Supporting Information

Interpretable Machine Learning Model-Driven Electrochemical Impedance Spectroscopic Analysis for Determination of Nitrate, Nitrite Ammonium Ions in Water

Reji K.[a], Ramakrishnan Vishnuraj[b,c], Adithyakrishna Narayanaswamy[d], Vishnuvarthan Rajagopal[e], Vishnu S.[e]* Murali Rangarajan[b,c,f]*

[a] Department of Electrical and Electronics Engineering, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, India.

[b]Center of Excellence in Advanced Materials and Green Technologies, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, India.

[c]Department of Chemical Engineering and Materials Science, Amrita School of Engineering Coimbatore, Amrita Vishwa Vidyapeetham, India.

[d]Department of Chemical Engineering, University of Florida, Gainesville, Florida.

[e]Department of Computer Science and Engineering, Amrita School of Computing Coimbatore, Amrita Vishwa Vidyapeetham, India.

[f]Gurukripa Electrolyzers Private Limited, Coimbatore, India.

**Corresponding Author, Email: r_murali@cb.amrita.edu; s_vishnu1@cb.amrita.edu.*

Measurements Techniques used:

Conductivity and pH were analysed for the simulated water samples using LABMAN Multiparameter LMMP-30 instrument. Electrochemical impedance spectra (EIS) were obtained using CH Instruments, Inc., (CHI604E. USA) electrochemical workstation. The electrochemical cell constitutes of working & counter electrodes of stainless steel and Ag/AgCl reference electrode were immersed in the simulated water solution. A 10mV AC signal superimposed on the open circuit potential (OCP) of the system and the frequency response of the impedance was measured in the range from 1Hz - 1MHz. Frequency response of the measured impedance was depicted either as a Bode plot or as a Nyquist plot. Appropriate equivalent electrical circuit models were developed to capture the behaviour of the stainless steel-solution interface. The magnitudes of the components of the equivalent circuit were obtained by non-linear regression (Aftermath®). Fits with rms errors less than 5% were accepted. Further, to develop the electrochemical impedance-based sensor for the analytes NO_3^-, NO_2^- and NH_4^+, known quantities of the three constituents

were added to the base solution thus prepared. The amounts of the three analytes added are presented in **Table S1**.

Table S1. Amounts of NO_3^-, NO_2^- and NH_4^+ added to the Base Solution to Prepare Simulated Solutions:

Set - I Nitrate Concentration (ppm)	**Set - II Nitrite Concentration (ppm)**	**Set - III Ammonium Concentration (ppm)**
7	0.5	0.5
14	1.5	1.5
21	3.0	3.0
28	4.5	4.5
35	6.0	6.0
42	8.5	8.5
50	10.0	10.0
NO_2^- : 4.5 ppm NH_4^+ : 4.5 ppm	**NO_3^- : 28 ppm NH_4^+ : 4.5 ppm**	**NO_3^- : 28 ppm NO_2^- : 4.5 ppm**

Table S2. Measured pH value for NO_3^-, NO_2^- and NH_4^+ added simulated solution:

DI water pH: **0.055** μS/cm					
Concentration (ppm)	**pH of varied NO_3^- ions**	**Concentration (ppm)**	**pH of varied NO_2^- ions**	**Concentration (ppm)**	**pH of varied NH_4^+ ions**
7	7.305	0.5	7.347	0.5	7.589
14	6.879	1.5	7.111	1.5	7.158
21	6.586	3.0	6.810	3.0	6.811
28	6.214	4.5	6.637	4.5	6.605
35	6.155	6.0	6.508	6.0	6.375
42	6.078	8.5	6.395	8.5	6.202
50	5.992	10	6.319	10	6.143

