventeen-dimensional vector, fits a machine-learning regressor, namely a Random Forest as the primary model with XGBoost, SVR, and an MLP added for cross-model confirmation in Section 3.10, and reports the coefficient of determination R^2 and the mean absolute percentage error (MAPE) per concentration band with a statistical comparison between conditions. The remainder of this section describes each stage in detail.

2.5. Multi-scan feature extraction

The methodology of the study involved the extraction of seven peak-related features from the Savitzky-Golay smoothed signal for each scan [40,41] (window width 11, polynomial order 3): peak amplitude A, peak position E_p, full-width-at-half-maximum-derived σ, peak area, peak skewness, peak signal-to-noise ratio, and baseline value at the peak. The forward and reverse branches of each voltammogram were processed separately, and the branch with the larger peak amplitude within the analytical potential window was retained.

Across the nine non-conditioning scans, each base feature was summarised by its mean and standard deviation. The coefficient of variation was additionally computed for amplitude, position, and area, yielding a total of 17 features per voltammogram. These descriptors were prioritised because they summarise the analytically meaningful properties of a stripping peak, namely its height and area that scale with concentration, its position and width that encode redox identity and reversibility, its skewness and local signal-to-noise that capture peak distortion, and the baseline at the peak that captures background drift. The across-scan dispersion summaries, namely standard deviation and coefficient of variation, were retained because they quantify the run-to-run stability of these estimates, which the mechanism in Section 4.1 predicts to become informative under noise injection. The construction of the feature vector is illustrated in Fig. 2.

2.6. Synthetic noise injection

The present study sought to investigate the relationship between feature extraction performance and signal noise levels. Synthetic noise was therefore introduced artificially into the raw scans prior to the extraction of features. The peak-relative SNR was defined following IUPAC convention [56] as

$$\mathrm{SNR_{dB}} = 20\log_{10}\left(\frac{\text{peak height}}{\sigma_{\text{noise}}}\right), \tag{1}$$

where the peak height is the maximum baseline-corrected signal amplitude in the analytical peak region.

Four noise distributions were considered, each drawn independently for every scan and rescaled so that its standard deviation equals the target σ_{noise} implied by the chosen SNR:

1. Gaussian, independent zero-mean normal samples $\mathcal{N}(0, \sigma^2_{\text{noise}})$, the assumption underlying most denoising work.
2. Student's-t with $df = 1.5$, heavy-tailed samples drawn from a t distribution and rescaled to the target standard deviation, matching the near-Cauchy tails found in Section 3.1.
3. Pink ($1/f^3$ coloured), synthesised in the frequency domain by weighting a white spectrum by $f^{-\beta/2}$ with $\beta = 3$, assigning uniform random phases, applying an inverse real FFT, and rescaling to the target standard deviation, matching the empirical power spectral density.

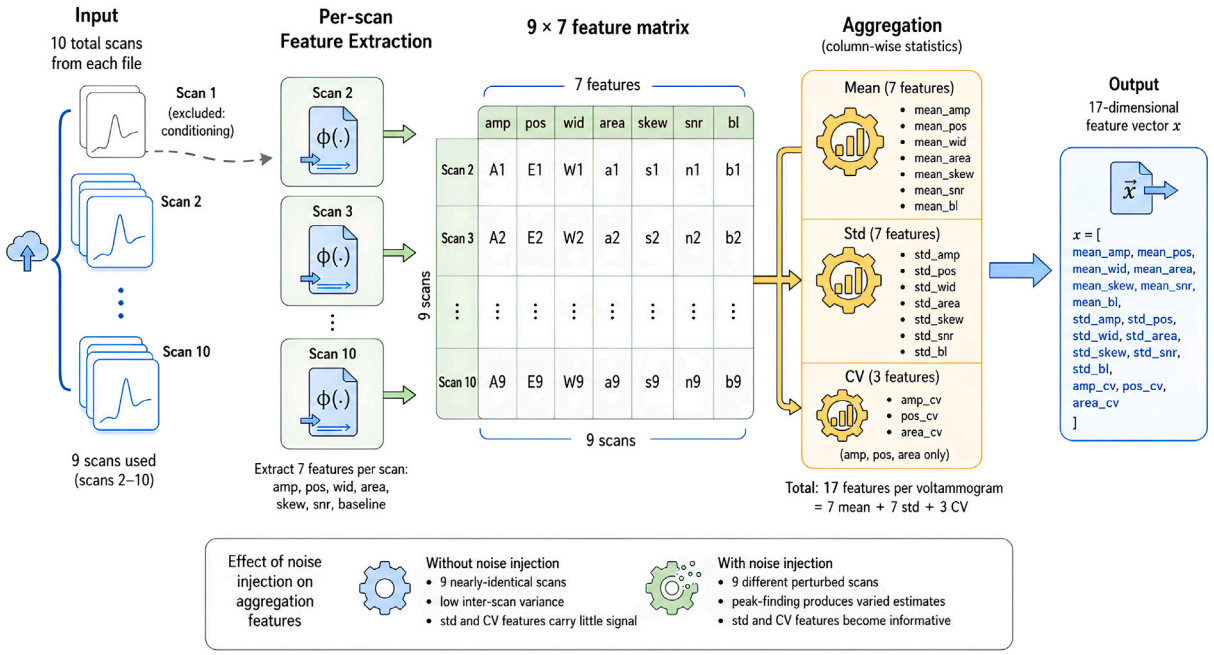

Fig. 2. Multi-scan feature extraction and aggregation. Each of the nine non-conditioning scans passes through the per-scan feature extractor $\phi(\cdot)$, which returns seven peak features, namely amplitude, position, width, area, skewness, signal-to-noise, and baseline at the peak. The resulting 9×7 matrix is aggregated column-wise into a seventeen-dimensional feature vector, namely the mean and standard deviation of all seven features together with the coefficient of variation of amplitude, position, and area. The lower legend contrasts the two regimes that drive the central finding. Without noise injection the nine scans are nearly identical, and thus the standard-deviation and coefficient-of-variation features carry little signal. With per-scan noise injection the peak finder produces varied estimates across scans, and these dispersion features become informative inputs to the regressor (Stage 2 in Fig. 1).

4. Heteroscedastic, Gaussian samples whose local standard deviation scales with the absolute deviation of the signal from its median, so that the variance concentrates in the high-current peak regions.

The target standard deviation follows from the peak-relative SNR as $\sigma_{\text{noise}} = (\text{peak height}) \cdot 10^{-\text{SNR}_{\text{dB}}/20}$, where the peak height is the maximum absolute amplitude of the quadratically detrended scan within the analytical window. The complete implementation of all four generators is available in the public code repository (Section 5).

Six SNR levels were evaluated: 30, 25, 20, 15, 10, and 5 dB. Each (noise type, SNR level) configuration was evaluated at three independent random seeds (42, 7, and 123). The complete sweep covered $4 \times 6 \times 3 = 72$ noisy configurations in addition to the clean baseline. Representative voltammograms before and after injection are shown in Figs. 3 and 4.

2.7. Random Forest prediction

Random Forest [57] was selected as the primary regressor because it suits a small tabular feature set with nonlinear and saturating concentration responses, requires little tuning, is robust to correlated and differently scaled features, and exposes feature importances on which the mechanistic analysis in Section 3.6 relies. The regressor used 100 trees, maximum depth 10, and minimum leaf size 2, taking the multi-scan feature vector (clean or noise-injected) as input to predict concentration in the original units (ppm). Predictions were clipped at 0.1 ppm. Hyperparameters were not tuned per condition, and the same configuration was used throughout. Their influence is examined directly in the sensitivity analysis of Section 3.7. The cross-model confirmation in Section 3.10 additionally evaluates three alternative regressors, namely XGBoost, SVR, and an MLP, under matched conditions, with hyperparameter settings detailed in that section.

2.8. Cross-validation evaluation protocol

Stratified five-fold cross-validation [58] was used, with concentration bins $[0, 10, 30, 100, 500, 1100]$ ppm to ensure proportional representation across folds. These six values are the boundaries of five concentration strata that group the fifteen calibration levels into analytically distinct ranges, namely a near-detection region below 10 ppm, a lower mid-range from 10 to 30 ppm, an upper mid-range from 30 to 100 ppm, and two higher ranges from 100 to 500 ppm and from 500 ppm to the maximum of 1000 ppm. The boundaries were chosen to follow the order-of-magnitude spacing of the calibration series and to align with the low, mid, and high reporting bands used for the error metrics. Stratifying the folds on these bins ensures that each fold contains samples from every range, which prevents any fold from being dominated by a single part of the calibration curve and yields per-band error estimates that can be interpreted against the same analytically meaningful ranges. Feature aggregation collapses each measurement file to a single 17-dimensional row, and thus the unit of cross-validation is the file and no scan from a file can appear in both the training and the test fold. Replicate files at the same concentration are distributed across folds by the stratified sampling, which does not constitute leakage because the regression target is the concentration and the replicates are independent measurements. The standardiser and the regressor were fit on the training fold only, and the test fold was transformed with the training-fold statistics. Noise injection acts on the raw per-scan signal before feature extraction and uses no concentration information, and thus it cannot leak the target. Performance metrics were computed on out-of-fold predictions, namely R^2, mean absolute error (MAE), MAPE, and per-band MAPE for the low band below 20 ppm, the mid band from 20 to 100 ppm, and the high band at or above 100 ppm. The mid-range MAPE (MAPE_{mid}) is reported as the primary metric, for two reasons. First, MAPE is scale-relative and weights trace and mid-range concentrations comparably rather than letting the high-concentration points dominate, as an absolute error would. Second, the clean baseline is already accurate above the mid band and is limited by detectability