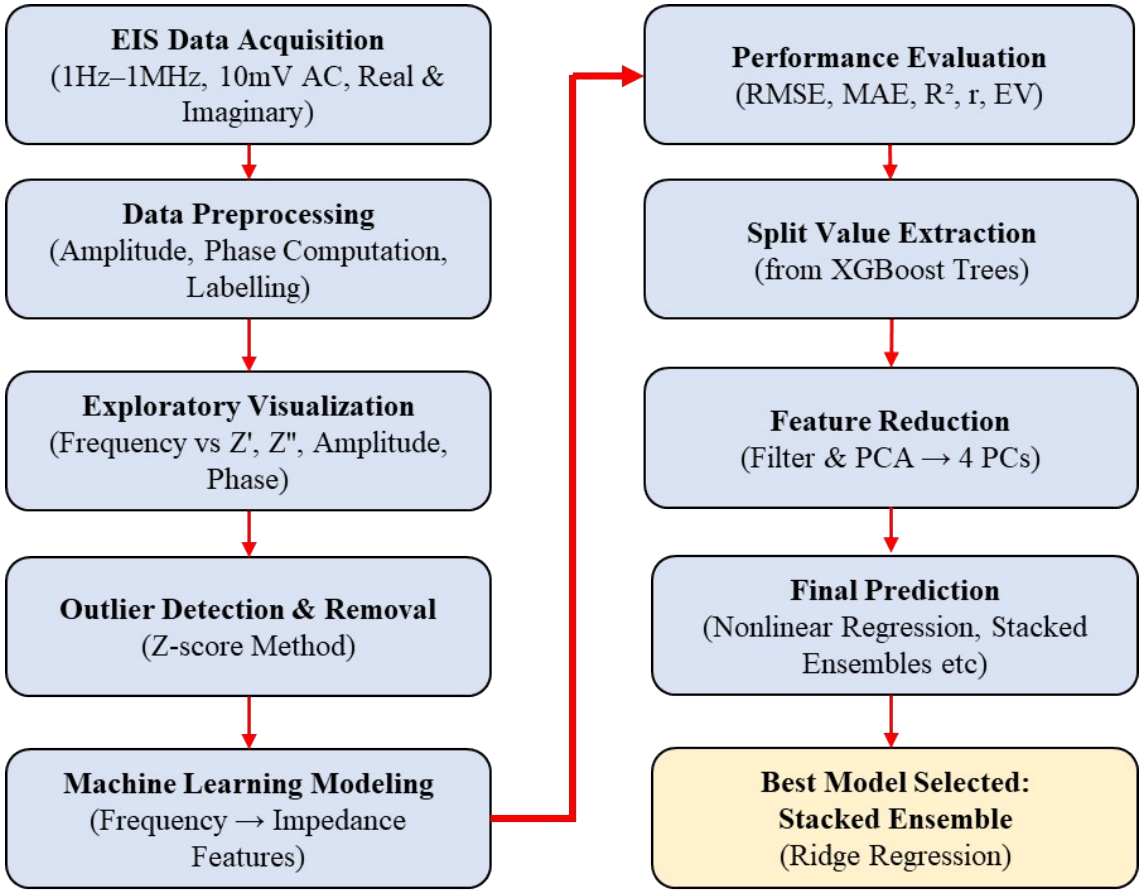

Figure S1 Flowchart of the methodology for NO_3^-, NO_2^-, and NH_4^+ concentration prediction using EIS and machine learning.

Table S3. Evaluation metrics for all ML models of NO_3^-, NO_2^- and NH_4^+.

Parameter	Analyte	ML Model	MSE	RMSE	MAE	R^2	EV
Real value	Nitrate (28ppm)	RF	632.826	25.156	11.011	0.9976	0.9977
		KNN	8297.438	91.090	43.059	0.9680	0.9698
		SVR	227207.906	476.663	203.481	0.1230	0.2675
		ANN	4642256.946	2154.590	1298.716	-16.9190	-16.8563
		XGBoost	**0.004**	**0.064**	**0.038**	**0.9999**	**0.9999**
	Nitrite (4.5ppm)	RF	1144.491	33.830	15.460	0.9978	0.9980
		KNN	14882.268	121.993	60.305	0.9717	0.9737
		SVR	415173.416	644.340	287.483	0.2110	0.3326
		ANN	5411004.249	2326.157	1627.279	-9.2830	-9.1773
		XGBoost	**0.010**	**0.099**	**0.055**	**0.9999**	**0.9999**