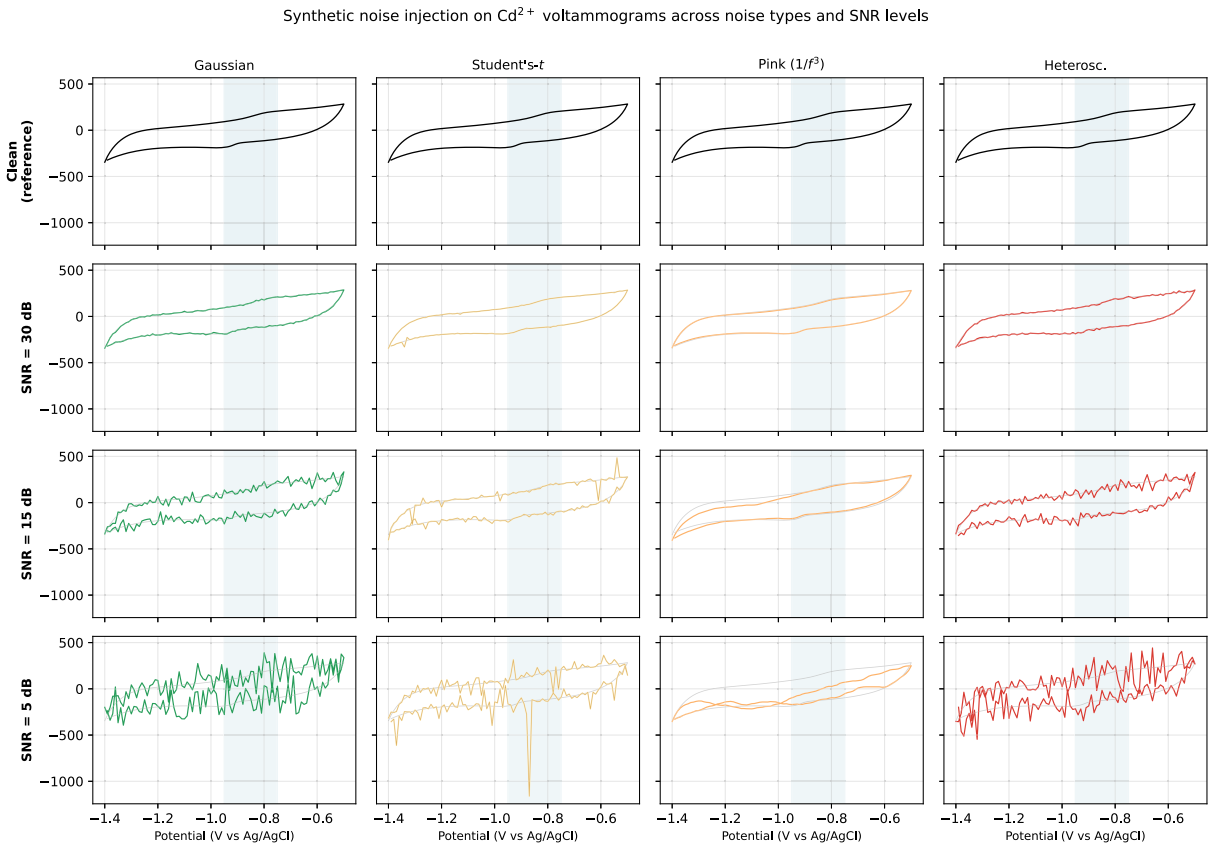

Fig. 3. Synthetic noise injection on Cd^{2+} voltammograms at 60 ppm across the four noise distributions (columns) and three representative SNR levels (rows). The clean reference is shown in the top row; rows 2–4 correspond to SNR = 30, 15, and 5 dB respectively, with the clean reference superimposed in light grey for direct comparison. Shaded vertical bands indicate the analytical peak window. The four distributions produce qualitatively different perturbation textures: Gaussian and Student's-t are pointwise (with the latter exhibiting heavy-tailed spikes visible at low SNR), pink noise produces low-frequency wandering that obscures the peak shape entirely at SNR = 5 dB, and the heteroscedastic perturbation is amplitude-dependent and concentrated in the high-current regions.

below it, so the mid band is where the method can change the analytical outcome. The coefficient of determination R^2 over the full range is retained as a secondary, aggregate-fit metric.

2.9. Statistical analysis

The SNR sensitivity analysis is based on point estimates, reported as the mean and one standard deviation across three seeds (42, 7, and 123). The noise-augmentation experiment was expanded to twenty random seeds per condition to support robust inferential claims. The mean $MAPE_{mid}$, an exact Wilcoxon signed-rank test on the per-seed paired differences against the clean baseline, the standardised effect size Cohen's d_z, and a bootstrap 95% confidence interval on the change relative to the baseline with 10 000 resamples are reported for each metal and noise distribution at 30 dB. The Wilcoxon test is used in place of a t-test because it makes no normality assumption on the per-seed differences. Exact p-values are reported throughout, and the word significant is reserved for results that passed the test at $\alpha = 0.05$. Where several distributions are compared for one metal, a Bonferroni correction over the eight single-distribution comparisons is applied ($\alpha = 0.05/8 = 0.006$).

3. Results

The central result of this paper is the asymmetric noise-augmentation effect and its statistical validation, reported in Section 3.5, together with the mechanism that explains it (Sections 3.6 and 4.1) and the diagnostic dissociation between a stable R^2 and a falling MAPE (Section 4.4). The present study first establishes the measurement-noise characterisation that motivated the augmentation design and then reports the clean baseline before turning to that core finding.

3.1. Empirical noise characterisation

Replicate residual analysis showed that the CV measurement noise departs from the ideal Gaussian assumption (Figs. 5 and 6). The power spectral density followed a $1/f^{\beta}$ law, with β near 3 at the concentrations where the peak signal dominates and an across-concentration mean of about 2.7, which is steeper than Brownian noise and motivates testing coloured noise. The within-replicate noise level rose with concentration, scaling roughly as $C^{0.4}$ to $C^{0.5}$ for both metals, although with substantial scatter between concentrations. The residual distribution was non-Gaussian and heavy-tailed, with a large excess kurtosis driven by occasional large excursions across scans rather than by a smooth heavy-tailed law, and thus the present study treats the tail behaviour qualitatively and does not fit a single parametric tail model.

These properties, namely a $1/f^{\beta}$ spectrum with β near 3, heteroscedastic concentration scaling, non-Gaussian heavy tails, and region dependence within voltammograms, motivated testing four distinct noise distributions in the SNR sensitivity analysis rather than Gaussian noise alone (Section 3.4). They also provide context for the noise-augmentation finding, since the effect reported below is not specific to Gaussian noise.

A further diagnostic examined how closely the injected noise reproduces the temporal structure of the real measurement noise. The autocorrelation of the injected Gaussian noise is near zero at all nonzero lags, consistent with the pointwise independence of the generator,

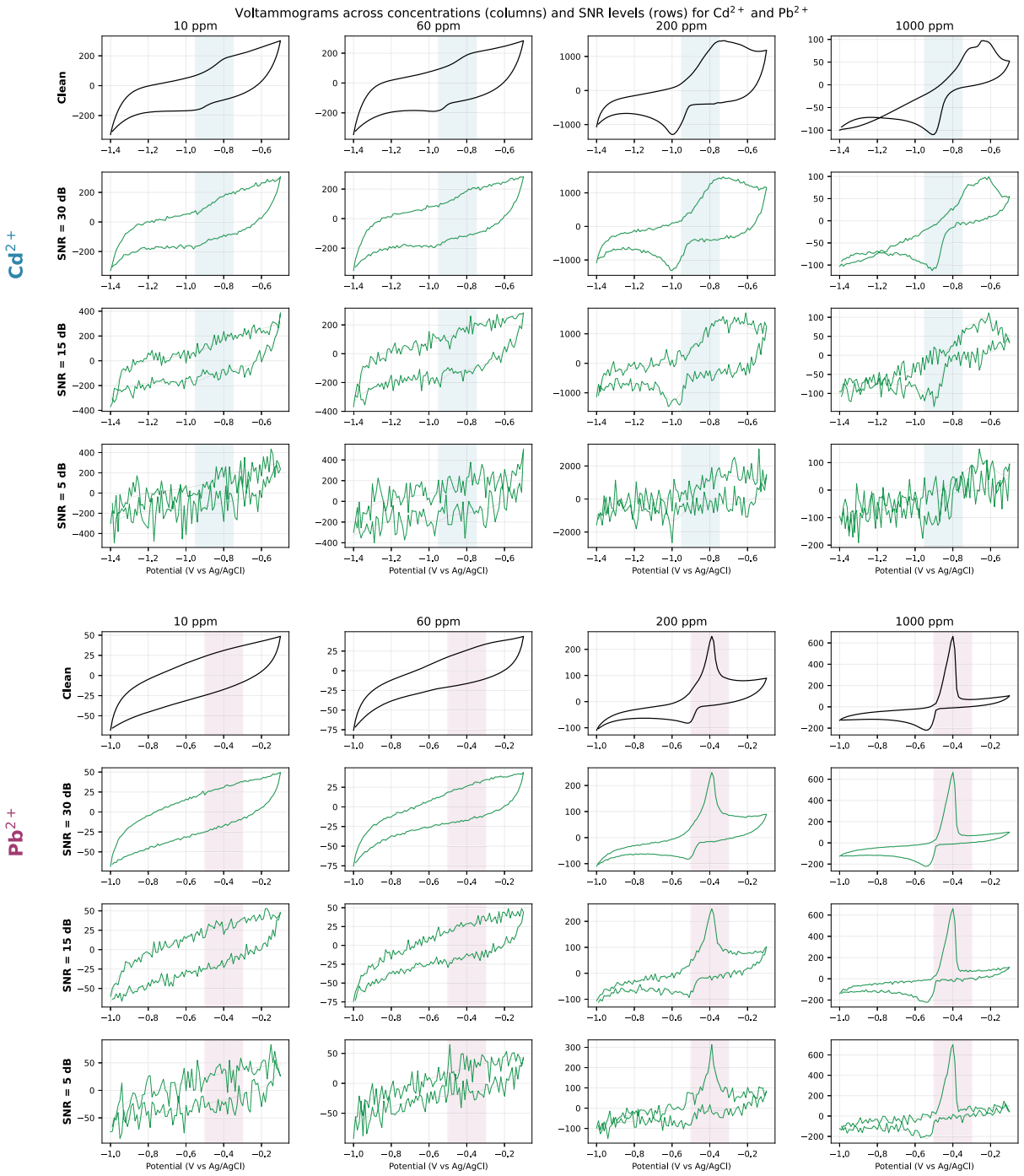

Fig. 4. Voltammograms across four representative concentrations (10, 60, 200, and 1000 ppm) and four signal conditions (clean, Gaussian noise injection at SNR = 30, 15, and 5 dB) for Cd^{2+} (top block, rows 1–4) and Pb^{2+} (bottom block, rows 5–8). Within each block, the top row shows the clean reference and the subsequent rows show progressively noisier injection. Shaded vertical bands mark the analytical peak window for each metal. The figure illustrates two regularities of the dataset: (i) characteristic peak features become clearly distinguishable from the background only above the low-concentration regime (approximately 200 ppm for Cd^{2+} and Pb^{2+} alike); and (ii) Pb^{2+} peaks at high concentration are noticeably sharper than Cd^{2+} peaks at the same concentration and remain identifiable even at SNR = 5 dB, consistent with the peak-shape interpretation discussed in Section 4.1.

whereas the injected pink noise carries the positive low-lag correlation expected of a $1/f^{\beta}$ process and therefore approximates the slow wandering seen in the replicate residuals. The mean in-window current drifted mildly and monotonically across the nine scans, by roughly 15% of the in-window mean for Cd^{2+} at 100 ppm, which indicates a small but nonzero within-file drift rather than a strictly flat baseline over the acquisition window. The injected models thus span both the uncorrelated and the slowly correlated regimes of the empirical noise. Strong long-timescale instrumental drift, which the present injection does not reproduce, is treated as a separate acquisition-side concern in Section 4.