	Ammonium (4.5ppm)	RF	1185.583	34.432	15.044	0.9979	0.9980
		KNN	16234.416	127.414	61.077	0.9709	0.9727
		SVR	478630.987	691.832	308.065	0.1412	0.2725
		ANN	5262061.103	2293.918	1558.762	-8.4416	-8.3891
		XGBoost	**0.015**	**0.124**	**0.074**	**0.9999**	**0.9999**
Imaginary value	Nitrate (28ppm)	RF	5943.444	77.094	30.157	0.9962	0.9964
		KNN	74177.237	272.355	113.039	0.9532	0.9543
		SVR	1825222.971	1351.008	576.276	-0.1524	0.0038
		ANN	4296007.676	2072.681	1149.936	-1.7125	-1.7099
		XGBoost	**0.024**	**0.154**	**0.097**	**0.9999**	**0.9999**
	Nitrite (4.5ppm)	RF	9172.728	95.774	38.521	0.9962	0.9963
		KNN	113609.447	337.060	143.475	0.9534	0.9546
		SVR	2780516.113	1667.488	722.362	-0.1417	0.0058
		ANN	7146115.128	2673.222	1475.143	-1.9342	-1.9326
		XGBoost	**0.058**	**0.241**	**0.148**	**0.9999**	**0.9999**
	Ammonium (4.5ppm)	RF	10502.596	102.482	39.675	0.9966	0.9967
		KNN	145174.910	381.018	158.156	0.9534	0.9546
		SVR	3560285.303	1886.872	802.665	-0.1424	0.0081
		ANN	6095350.062	2468.876	1367.717	-0.9558	-0.9163
		XGBoost	**0.101**	**0.318**	**0.203**	**0.9999**	**0.9999**
Amplitude	Nitrate (28ppm)	RF	6434.928	80.218	29.981	0.9962	0.9963
		KNN	80602.759	283.906	115.869	0.9522	0.9536
		SVR	1850489.748	1360.327	539.792	-0.0971	0.0699
		ANN	6655575.773	2579.840	1631.500	-2.9458	-2.9164
		XGBoost	**0.018**	**0.136**	**0.091**	**0.9999**	**0.9999**
	Nitrite (4.5ppm)	RF	9979.890	99.899	38.086	0.9962	0.9964
		KNN	123961.832	352.082	146.686	0.9534	0.9549
		SVR	2819016.422	1678.993	670.207	-0.0608	0.0987
		ANN	8550436.946	2924.113	2014.805	-2.2176	-2.0581
		XGBoost	**0.093**	**0.304**	**0.192**	**0.9999**	**0.9999**
	Ammonium (4.5ppm)	RF	11399.578	106.769	39.269	0.9966	0.9967
		KNN	157179.674	396.459	162.091	0.9534	0.9548
		SVR	3663939.463	1914.142	763.285	-0.0870	0.0756
		ANN	9212142.981	3035.151	2018.662	-1.7329	-1.6051
		XGBoost	**0.036**	**0.190**	**0.129**	**0.9999**	**0.9999**
Phase	Nitrate (28ppm)	RF	0.321	0.567	0.483	0.9993	0.9993
		KNN	4.239	2.059	1.699	0.9901	0.9903
		SVR	449.329	21.197	14.061	-0.0468	0.0800
		ANN	2384012.562	1544.025	606.093	-5552.9872	4786.0572
		XGBoost	**0.000**	**0.004**	**0.002**	**0.9999**	**0.9999**
	Nitrite	RF	0.313	0.559	0.486	0.9992	0.9992

	(4.5ppm)	KNN	4.107	2.027	1.687	0.9900	0.9901
		SVR	388.657	19.714	12.817	0.0562	0.1655
		ANN	4211991.147	2052.314	800.073	10227.4467	8783.0821
		XGBoost	**0.000**	**0.002**	**0.001**	**0.9999**	**0.9999**
		RF	0.305	0.552	0.481	0.9993	0.9993
	Ammonium (4.5ppm)	KNN	3.989	1.997	1.667	0.9904	0.9906
		SVR	411.250	20.279	13.722	0.0105	0.1043
		ANN	2294549.145	1514.777	589.279	-5519.6885	4771.0737
		XGBoost	**0.000**	**0.004**	**0.003**	**0.9999**	**0.9999**

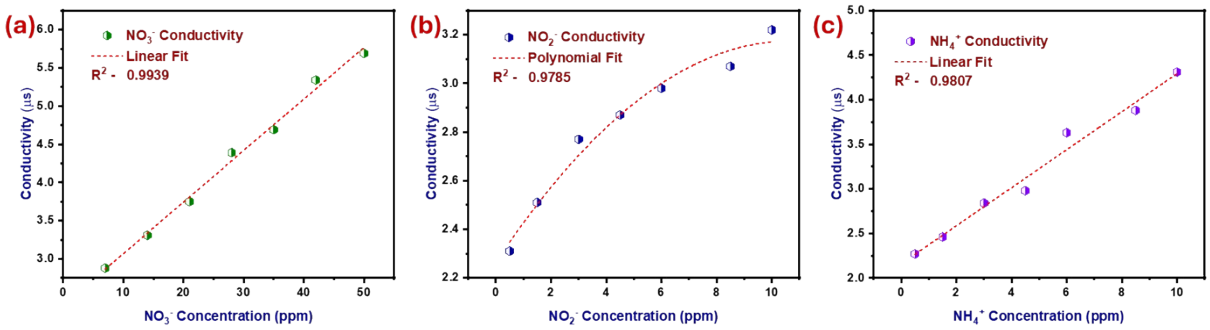

Figure S2. Conductivity Plot for **(a)** Nitrate **(b)** Nitrite and **(c)** Ammonium ions in simulated solution.

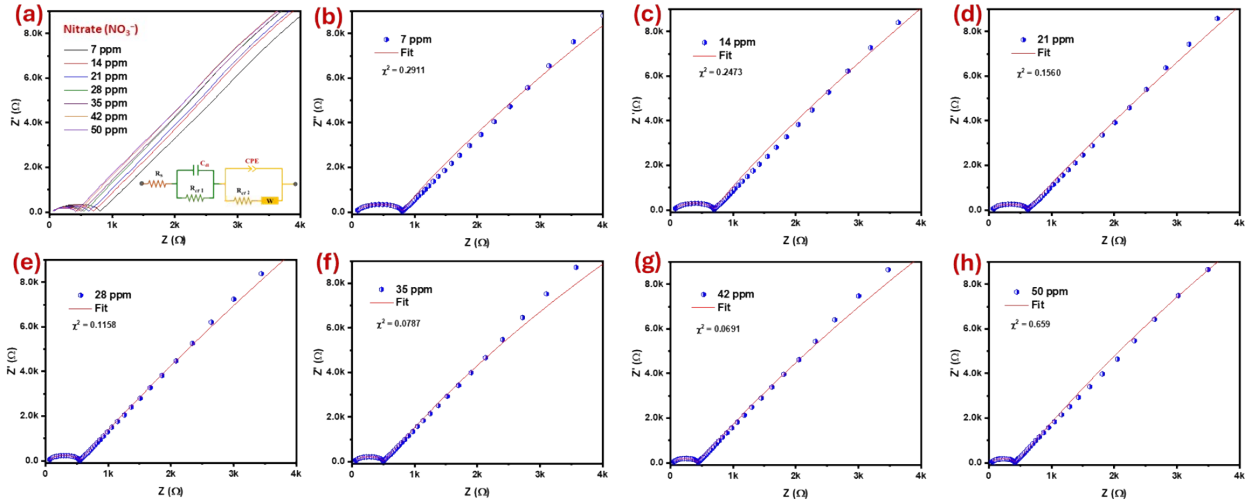

Figure S3 (a) EIS Spectra of varying NO_3^- in simulated solution and **(b-h)** EIS Spectra fitted with equivalent circuit for varying NO_3^- concentration (7 – 50 ppm).

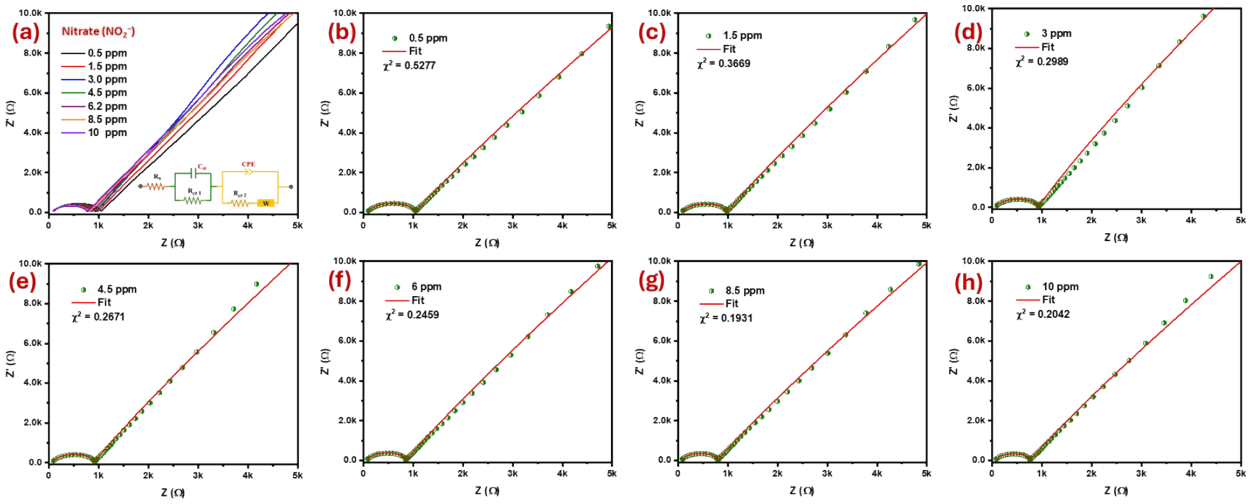

Figure S4 (a) EIS Spectra of varying NO_2^- in simulated solution and **(b-h)** EIS Spectra fitted with equivalent circuit for varying NO_2^-concentration (0.5 – 10 ppm).