3.2. Clean-baseline performance

The multi-scan pipeline with Random Forest regression achieved the following mid-range MAPE values under stratified five-fold cross-validation on clean signals (mean across twenty seeds, bins 20–100 ppm):

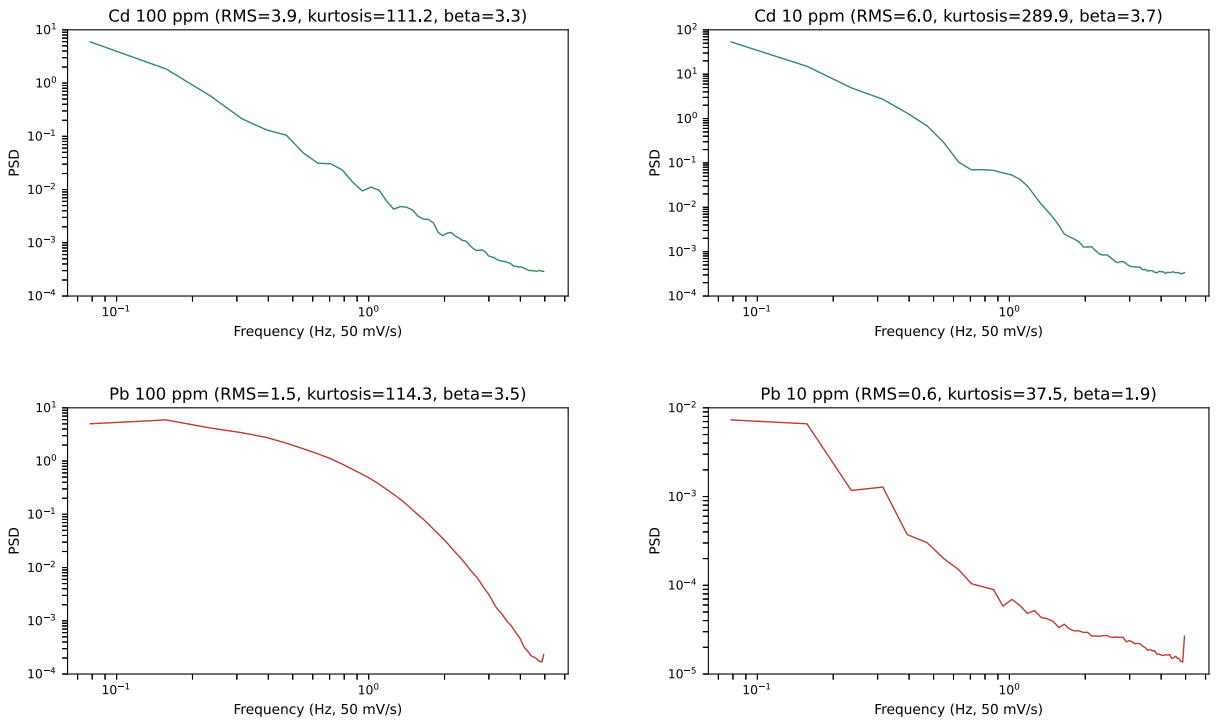

Fig. 5. Power spectral density of replicate residuals follows a $1/f^{\beta}$ law, with β near 3 at the concentrations where the peak signal dominates. Example concentrations are shown.

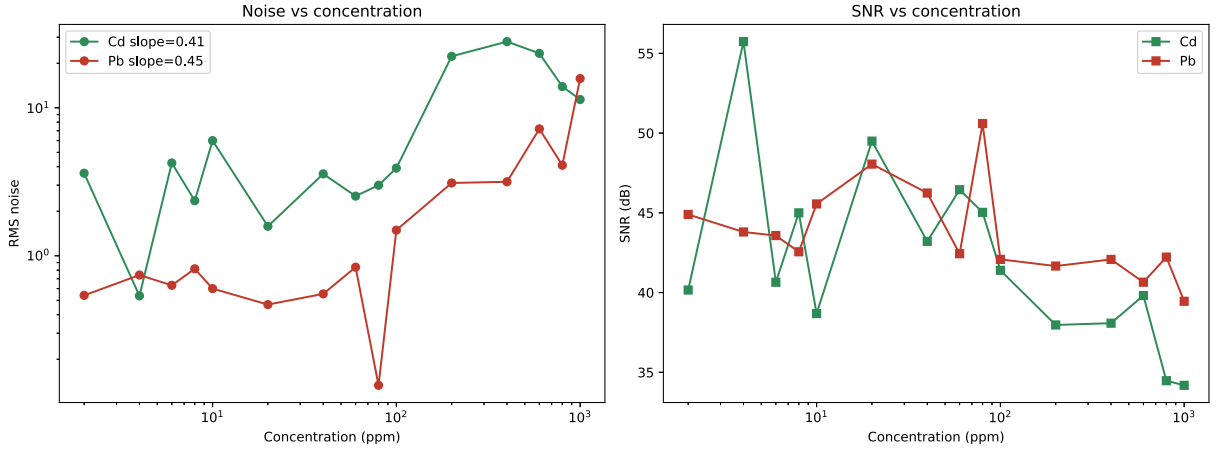

Fig. 6. Within-replicate noise standard deviation scales approximately as $C^{0.4}$ to $C^{0.5}$ for both Cd^{2+} and Pb^{2+}, with substantial scatter between concentrations. Example concentrations are shown.

- Cd^{2+}: $MAPE_{mid} = 40.4\%$, $R^2 \approx 0.96$
- Pb^{2+}: $MAPE_{mid} = 10.2\%$, $R^2 \approx 0.99$

These baseline values are not claimed as a methodological contribution. The more interesting finding lies in the way this baseline changes under noise injection, which is examined in Section 3.5.

3.3. Single-scan versus multi-scan benchmark

The 17-feature multi-scan vector was compared against a single-scan seven-feature vector under matched conditions to contextualise the value of the multi-scan representation (Table 1). Under clean acquisition the multi-scan representation is no better than single-scan for Cd^{2+}, at 40.1 against 40.3%, and is slightly worse for Pb^{2+}, at 10.1 against 7.1%, which confirms the statement that the clean multi-scan baseline is comparable to or slightly worse than a simpler single-scan extraction. Under Gaussian injection at 30 dB the multi-scan representation reduces the Cd^{2+} mid-range error from 42.0 to 13.0%, a 69% reduction relative to single-scan, while the Pb^{2+} values remain close, at 9.5 against 14.0%. The aggregation therefore adds value only when injection makes the across-scan dispersion features informative, which links this benchmark to the mechanism in Section 4.1.

3.4. SNR sensitivity analysis

Fig. 7 reports MAPE versus SNR for the four noise distributions, separately for the mid-range and high-range concentration bands. Three patterns are visible.

First, Pb^{2+} at concentrations ≥100 ppm remained at MAPE 2.7 to 4.0% across all 24 noise configurations (standard deviation < 0.7%). High-concentration Pb^{2+} detection therefore appears to be largely insensitive to noise across the SNR range tested.

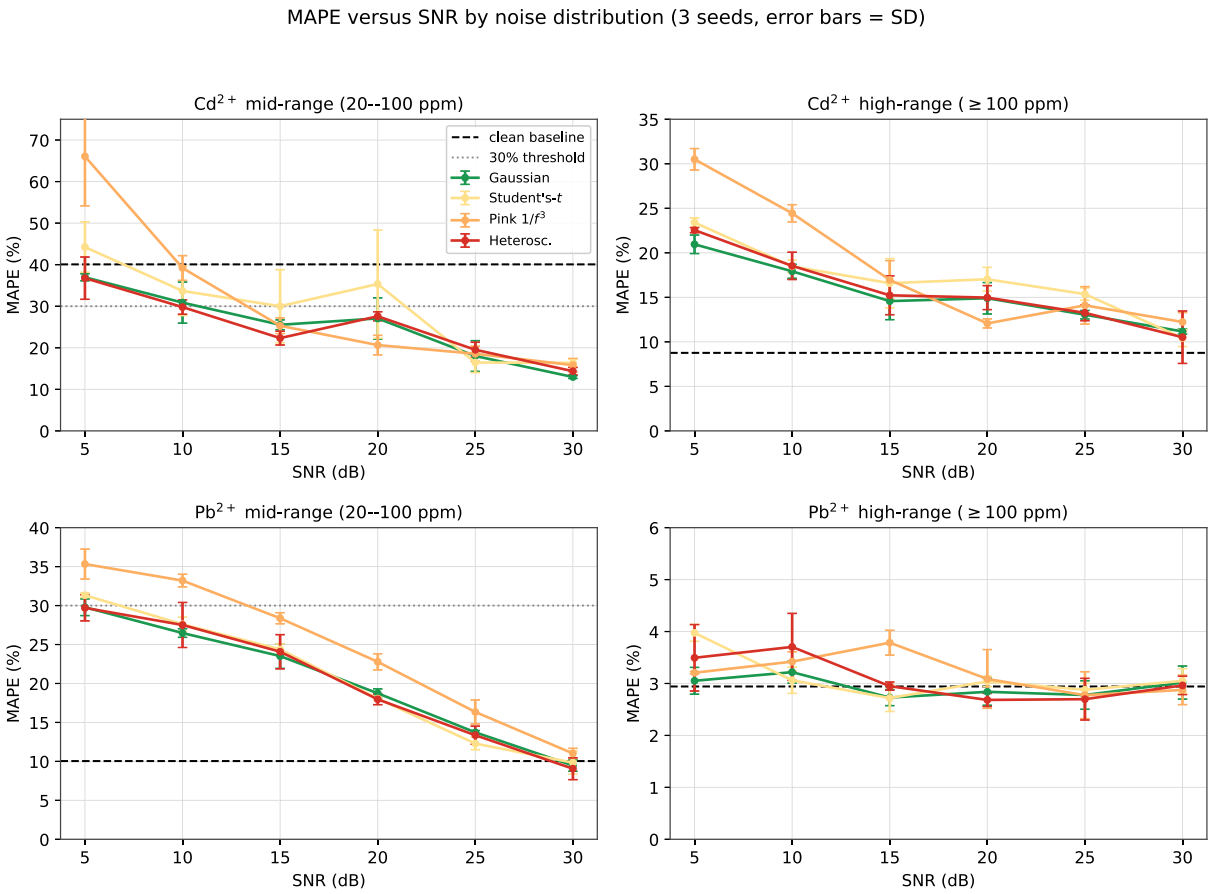

Fig. 7. MAPE versus SNR for the four noise distributions, in the mid-range (20–100 ppm) and high-range (≥100 ppm) concentration bands. Error bars show the standard deviation across three random seeds. The 30% MAPE threshold and the clean baseline are indicated for reference.