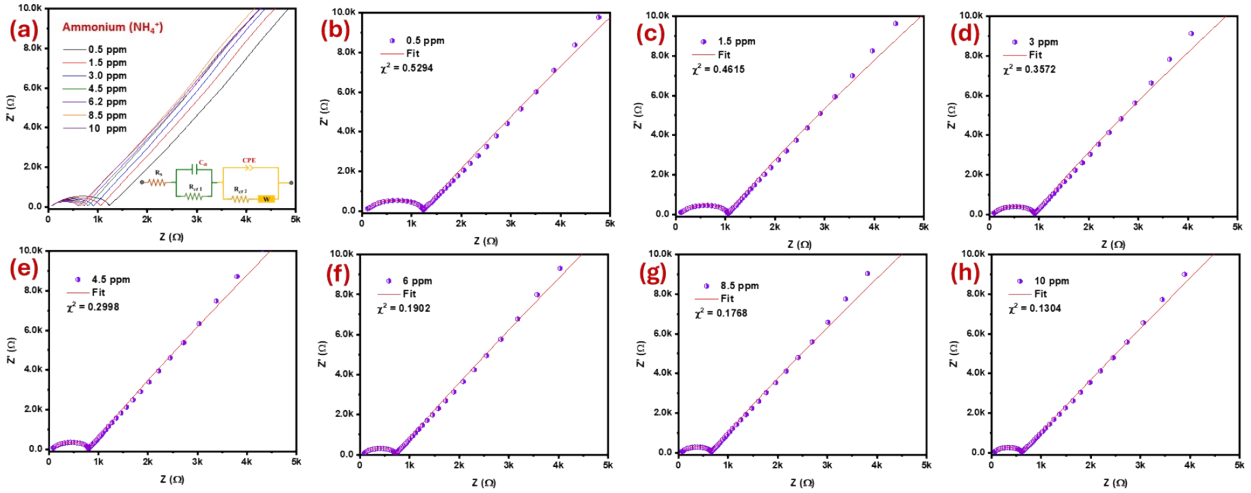

Figure S5 (a) EIS Spectra of varying NH_4^+ in simulated solution and **(b-h)** EIS Spectra fitted with equivalent circuit for varying NH_4^+ concentration (0.5 – 10 ppm).

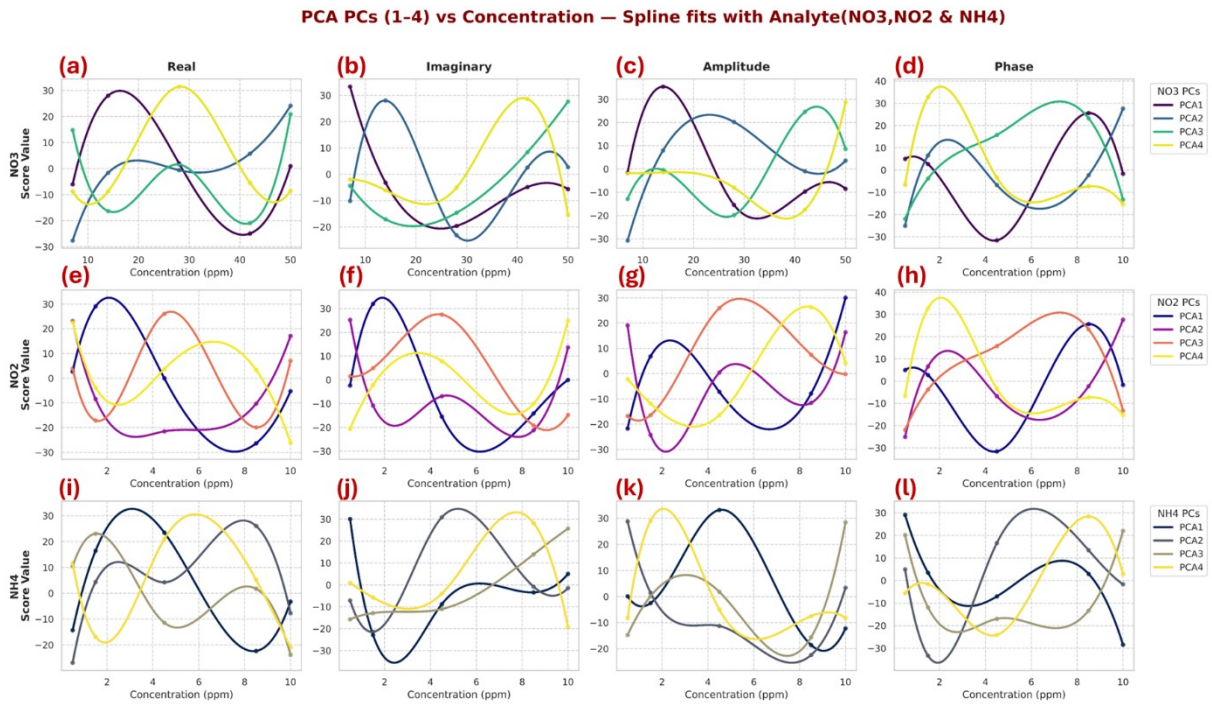

Figure S6 Spline fits of PC_1-PC_4 versus concentrations of **(a-d)** NO_3^-, **(e-h)** NO_2^-, and **(i-l)** NH_4^+ ions.

S1. ***Dataset Preparation:***

For each experiment, the impedance was recorded as the real component (Z') and the imaginary component (Z''), from which the amplitude and phase of the impedance were derived using the following eqns:

$$|Z| = \sqrt{Z'^2 + Z''^2} \qquad \text{(S1)}$$

$$\theta = \tan^{-1}\frac{Z''}{Z'} \qquad \text{(S2)}$$

Eqns. S1 and **S2** provided two additional parameters that offered a more comprehensive characterization of the electrochemical response.

The concentrations selected for each analyte were as follows: five levels for nitrate (7, 14, 28, 42, and 50ppm) and seven levels for both nitrite and ammonium (0.5, 1.5, 4.5, 8.5, and 10ppm). **Table S3** presents an organized summary of the concentration ranges and fixed median values.

Table S3: Concentration ranges and fixed constants for nitrate, nitrite, and ammonia experiments.

Analyte (Target)	Concentration Levels (ppm)	Constants (ppm)
Nitrate (NO_3^-)	7, 14, 28, 42, 50	Nitrite = 4.5, Ammonia = 4.5
Nitrite (NO_2^-)	0.5, 1.5, 4.5, 8.5, 10	Nitrate = 28, Ammonia = 4.5
Ammonium (NH_4^+)	0.5, 1.5, 4.5, 8.5, 10	Nitrate = 28, Nitrite = 4.5

S2. ***Exploratory Visualization:***

Electrochemical impedance spectroscopy describes impedance as a complex function of angular frequency (ω) **(eqn. S3)**

$$Z(\omega) = Z^{'}(\omega) + jZ"(\omega) \quad \text{(S3)}$$

Where Z'(ω) and Z"(ω) represent the real and imaginary components, respectively, and ω = 2πf is the angular frequency. For a simple parallel RC system, this relationship can be expressed as **eqn. S4**

$$Z(\omega) = \frac{R}{1 + j\omega RC} \quad \text{(S4)}$$

which can be separated into the real and imaginary parts **(eqn. S5)**

$$Z^{'}(\omega) = \frac{R}{1 + (\omega RC)^2} \quad ; \quad Z"(\omega) = \frac{\omega R^2 C}{1 + (\omega RC)^2} \quad \text{(S5)}$$

From these components, the impedance amplitude and phase are derived **eqn. S6** and **S7**

$$|Z(\omega)| = \sqrt{Z'(\omega)^2 + Z"(\omega)^2} \quad \text{(S6)}$$

$$\theta(\omega) = \tan^{-1}\frac{Z"(\omega)}{Z'(\omega)} \quad \text{(S7)}$$

Plots of frequency versus Z'(ω), Z"(ω), |Z(ω)| and θ(ω) were generated for all concentration levels of NO_3^-, NO_2^-, and NH_4^+.

S3. ***Data Preprocessing and Outlier Removal:***

A systematic outlier detection procedure was employed using the Z-score method which assures normal distribution of the measured data. For each impedance parameter (real component, imaginary component, amplitude, and phase), the Z-score was calculated as **eqn. S8**

$$Z = \frac{x - \mu}{\sigma} \quad \text{(S8)}$$

where μ and σ are the mean and standard deviation of the feature distribution, respectively. These were computed as **eqn. S9** and **S10**

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \text{(S9)}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad \text{(S10)}$$

Here, N denotes the total number of data points in the dataset, x_i represents the i^{th} data value, and x in **eqn. S8** corresponds to the individual observation for which the Z-score is calculated.

A data point was considered an outlier if it satisfied the condition given in **eqn. S11**

$$|Z| > Z_{threshold} \quad \text{(S11)}$$

where $Z_{threshold} = 3$ was selected in this study. All such outliers were removed from the dataset to improve consistency and reduce noise.