Table 1
Single-scan versus multi-scan mid-range MAPE (%) under clean and Gaussian-noise (30 dB) conditions, reported as a self-contained three-seed benchmark. The twenty-seed augmentation experiment in Table 2 uses a separate seed set, and thus the multi-scan noise value differs slightly. The aggregation benefit appears only under noise injection.

Condition	Metal	Single-scan	Multi-scan
Clean	Cd^{2+}	40.3	40.1
Clean	Pb^{2+}	7.1	10.1
Gaussian 30 dB	Cd^{2+}	42.0	13.0
Gaussian 30 dB	Pb^{2+}	14.0	9.5

Second, mid-range performance crossed the 30% MAPE threshold between approximately 10 and 15 dB for Pb^{2+} and between 15 and 25 dB for Cd^{2+}. Pink ($1/f^3$) noise produced the largest mid-range error at the lowest SNR, particularly for Cd^{2+}, where $MAPE_{mid}$ reached 66% at 5 dB compared with 37 to 44% under the other three distributions, which identifies pink noise as the dominant failure mode at low SNR.

Third, and most relevant to the present paper, the curves at SNR = 30 dB lie *well below* the clean baseline for Cd^{2+} in the mid-range, while remaining at or near the clean baseline for Pb^{2+}. The Pb^{2+} panels show the curves intersecting the clean baseline near 30 dB, whereas the Cd^{2+} mid-range panel shows all four curves substantially below the clean baseline at high SNR. This pattern motivated the focused statistical analysis presented next.

A complementary view of the same data is given as a heatmap in Fig. 8, which makes three features more immediately legible than the line plots in Fig. 7: the diagonal-stripe pattern by which pink noise dominates the failure mode for Cd^{2+} at low SNR (top-left quadrant); the near-uniform colour of the Pb^{2+} high-range panel, indicating insensitivity to noise; and the contrast between the bold (low-MAPE) cells at SNR = 30 dB for Cd^{2+} mid-range and the corresponding Pb^{2+} cells, which match the clean baseline.

3.5. The asymmetric noise-augmentation effect

The objective of the present study was to perform a focused confirmation experiment in order to ascertain whether the observed SNR of 30 dB reflected a genuine augmentation effect: twenty random seeds per noise distribution at SNR = 30 dB, with exact Wilcoxon signed-rank tests against the clean baseline (Table 2, Fig. 9).

All four noise distributions reduced the Cd^{2+} $MAPE_{mid}$ relative to the clean baseline, with reductions of 55.0 to 63.0%, exact Wilcoxon $p < 10^{-6}$, and large effect sizes (d_z between −3.1 and −3.8) in every case. The bootstrap 95% confidence intervals exclude zero by a wide margin, spanning roughly 49 to 67% reduction. The effect is consistent across distributions with substantially different shape and spectral characteristics, namely a heavy-tailed distribution (Student's-t with df = 1.5), a short-tailed approximately symmetric distribution (Gaussian), a coloured-noise distribution emphasising low frequencies (pink $1/f^3$), and an amplitude-modulated distribution (heteroscedastic).

The Pb^{2+} response is an order of magnitude smaller and is not reliable. Only heteroscedastic injection produced a reduction that was significant at the uncorrected level (−5.5%, p = 0.024, d_z = −0.58), and it does not survive the Bonferroni correction over the eight single-distribution comparisons ($\alpha = 0.05/8 = 0.006$). Gaussian injection was not significant (−3.7%, p = 0.058), and Student's-t and pink injection produced no reduction at all (+0.3% and +2.9%). The hyperparameter analysis below sharpens this picture, since tuning reverses the sign of

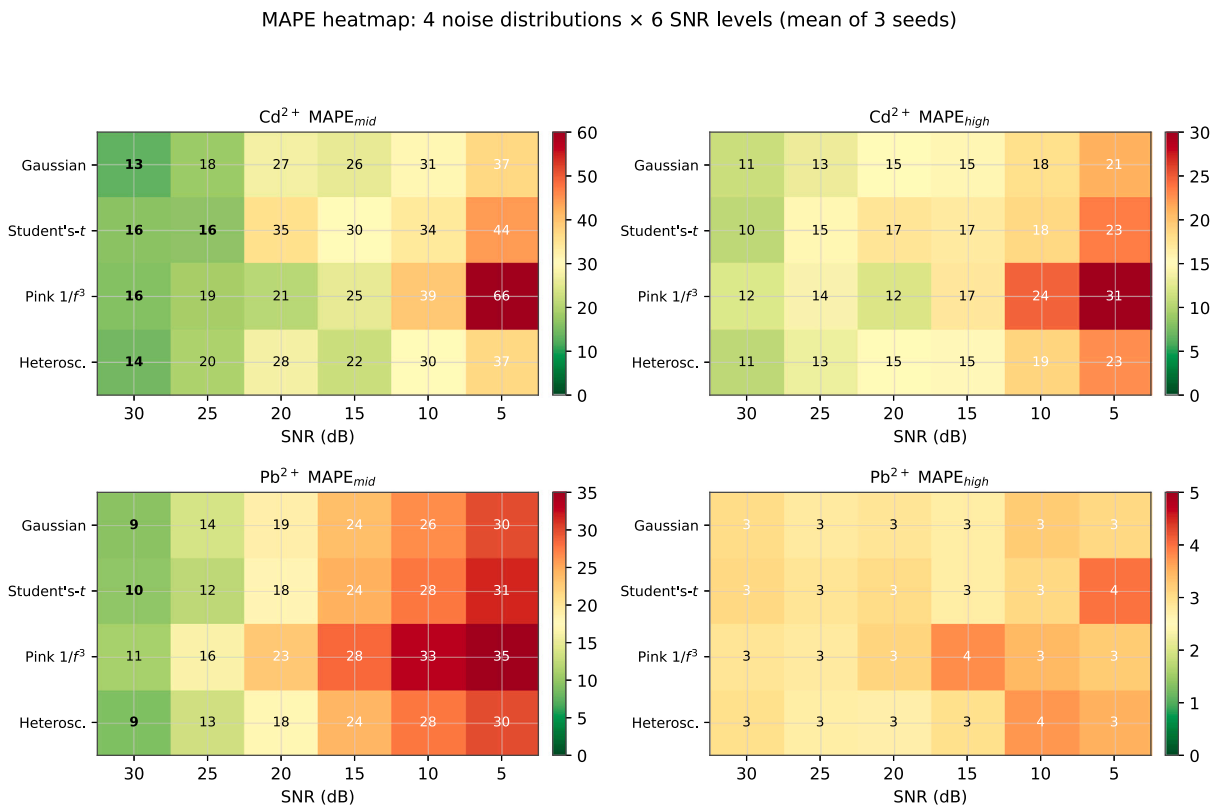

Fig. 8. MAPE heatmap across the full SNR sweep (4 noise distributions × 6 SNR levels). Each cell reports the mean MAPE across three random seeds. Colour scales are panel-specific (capped at 60%, 30%, 35%, and 5% for Cd^{2+} mid, Cd^{2+} high, Pb^{2+} mid, and Pb^{2+} high, respectively); cells with values below approximately 30% of the panel maximum are rendered in bold to highlight the augmentation regime.

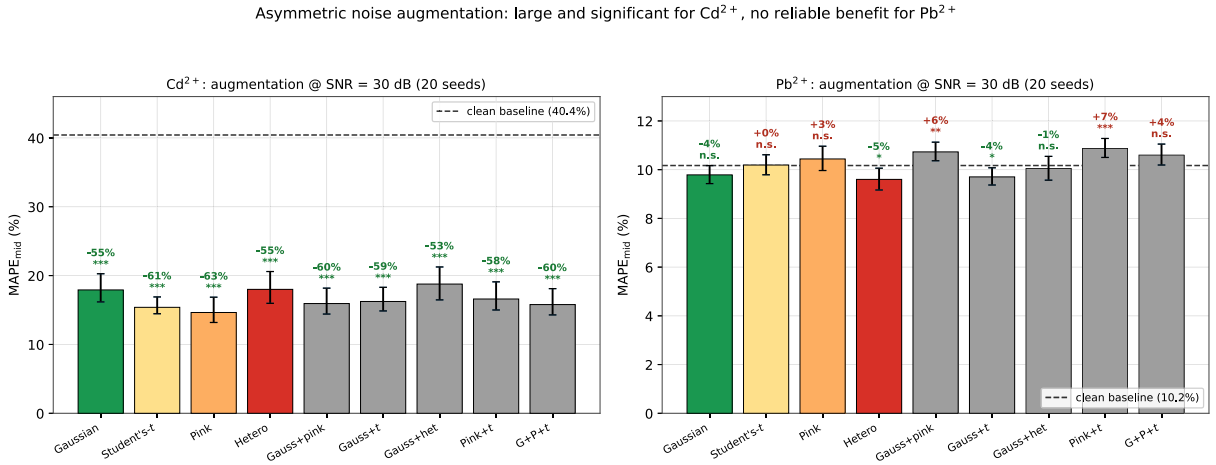

Fig. 9. Change in mid-range MAPE relative to the clean baseline at SNR = 30 dB, over twenty seeds with bootstrap 95% confidence intervals (error bars), with bars coloured by noise distribution. In the left panel, all four distributions reduce the Cd^{2+} MAPE by 55.0 to 63.0%. In the right panel, the Pb^{2+} change is an order of magnitude smaller, and only heteroscedastic injection is significant at the uncorrected level. Markers denote $p < 0.001$ (***), $p < 0.05$ (*), or a non-significant change (n.s.).

the Pb^{2+} change. The asymmetry between the two metals is therefore one of presence rather than degree, namely a large and robust Cd^{2+} benefit against no reliable Pb^{2+} benefit.