S4. Selection criteria of *Machine Learning Algorithms:*

To establish the relationship between frequency-dependent impedance parameters and analyte concentration, several nonlinear machine learning algorithms were systematically explored. A variety of approaches were investigated, including Random Forests, k-Nearest Neighbors (KNN), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost). While most of these models demonstrated a partial ability to approximate the complex spectral trends, none - except XGBoost - consistently provided a reliable and robust fit across nitrate, nitrite, and ammonium datasets. In many cases, the models either tended to overfit specific regions of the spectra or exhibited poor generalization when applied to unseen concentration levels, limiting their practical applicability is shown in **Figure S7**. In contrast, XGBoost emerged as the only method capable of capturing the intricate nonlinearities of the data while maintaining stability across all experimental conditions.

For consistency with electrochemical impedance spectroscopy (EIS) notation, the real and imaginary parts of impedance are denoted as Z' and Z'', respectively. In the regression framework, we adopt the equivalent notation $Z_r \equiv Z'$ and $Z_i \equiv Z''$ for clarity, while the amplitude and phase angle remain as $|Z|$ and ϕ.

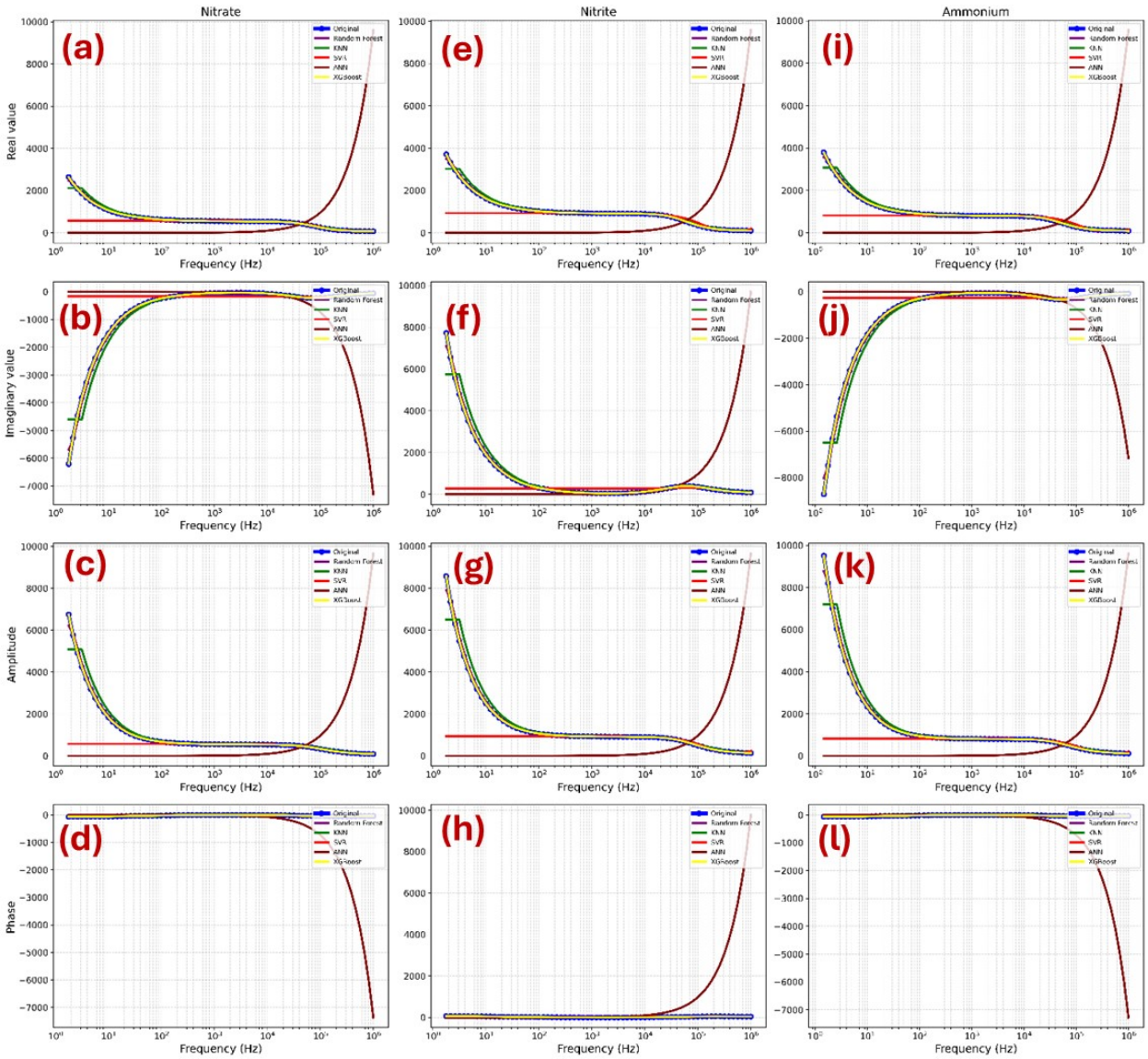

Figure S7 Machine Learning model comparison at median concentrations of **(a-d)** Nitrate (28ppm), **(e-h)** Nitrite (4.5ppm) and **(i-l)** Ammonium (4.5ppm).