The persistence of a large Cd^{2+} effect across distributions with such different characteristics rules out distribution-specific explanations. The single property shared by all four distributions is the independent stochasticity of the perturbation across scans within each measurement file. The mechanistic implications are discussed in Section 4.1.

3.6. Feature importance before and after augmentation

Permutation feature importances were computed under clean and Gaussian-injected (30 dB) conditions to test the proposed mechanism directly, grouping the six across-scan dispersion features (the standard-deviation and coefficient-of-variation columns) and expressing their summed importance as a share of the total (Table 3). The Cd^{2+} dispersion features carry only 0.5% of the total importance under clean

Table 2
Statistical analysis of the noise-augmentation effect at SNR = 30 dB over twenty random seeds. Exact Wilcoxon signed-rank tests compare $MAPE_{mid}$ against the clean-baseline mean (Cd^{2+}: 40.4%; Pb^{2+}: 10.2%). d_z is Cohen's standardised effect size, and the 95% confidence interval is a bootstrap interval on the change Δ relative to the baseline (10 000 resamples). The coefficient of determination remained near 0.96 for Cd^{2+} and 0.99 for Pb^{2+} across all conditions. Only the heteroscedastic Pb^{2+} effect is significant at $\alpha = 0.05$, and it does not survive the Bonferroni correction over the eight single-distribution comparisons.

Metal	Noise	$MAPE_{mid}$ (%)	Δ (%)	95% CI (%)	d_z	p
Cd^{2+}	Clean baseline	40.4	—	—	—	—
Cd^{2+}	Gaussian	17.9	−54.9	[−60.0, −49.9]	−3.16	9.5×10^{-7}
Cd^{2+}	Student's-t	15.4	−61.2	[−64.2, −58.2]	−3.82	9.5×10^{-7}
Cd^{2+}	Pink $1/f^3$	14.6	−62.8	[−67.4, −58.3]	−3.47	9.5×10^{-7}
Cd^{2+}	Heteroscedastic.	18.0	−54.8	[−60.5, −49.1]	−3.09	9.5×10^{-7}
Pb^{2+}	Clean baseline	10.2	—	—	—	—
Pb^{2+}	Gaussian	9.8	−3.7	[−7.3, −0.0]	−0.47	0.058
Pb^{2+}	Student's-t	10.2	+0.3	[−3.7, +4.3]	+0.03	1.0
Pb^{2+}	Pink $1/f^3$	10.4	+2.9	[−2.0, +7.8]	+0.27	0.37
Pb^{2+}	Heteroscedastic.	9.6	−5.5	[−9.9, −1.1]	−0.58	0.024

Table 3
Summed importance of the six across-scan dispersion features (standard deviation and coefficient of variation) as a percentage of total permutation feature importance, under clean and Gaussian-noise (30 dB) conditions, averaged over seeds.

Metal	Clean (%)	Noise (%)
Cd^{2+}	0.5	2.8
Pb^{2+}	2.8	3.3

Table 4
Augmentation effect under simultaneously combined noise distributions at SNR = 30 dB. Values are the change in $MAPE_{mid}$ relative to the clean baseline.

Combination	Cd^{2+} Δ (%)	Pb^{2+} Δ (%)
Gaussian + pink	−59.7	+5.7
Gaussian + Student's-t	−59.0	−4.4
Gaussian + heteroscedastic	−53.3	−1.1
Pink + Student's-t	−57.8	+7.1
Gaussian + pink + Student's-t	−59.9	+4.4

acquisition but 2.8% under injection, a fivefold increase. The same Pb^{2+} share is essentially unchanged, at 2.8 against 3.3%. This is the pattern predicted by the mechanism in Section 4.1, namely that injection makes the across-scan dispersion of the Cd^{2+} peak estimates informative, whereas for the already-stable Pb^{2+} peak the dispersion features carry only a small share that injection does not enlarge.

3.7. Hyperparameter sensitivity

The hyperparameters were tuned by nested cross-validation to verify that the conclusions do not depend on the fixed Random Forest configuration, with the number of trees, the maximum depth, and the minimum leaf size selected on inner folds and the augmentation effect evaluated on the outer folds. Tuning lowered the clean baselines, to 37.0% for Cd^{2+} and 7.9% for Pb^{2+}, yet the asymmetry persisted and sharpened. The tuned Cd^{2+} Gaussian-injection effect was −63.5%, from 37.0% to 13.5%, even larger than the untuned effect. The tuned Pb^{2+} change was +21.2%, from 7.9% to 9.5%, and thus once the clean Pb^{2+} model is properly tuned, injection makes it worse rather than better. The Cd^{2+} benefit is therefore robust to tuning, whereas the apparent Pb^{2+} effect is not.

3.8. Combined noise distributions

The augmentation experiment was repeated with simultaneously injected combinations at 30 dB because real acquisition noise mixes several components (Table 4). The Cd^{2+} reduction persisted under all combinations, between 53 and 60%, indistinguishable from the single-distribution results. The Pb^{2+} change remained small, and several combinations significantly increased the error rather than reducing it. Combining distributions therefore neither amplifies nor cancels the Cd^{2+} effect, which is consistent with the mechanism resting on the independent across-scan stochasticity shared by every component rather than on any single spectral signature, while confirming that Pb^{2+} derives no benefit.

3.9. Train-only versus acquisition-time injection

A natural question is whether the benefit is an ordinary training-time data augmentation, in which case it should appear when noise is added only to the training folds. Three protocols were compared under the same leakage-safe cross-validation, namely a clean baseline (A), noise injected at acquisition so that both the training and the test folds are noisy (B), and noise injected only on the training folds while the test folds stay clean (C). In Cd^{2+}, the protocols gave 40.4% (A), 17.9% (B), and 637.1% (C). In Pb^{2+}, they gave 10.2% (A), 9.8% (B), and 75.3% (C). Train-only injection is therefore catastrophic, because a model trained on noisy multi-scan features and applied to clean ones faces a severe distribution shift. The benefit appears only when the noise is present at acquisition, which means the result is not a training-time regulariser but a property of how the multi-scan features are measured. This reframes the practical recommendation as acquiring deliberately noisier scans rather than augmenting a fixed clean dataset.

3.10. Generalisation across regressor families

Consequently, in order to assess whether the augmentation effect is specific to Random Forest or generalises across model classes, the multi-scan pipeline was evaluated with three additional regressors under matched conditions (clean baseline and Gaussian noise injection at SNR = 30 dB, three random seeds): gradient-boosted trees [XGBoost, 59], ε-insensitive support vector regression with an RBF kernel (SVR), and a multilayer perceptron (MLP). Results are summarised in Fig. 10 and Table 5.

The Cd^{2+} augmentation effect appears in all four regressor families. The effect is largest for the two tree-based models, at −67.6% for Random Forest and −41.4% for XGBoost, and smaller but in the same direction for the continuous-decision-function models, at −20.4% for SVR and −25.0% for the MLP. The effect is therefore not specific to Random Forest. The Pb^{2+} picture is mixed and small. Random Forest and SVR show reductions of a few percent, whereas XGBoost and the MLP show increases of 41.5 and 17.3%, so no consistent benefit emerges for Pb^{2+} in any model.

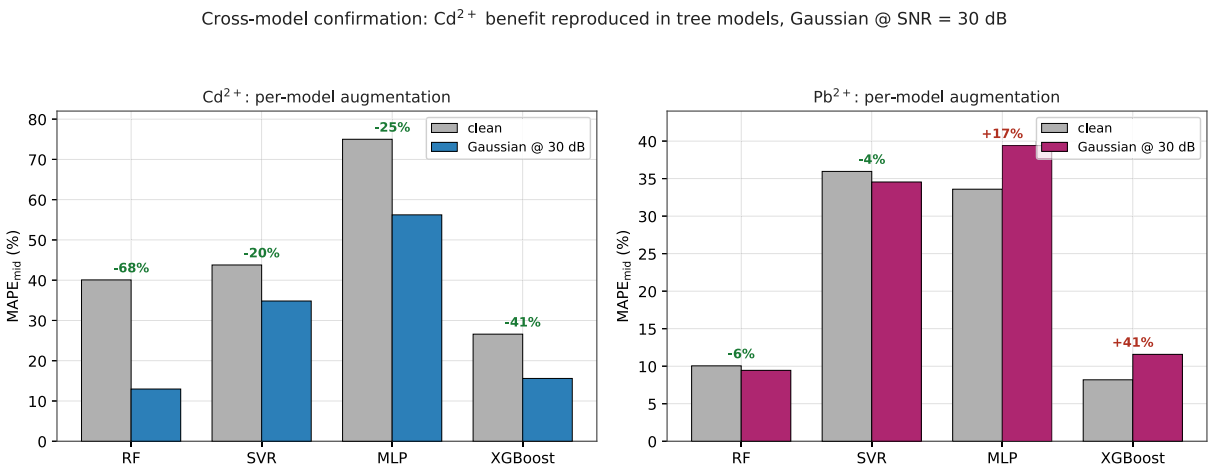

Fig. 10. Generalisation across regressor families under Gaussian injection at SNR = 30 dB, averaged over three seeds, showing the change in mid-range MAPE relative to the clean baseline with bars coloured by model. In the left panel, the Cd^{2+} reduction appears in all four families and is largest for the tree-based models, namely RF and XGBoost. In the right panel, the Pb^{2+} change is small and sign-variable, with XGBoost and the MLP showing an increase.

Table 5
Augmentation effect across regressor families at SNR = 30 dB Gaussian noise. $MAPE_{mid}$ values are means across three seeds. The Cd^{2+} effect appears in all four families and is largest for the tree-based models. SVR and the MLP show much higher baseline MAPE, which indicates that they underfit the multi-scan representation.

Model	Metal	Clean $MAPE_{mid}$ (%)	Noise $MAPE_{mid}$ (%)	Δ (%)
RF	Cd^{2+}	40.1	13.0	−67.6
XGBoost	Cd^{2+}	26.6	15.6	−41.4
SVR	Cd^{2+}	43.8	34.8	−20.4
MLP	Cd^{2+}	75.0	56.2	−25.0
RF	Pb^{2+}	10.1	9.5	−6.0
XGBoost	Pb^{2+}	8.2	11.6	+41.5
SVR	Pb^{2+}	36.0	34.6	−3.9
MLP	Pb^{2+}	33.6	39.4	+17.3

The pattern is consistent with the proposed mechanism (Section 4.1). Tree-based regressors can directly partition the feature space along the standard-deviation and coefficient-of-variation dimensions of the multi-scan feature vector, which become informative under noise injection, whereas SVR and the MLP construct continuous decision functions that interact with these features less directly and therefore extract less of the augmentation benefit. No model class extracts a benefit for Pb^{2+}, in line with the peak-shape interpretation that when deterministic peak finding is already stable, the additional stochasticity yields no advantage.

Notably, SVR and MLP achieve substantially higher baseline MAPE values than the tree-based regressors on this dataset (Table 5), indicating that they underfit the multi-scan representation. A more thorough hyperparameter search and architecture exploration could narrow this gap, but the focus of the present analysis is the relative change between clean and noise-injected conditions rather than absolute performance.

3.11. Controlled synthetic-peak study

Synthetic voltammograms were generated to isolate peak shape as the causal factor, with a broad and slightly asymmetric peak resembling Cd^{2+} and a narrow and symmetric peak resembling Pb^{2+}, then passed through the same feature-extraction and noise-injection pipeline. Gaussian injection reduced the mid-range error of the Cd^{2+}-like peak from 11.4% to 10.8%, a 4.3% reduction, while it left the Pb^{2+}-like peak essentially unchanged at a small increase of 3.4%. The synthetic study therefore reproduces the direction of the asymmetry from peak shape alone, namely a benefit for the broader and less symmetric peak and none for the sharp peak. The synthetic model does not reproduce the full magnitude of the effect seen on the real data, which indicates that the simplified peak model captures the qualitative mechanism rather than the complete peak-finding behaviour of the real electrode. This result supports the proposed mechanism while making clear what the controlled model can and cannot establish.

4. Discussion

4.1. A tentative mechanistic interpretation

The asymmetry between Cd^{2+}, where noise injection lowers MAPE by 55.0 to 63.0%, and Pb^{2+}, where no distribution produces a reliable reduction, calls for explanation. The augmentation effect is observed across all four noise distributions tested and cannot rest on properties unique to any single distribution. The shared property is that each scan receives an independent random perturbation, which propagates through the deterministic feature-extraction pipeline and produces per-scan estimates that differ slightly from the clean-signal estimate.

The present study proposes a tentative interpretation linked to peak shape. In the present dataset, Cd^{2+} peaks at the GCE working electrode are broader and somewhat asymmetric, with subtle local extrema near the global maximum. Deterministic peak finding on a single clean scan can settle on one of these local extrema, biasing the peak-position and amplitude estimates that subsequently propagate through feature aggregation. With independent noise injection across nine scans, each scan's peak-finder produces a slightly different position estimate, and the multi-scan summary statistics (mean, standard deviation, coefficient of variation) become more reliable estimators than they would be for nearly identical scan-level estimates. The bottom panels of Fig. 2 sketch this contrast at the level of the feature matrix. Under clean

acquisition the standard-deviation and coefficient-of-variation columns of the aggregated vector carry little signal, whereas under independent noise injection these same columns become informative inputs to the regressor. The permutation feature-importance analysis in Section 3.6 confirms this empirically, since the summed importance of the dispersion features for Cd^{2+} rises from 0.5 to 2.8% of the total under injection. Pb^{2+} peaks are narrower and closer to Gaussian, so deterministic peak finding is already stable and the dispersion features carry only a small share under clean acquisition (2.8%) that injection does not enlarge (3.3%), which accounts for the absence of a reliable Pb^{2+} benefit. The controlled synthetic-peak study in Section 3.11 reproduces the direction of this contrast from peak shape alone.

This interpretation is conceptually similar to dropout regularisation in neural networks [46], where stochasticity prevents overfitting to deterministic artefacts, and is consistent with the broader theoretical result that training with noise is equivalent to a form of Tikhonov regularisation [45,52]. Related effects are documented in the time-series augmentation literature [47]. No prior reports of an analogous effect in cyclic voltammetric signal processing are known to the authors, although the analogy is general enough that the connection may have been noticed in adjacent literature beyond the present survey.

This interpretation was tested directly. The synthetic-peak study in Section 3.11 isolates peak shape as the experimental variable, and it reproduces the direction of the asymmetry, namely a benefit for the broad and asymmetric peak and none for the sharp peak. The synthetic model does not reproduce the full magnitude of the real-data effect, so it supports the mechanism without establishing it completely. A signal-level architecture that bypasses deterministic peak finding would provide a complementary test.

4.2. Practical implications

Should the augmentation effect generalise to similar electrochemical settings, it would suggest a counterintuitive design principle: for some analytes, deliberately injecting controlled noise during multi-scan acquisition — or, equivalently, accepting modestly noisier acquisition settings without aggressive post-acquisition filtering — may improve mid-range quantification accuracy relative to an idealised clean acquisition. This caveat is emphasised: the SNR analysis used synthetic noise injection on laboratory-acquired signals rather than direct field measurements, and the augmentation effect was demonstrated specifically with the multi-scan feature pipeline and Random Forest regression. Generalisation to other analytes, model classes, and acquisition setups remains to be verified.

A separate practical observation arises from the SNR sensitivity analysis (Section 3.4). The dominance of pink noise as the failure mode at low SNR is, if generalisable, a counterweight to the field's prevailing focus on high-frequency denoising. Slow drift and $1/f$-type interference [39] are largely orthogonal to the bandwidths targeted by Savitzky-Golay smoothing [40,41] and wavelet shrinkage [42,43], and may be better addressed by acquisition-side stabilisation — temperature control, vibration isolation, and electrode reconditioning — than by post-acquisition signal processing [36,37,44].

The mechanism is governed by peak morphology rather than by the chemical identity of the analyte, which indicates where the approach should and should not transfer. Any electroactive species whose stripping peak is broad, asymmetric, or partially convoluted, so that deterministic peak finding on a single scan can settle on a spurious local extremum, is a candidate for the same mid-range benefit, whereas species that produce sharp, symmetric, well-separated peaks, as Pb^{2+} does here, are not expected to benefit because their peak localisation is already stable. Heavy metals commonly co-determined by anodic stripping voltammetry, together with organic electroactive compounds that give broad oxidation or reduction waves, therefore fall within the plausible scope, while the criterion for any specific system remains empirical, namely whether its peaks are broad enough for deterministic peak finding to be unstable. This morphological criterion, rather than the Cd^{2+} and Pb^{2+} pair alone, defines the intended reach of the method, and it is directly testable on a new analyte through the permutation-importance and synthetic-peak diagnostics used here.

4.3. Computational cost and scalability

The workflow is computationally light and scales gracefully to larger datasets. Feature extraction is linear in the number of voltammograms and compresses each measurement file, namely nine scans of 180 points, into a single seventeen-dimensional row, so the regressor operates on a compact tabular matrix rather than on the raw signal and its cost grows with the number of files rather than with the length of each scan. A single Random Forest fit on this representation trains in seconds on a standard workstation without a GPU, and the memory footprint is modest because only the aggregated feature matrix is retained after extraction. The heavier cost in the present study comes from the validation suite rather than from the method itself, namely the twenty-seed augmentation experiment, the nested cross-validation, the permutation-importance analysis, and the ten-thousand-resample bootstrap, each of which repeats the same inexpensive fit and parallelises cleanly across seeds and folds. Scaling to substantially larger datasets therefore increases mainly the one-off feature-extraction pass, which is trivially parallel across files, while the model-fitting and inference cost remains governed by the low feature dimensionality. For datasets large enough to strain a single Random Forest, the pipeline accepts gradient-boosted trees, which Section 3.10 already evaluates and which offer favourable scaling on larger tabular inputs.

4.4. Relation to existing CV-ML pipelines

Several recent CV-ML pipelines have reported strong R^2 for heavy-metal quantification while showing elevated MAPE values [24,27]. One possible reading of the present findings is that the gap between R^2 and MAPE in such systems may, in part, reflect the deterministic peak-finding artefacts described in Section 4.1: a model can capture the overall concentration trend (yielding a high R^2) while accumulating per-sample relative errors (yielding a high MAPE) when its features encode small biases that depend on local noise realisations.

The present data support this reading directly. As shown in Tables 2 and 5, the noise-augmentation effect reduces $MAPE_{mid}$ by approximately 55.0 to 63.0% for Cd^{2+} while leaving R^2 essentially unchanged, near 0.96 across all conditions. The fact that R^2 is invariant under the same intervention that dramatically lowers MAPE indicates that the augmentation does not improve how well the model captures the concentration trend. Instead, the augmentation improves the per-sample accuracy by removing a per-sample bias. This dissociation between an aggregate-fit metric (R^2) and a per-sample error metric (MAPE) is consistent with the proposed mechanism and offers a quantitative diagnostic that future studies could use to detect similar peak-finding artefacts in their own pipelines.

4.5. Limitations

Several limitations should be acknowledged. The mechanistic interpretation in Section 4.1 is supported by a permutation feature-importance analysis and a controlled synthetic-peak study, but it is not established beyond doubt, and the noise-augmentation effect could plausibly involve mechanisms not considered here. The multi-regressor analysis in Section 3.10 addresses generalisation across feature-vector-based regressor families and shows that the effect is reproduced in tree-based models and weakened or absent in continuous decision-function models. The multi-regressor analysis does not, however, address whether the effect persists under deep learning architectures that operate directly on the raw signal and bypass deterministic peak finding

altogether. A 1D convolutional or recurrent architecture trained on stacked multi-scan signals would constitute a direct test of the peak-finding mechanism: an absence of the augmentation effect under such an architecture would strengthen the proposed mechanism, while a persistence would indicate that the mechanism hypothesis is incomplete. This comparison was not pursued in the present work owing to computational constraints, but it is a natural next step.

The dataset, while substantial at 1500 voltammograms, originates from a single laboratory setup, electrode material, and electrolyte. The internal consistency is therefore strong, but the transferability of the augmentation effect to other electrode systems, scan rates, and real environmental matrices is not established by the present data. Two considerations bound the expected generalisation. The mechanism depends on deterministic peak finding being biased for broad and asymmetric peaks, which is an electrode- and analyte-specific property rather than a feature of the present setup alone, and thus the effect is expected wherever such peak shapes occur. At the same time, real matrices introduce interferents and matrix effects that the laboratory standards here do not contain, and thus external validation on spiked real-water samples and on at least one alternative electrode remains necessary before the design principle can be recommended in practice.

The augmentation experiment used twenty seeds per condition, which yields the narrow confidence intervals reported in Table 2, whereas the SNR sweep used three seeds and therefore has wider intervals at low SNR, where seed-level variability is highest. The controlled synthetic-peak study reproduces the direction of the asymmetry but not its full magnitude, which indicates that the simplified peak model captures the qualitative mechanism rather than the complete peak-finding behaviour of the real electrode. Concentrations below approximately 15 to 20 ppm are not reliably quantified by the present pipeline. This matters for environmental monitoring, since regulatory limits for Cd^{2+} and Pb^{2+} in drinking water lie well below this range, so the method in its present form suits screening and mid-range quantification rather than trace-level compliance testing. Closing this gap would require preconcentration techniques such as adsorptive stripping voltammetry [22,26,28], which lower the detectable concentration and are a natural extension of this work.

5. Conclusion

The present study reports an asymmetric noise-augmentation effect in machine-learning-based cyclic voltammetric quantification of Cd^{2+} and Pb^{2+}. Stochastic noise injection during multi-scan acquisition reduces Cd^{2+} mid-range MAPE by 55.0 to 63.0% across four distinct noise distributions (Gaussian, Student's-t, pink $1/f^3$, heteroscedastic), with exact Wilcoxon $p < 10^{-6}$, large effect sizes, and bootstrap 95% confidence intervals that exclude zero over twenty random seeds. Pb^{2+} shows no reliable benefit, since only heteroscedastic injection is significant at the uncorrected level and the change reverses sign under hyperparameter tuning, and thus the asymmetry is one of presence rather than degree. The persistence of the large Cd^{2+} effect across distributions with substantially different shape and spectral characteristics rules out distribution-specific mechanisms, and a permutation feature-importance analysis confirms that the across-scan dispersion features gain importance under injection for Cd^{2+} but not for Pb^{2+}. The effect is robust to nested-cross-validation tuning and is reproduced across four regressor families, a train-only control shows that the benefit requires noise at acquisition rather than only during training, and a controlled synthetic-peak study reproduces the direction of the asymmetry from peak shape alone. A further step is to reproduce the effect with experimentally acquired noise rather than synthetic injection, by acquiring replicate voltammograms at a controlled and moderately reduced signal-to-noise ratio near the 30 dB value at which the injected benefit is largest, through acquisition-side means such as a higher scan rate, a shortened equilibration time, or reduced instrumental filtering. The train-only control in Section 3.9 already localises the benefit to acquisition rather than training, so genuine hardware noise is expected to reproduce the Cd^{2+} mid-range improvement, and a matched clean-versus-noisy acquisition campaign on the same electrode would confirm this directly. Testing generalisation to signal-level deep learning architectures that bypass deterministic peak finding, extending the synthetic study to overlapping multi-analyte peaks, and external validation on real-water samples are natural directions for future work.

CRediT authorship contribution statement

Alya Kamilah: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Riyanarto Sarno:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Kelly Rossa Sungkono:** Writing – review & editing, Validation, Methodology. **Rizqy Ahsana Putri:** Writing – review & editing, Resources, Methodology, Data curation. **Taufiq Choirul Amri:** Writing – review & editing, Validation, Methodology. **Fadlilatul Taufany:** Writing – review & editing, Validation, Resources. **Arif Abdullah Sagran:** Writing – review & editing, Validation. **Sang-Seok Lee:** Writing – review & editing, Validation.

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.

Acknowledgements

This research is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No. 4299/B3/DT.03.08/2025 & No 3029/PKS/ITS/2025).

Data availability

The analysis code is publicly available at GitHub repository.

References

[1] W. Xu, Y. Jin, G. Zeng, Introduction of heavy metals contamination in the water and soil: a review on source, toxicity and remediation methods, Green Chem. Lett. Rev. 17 (1) (2024) 2404235, http://dx.doi.org/10.1080/17518253.2024.2404235.

[2] Q. Zhou, N. Yang, Y. Li, B. Ren, X. Ding, H. Bian, X. Yao, Total concentrations and sources of heavy metal pollution in global river and lake water bodies from 1972 to 2017, Glob. Ecol. Conserv. 22 (2020) e00925, http://dx.doi.org/10.1016/j.gecco.2020.e00925.

[3] J. Briffa, E. Sinagra, R. Blundell, Heavy metal pollution in the environment and their toxicological effects on humans, Heliyon 6 (9) (2020) e04691, http://dx.doi.org/10.1016/j.heliyon.2020.e04691.

[4] H. Ali, E. Khan, I. Ilahi, Environmental chemistry and ecotoxicology of hazardous heavy metals: environmental persistence, toxicity, and bioaccumulation, J. Chem. 2019 (2019) 6730305, http://dx.doi.org/10.1155/2019/6730305.

[5] M.S. Collin, S.K. Venkatraman, N. Vijayakumar, V. Kanimozhi, S.M. Arbaaz, R.G.S. Stacey, J. Anusha, R. Choudhary, V. Lvov, G.I. Tovar, F. Senatov, S. Koppala, S. Swamiappan, Bioaccumulation of lead (Pb) and its effects on human: a review, J. Hazard. Mater. Adv. 7 (2022) 100094, http://dx.doi.org/10.1016/j.hazadv.2022.100094.

[6] M. Balali-Mood, K. Naseri, Z. Tahergorabi, M.R. Khazdair, M. Sadeghi, Toxic mechanisms of five heavy metals: mercury, lead, chromium, cadmium, and arsenic, Front. Pharmacol. 12 (2021) 643972, http://dx.doi.org/10.3389/fphar.2021.643972.

[7] S. Mitra, A.J. Chakraborty, A.M. Tareq, T.B. Emran, F. Nainu, A. Khusro, A.M. Idris, M.U. Khandaker, H. Osman, F.A. Alhumaydhi, J. Simál-Gandara, Impact of heavy metals on the environment and human health: novel therapeutic insights to counter the toxicity, J. King Saud Univ. – Sci. 34 (3) (2022) 101865, http://dx.doi.org/10.1016/j.jksus.2022.101865.

[8] G. Genchi, M.S. Sinicropi, G. Lauria, A. Carocci, A. Catalano, The effects of cadmium toxicity, Int. J. Environ. Res. Public Health 17 (11) (2020) 3782, http://dx.doi.org/10.3390/ijerph17113782.

[9] A.L. Wani, A. Ara, J.A. Usmani, Lead toxicity: a review, Interdiscip. Toxicol. 8 (2) (2015) 55–64, http://dx.doi.org/10.1515/intox-2015-0009.

[10] K. Renu, R. Chakraborty, H. Myakala, R. Koti, A.C. Famurewa, H. Madhyastha, B. Vellingiri, A. George, A. Valsala Gopalakrishnan, Molecular mechanism of heavy metals (lead, chromium, arsenic, mercury, nickel and cadmium) – induced hepatotoxicity, Chemosphere 271 (2021) 129735, http://dx.doi.org/10.1016/j.chemosphere.2021.129735.

[11] S. Kumari, M. Kumar, S. Agrawal, A. Kaushik, A. Hashem, E.F. Abd_Allah, A. Kumar, M.C. Garg, Research hotspot and trend analysis of heavy metals decontamination of wastewater in past two decades: A bibliometric analysis, J. Hazard. Mater. Adv. 17 (2025) 100532, http://dx.doi.org/10.1016/j.hazadv.2024.100532.

[12] S. Kumari, J. Chowdhry, R.V.S. Uppaluri, P.K. Patel, Chitosan-based materials: Bridging the gap between pollutant removal and sustainability, Process. Saf. Environ. Prot. 202 (2025) 107671, http://dx.doi.org/10.1016/j.psep.2025.107671.

[13] S. Kumari, J. Chowdhry, M.C. Garg, AI-enhanced adsorption modeling: Challenges, applications, and bibliographic analysis, J. Environ. Manag. 351 (2024) 119968, http://dx.doi.org/10.1016/j.jenvman.2023.119968.

[14] S. Anwar, Q.A. Syed, A. Ullah, M. Arshad, S.N. Khan, M.I. Hussain, Detection of heavy metals using atomic absorption and emission spectroscopy in drinking water of Faisalabad, Pakistan: microbial safety and quality status assessment, Toxin Rev. 43 (1) (2024) 1–17, http://dx.doi.org/10.1080/15569543.2023.2272174.

[15] I.A. Alhagri, A.N. Al-Hakimi, S.M. Al-Hazmy, A.E. Albadri, Determination of trace and heavy metals in bottled drinking water in Yemen by ICP-MS, Results Chem. 8 (2024) 101558, http://dx.doi.org/10.1016/j.rechem.2024.101558.

[16] J.E. Vilasó-Cadre, I.A. Reyes-Domínguez, J.G. González-Fontanet, J. Hidalgo-Viteri, L.A. González-Fernández, M.d.l.Á. Arada-Pérez, G.L. Turdean, Voltammetry and related electrochemical methods based on low-cost instrumentation: a review from basic to advanced, J. Anal. Chem. 79 (2024) 520–539, http://dx.doi.org/10.1134/s1061934824050150.

[17] S. Hu, S. Zhang, J. Qin, K. Cai, C. Peng, L. Luo, Y. Gu, Y. Mei, Simultaneous determination of lead and cadmium in water by metal oxide framework complex-modified glassy carbon electrodes, Microchem. J. 205 (2024) 111154, http://dx.doi.org/10.1016/j.microc.2024.111154.

[18] A. Ait Lahcen, E. Zor, Trace-level analysis of heavy metals in the environment using electrochemical sensors, in: S.A. Ozkan (Ed.), Recent Trends and Perspectives on Electrochemical Sensors for Environmental Monitoring, Elsevier, 2024, pp. 411–445, http://dx.doi.org/10.1016/B978-0-443-13388-6.00013-9.

[19] B. Bansod, T. Kumar, R. Thakur, S. Rana, I. Singh, A review on various electrochemical techniques for heavy metal ions detection with different sensing platforms, Biosens. Bioelectron. 94 (2017) 443–455, http://dx.doi.org/10.1016/j.bios.2017.03.031.

[20] L. Pujol, D. Evrard, K. Groenen-Serrano, M. Freyssinier, A. Ruffien-Cizsak, P. Gros, Electrochemical sensors and devices for heavy metals assay in water: the French groups' contribution, Front. Chem. 2 (2014) 19, http://dx.doi.org/10.3389/fchem.2014.00019.

[21] D. Valera, P.J. Espinoza-Montero, J. Alvarado, P. Carrera, P. Bonilla, L. Cumbal, L. Fernández, Development and evaluation of a glassy carbon electrode modified with silver and mercury nanoparticles for quantification of cysteine rich peptides, Sensors Actuators B: Chem. 253 (2017) 1170–1179, http://dx.doi.org/10.1016/j.snb.2017.07.153.

[22] D. Zhao, T. Wang, D. Han, C. Rusinek, A.J. Steckl, W.R. Heineman, Electrospun carbon nanofiber modified electrodes for stripping voltammetry, Anal. Chem. 87 (18) (2015) 9315–9321, http://dx.doi.org/10.1021/acs.analchem.5b02017.

[23] G. March, T.D. Nguyen, B. Piro, Modified electrodes used for electrochemical detection of metal ions in environmental analysis, Biosensors 5 (2) (2015) 241–275, http://dx.doi.org/10.3390/bios5020241.

[24] J. Liu, J. Li, Y. Xu, S.S. Low, H. Ning, S. Yu, D. Xu, Q. Liu, A fused convolutional transformer for voltammetric electronic tongue analysis tasks, J. Environ. Chem. Eng. 12 (2024) 113462, http://dx.doi.org/10.1016/j.jece.2024.113462.

[25] J.X. Leon-Medina, D.A. Tibaduiza, J.C. Burgos, M. Cuenca, D. Vasquez, Classification of As, Pb and Cd heavy metal ions using square wave voltammetry, dimensionality reduction and machine learning, IEEE Access 10 (2022) 7684–7694, http://dx.doi.org/10.1109/access.2022.3143451.

[26] N. Liu, G. Zhao, G. Liu, Accurate SWASV detection of Cd(II) under the interference of Pb(II) by coupling support vector regression and feature stripping currents, J. Electroanal. Chem. 889 (2021) 115227, http://dx.doi.org/10.1016/j.jelechem.2021.115227.

[27] R.A. Putri, R. Sarno, W.P. Utomo, F. Taufany, K.R. Sungkono, T.C. Amri, A. Kamilah, R. Handayani, S.-S. Lee, A.M. Tjoa, A.A. Sagran, Cross-attention feature fusion network for robust estimation of Cd2+ and Pb2+ in water samples using cyclic voltammetry, Chem. Eng. J. Adv. 26 (2026) 101065, http://dx.doi.org/10.1016/j.ceja.2026.101065.

[28] N. Liu, W. Ye, G. Liu, G. Zhao, Improving the accuracy of stripping voltammetry detection of Cd2+ and Pb2+ in the presence of Cu2+ and Zn2+ by machine learning: understanding and inhibiting the interactive interference among multiple heavy metals, Anal. Chim. Acta 1213 (2022) 339956, http://dx.doi.org/10.1016/j.aca.2022.339956.

[29] J. Zhang, F. Chen, R. Zou, J. Liao, Y. Zhang, Z. Zhu, X. Yan, Z. Jiang, F. Tan, A CNN-based method for heavy-metal ion detection, Appl. Sci. 13 (8) (2023) 4520, http://dx.doi.org/10.3390/app13074520.

[30] N. Yadav, B.M. Maurya, D. Chettri, Pooja, C. Pulwani, M. Jajula, S.S. Kanda, H.W. Suresh Babu, A. Elangovan, P. Velusamy, M. Iyer, B. Vellingiri, Artificial intelligence in heavy metals detection: methodological and ethical challenges, Hyg. Environ. Health Adv. 7 (2023) 100071, http://dx.doi.org/10.1016/j.heha.2023.100071.

[31] M. Molinara, R. Cancelliere, A. Di Tinno, L. Ferrigno, M. Shuba, P. Kuzhir, A. Maffucci, L. Micheli, A deep learning approach to organic pollutants classification using voltammetry, Sensors 22 (20) (2022) 8032, http://dx.doi.org/10.3390/s22208032.

[32] Z. Yu, C. Zhang, N. Xiong, F. Chen, A new random forest applied to heavy metal risk assessment, Comput. Syst. Sci. Eng. 40 (1) (2022) 207–221, http://dx.doi.org/10.32604/csse.2022.018301.

[33] R.A. Putri, R. Sarno, W.P. Utomo, F. Taufany, D. Sunaryono, S.I. Sabilla, S.-S. Lee, K.R. Sungkono, T.C. Amri, R.A. Wahyuono, R. Handayani, M.A. Prakasa, Detecting water contaminants using an electrochemical sensor based on improved Transformer Encoder with Genetic Algorithm, J. Water Process. Eng. 87 (2026) 110111, http://dx.doi.org/10.1016/j.jwpe.2026.110111.

[34] N.D. Takarina, N. Matsue, E. Johan, A. Adiwibowo, M.F.N.K. Rahmawati, S.A. Pramudyawardhani, T. Wukirsari, Machine learning using random forest to model heavy metals removal efficiency using a zeolite-embedded sheet in water, Glob. J. Environ. Sci. Manag. 10 (1) (2024) 321–336, http://dx.doi.org/10.22034/gjesm.2024.01.20.

[35] R.A. Cottis, Interpretation of electrochemical noise data, Corrosion 57 (3) (2001) 265–285, http://dx.doi.org/10.5006/1.3290350.

[36] U. Bertocci, F. Huet, R.P. Nogueira, P. Rousseau, Drift removal procedures in the analysis of electrochemical noise, Corrosion 58 (4) (2002) 337–347, http://dx.doi.org/10.5006/1.3287684.

[37] F. Mansfeld, Z. Sun, C.H. Hsu, A. Nagiub, Concerning trend removal in electrochemical noise measurements, Corros. Sci. 43 (2) (2001) 341–352, http://dx.doi.org/10.1016/s0010-938x(00)00064-0.

[38] E.A. Astafev, Electrochemical noise of a Li-ion battery: measurement and spectral analysis, J. Solid State Electrochem. 23 (4) (2019) 1145–1153, http://dx.doi.org/10.1007/s10008-019-04209-5.

[39] B. Kaulakys, M. Alaburda, Modeling scaled processes and 1/f noise by nonlinear stochastic differential equations, J. Stat. Mech. Theory Exp. 2009 (2) (2009) P02051.

[40] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (8) (1964) 1627–1639, http://dx.doi.org/10.1021/ac60214a047.

[41] M. Schmid, D. Rath, U. Diebold, Why and how Savitzky-Golay filters should be replaced, ACS Meas. Sci. Au 2 (2) (2022) 185–196, http://dx.doi.org/10.1021/acsmeasuresciau.1c00054.

[42] D.L. Donoho, I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, J. Amer. Statist. Assoc. 90 (432) (1995) 1200–1224, http://dx.doi.org/10.1080/01621459.1995.10476626.

[43] D.R. Wijaya, R. Sarno, E. Zulaika, Information quality ratio as a novel metric for mother wavelet selection, Chemometr. Intell. Lab. Syst. 160 (2017) 59–71, http://dx.doi.org/10.1016/j.chemolab.2016.11.012.

[44] M.S.H. Ardani, R. Sarno, M. Malikhah, D.P. Purbawa, S.I. Sabilla, K.R. Sungkono, C. Fatichah, D. Sunaryono, R.I. Susilo, A new approach to signal filtering method using K-means clustering and distance-based Kalman filtering, Sens. Bio-Sensing Res. 38 (2022) 100539, http://dx.doi.org/10.1016/j.sbsr.2022.100539.

[45] C.M. Bishop, Training with noise is equivalent to Tikhonov regularization, Neural Comput. 7 (1) (1995) 108–116, http://dx.doi.org/10.1162/neco.1995.7.1.108.

[46] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.

[47] Q. Wen, L. Sun, F. Yang, X. Song, J. Gao, X. Wang, H. Xu, Time series data augmentation for deep learning: a survey, IJCAI Int. Jt. Conf. Artif. Intell. (2021) 4653–4660, http://dx.doi.org/10.24963/ijcai.2021/631.

[48] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (1) (2019) 60, http://dx.doi.org/10.1186/s40537-019-0197-0.

[49] I.J. Goodfellow, J. Shlens, C. Szegedy, Explaining and harnessing adversarial examples, in: International Conference on Learning Representations, 2015.

[50] N. Akhtar, A. Mian, Threat of adversarial attacks on deep learning in computer vision: a survey, IEEE Access 6 (2018) 14410–14430, http://dx.doi.org/10.1109/access.2018.2807385.

[51] S.H. Lim, N.B. Erichson, F. Utrera, W. Xu, M.W. Mahoney, Noisy feature mixup, in: International Conference on Learning Representations, 2022.

[52] A. Camuto, M. Willetts, U. Simsekli, S. Roberts, C. Holmes, Explicit regularisation in Gaussian noise injections, Adv. Neural Inf. Process. Syst. 34 (2021) 16603–16614.

[53] S. Jenni, P. Favaro, On stabilizing generative adversarial training with noise, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 12141–12150, http://dx.doi.org/10.1109/cvpr.2019.01242.

[54] F. Taufany, R.A. Putri, R. Sarno, W.P. Utomo, K.R. Sungkono, A.A. Sagran, iTransKAN: A deep chemometric model for enhanced heavy metal quantification via multi-scan cyclic voltammetry, Chem. Eng. J. Adv. 27 (2026) 101296, http://dx.doi.org/10.1016/j.ceja.2026.101296.

[55] P. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. 15 (2) (1967) 70–73, http://dx.doi.org/10.1109/tau.1967.1161901.

[56] A.D. McNaught, A. Wilkinson, Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"), Blackwell Scientific Publications, Oxford, 1997.

[57] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, http://dx.doi.org/10.1023/a:1010933404324.

[58] P. Heidari, A. Milan, Combining K-fold cross validation with Bayesian hyperparameter optimization for accuracy enhancement of land cover and land use classification, Sci. Rep. 15 (1) (2025) 39758, http://dx.doi.org/10.1038/s41598-025-23336-w.

[59] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.