In this work, XGBoost was employed to model the spectral response of each analyte (NO_3^-, NO_2^-, and NH_4^+). Frequency served as the independent variable, while the impedance parameters - the real part, imaginary part, amplitude, and phase, i.e., Z', Z", |Z| and ϕ respectively were treated as dependent variables. All these parameters are functions of both the analyte concentration and frequency, which can be expressed for nitrate, nitrite, and ammonium in **eqn. S12 - S14**:

For Nitrate (NO_3^-),

$$Z_r = Z_r\left(C_{NO_3^-}, f\right);\ Z_i = Z_i\left(C_{NO_3^-}, f\right);\ |Z| = |Z|\left(C_{NO_3^-}, f\right);\ \phi = \phi\left(C_{NO_3^-}, f\right) \quad \text{(S12)}$$

For Nitrite (NO_2^-),

$$Z_r = Z_r\,(C_{NO_2^-}, f)\,;\; Z_i = Z_i\,(C_{NO_2^-}, f)\,;\; |Z| = |Z|\,(C_{NO_2^-}, f)\,;\; \phi = \phi(C_{NO_2^-}, f) \quad \text{(S13)}$$

For Ammonium (NH_4^+)

$$Z_r = Z_r\,(C_{NO_4^+}, f)\,;\; Z_i = Z_i\,(C_{NO_4^+}, f)\,;\; |Z| = |Z|\,(C_{NO_4^+}, f)\,;\; \phi = \phi(C_{NO_4^+}, f) \quad \text{(S14)}$$

Here, $C_{NO_3^-}$, $C_{NO_2^-}$, and $C_{NH_4^+}$ denote the concentrations of nitrate, nitrite, and ammonium, respectively, expressed in parts per million (ppm), while f represents the frequency in Hz.

However, the regression models that ultimately need to be developed are the inverse mappings, where concentration is expressed as a function of impedance features across frequencies, as shown in **eqn. S15 - S17**:

For Nitrate (NO_3^-):

$$C_{NO_3^-} = C_{NO_3^-}(Zr(f), Zi\,(f), |Z|\,(f), \phi\,(f)) \quad \text{(S15)}$$

For Nitrite (NO_2^-):

$$C_{NO_2^-} = C_{NO_2^-}(Zr(f), Zi\,(f), |Z|\,(f), \phi\,(f)) \quad \text{(S16)}$$

For Ammonium (NH_4^+):

$$C_{NH_4^+} = C_{NH_4^+}(Zr(f), Zi\,(f), |Z|\,(f), \phi\,(f)) \quad \text{(S17)}$$

This formulation highlights that while impedance features are inherently functions of both frequency and concentration **eqn. S12 - S14**, the inverse problem - predicting concentration from spectral data (**eqn. S15 - S17**) - requires the construction of robust regression models.

S5. ***Performance evaluation:***

The coefficient of determination (R^2) (**eqn. S18**) measures how well the model explains the variance in the experimental data. A perfect score of $R^2 = 1.0$ was consistently obtained, confirming that the model fully captured the frequency–impedance relationships.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad \text{(S18)}$$

Here, N denotes the total number of data points, y_i represents the i^{th} experimental value, $\hat{y}_i$ is the corresponding predicted value, and $\bar{y}$ is the mean of the experimental values.

The mean squared error (MSE) (**eqn. S19**) represents the average squared deviation between predicted and experimental values, penalizing large errors more heavily. Its square root, the root means square error (RMSE) (**eqn. S20**), provides an interpretable measure of the typical error magnitude. The mean absolute error (MAE) (**eqn. S21**) instead averages absolute deviations, offering a more intuitive sense of prediction accuracy that is less influenced by extreme values.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_{i\_}\hat{y}_i)^2 \tag{S19}$$

$$RMSE = \sqrt{(\frac{1}{N}\sum_{i=1}^{N}(y_{i\_}\hat{y}_i)^2)} \tag{S20}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(y_{i\_}\hat{y}_i)| \tag{S21}$$

Finally, the explained variance (EV) (**eqn. S22**) evaluates how much of the variability in the data is preserved in the predictions. An EV value of 1.0 indicates that the model reproduces the spread of the experimental data without loss of information.

$$EV = 1 - \frac{Var(y - \hat{y})}{Var\ (y)} \tag{22}$$

Figure S7 illustrates the mean evaluation metrics of all the machine learning models represented through bar diagrams.

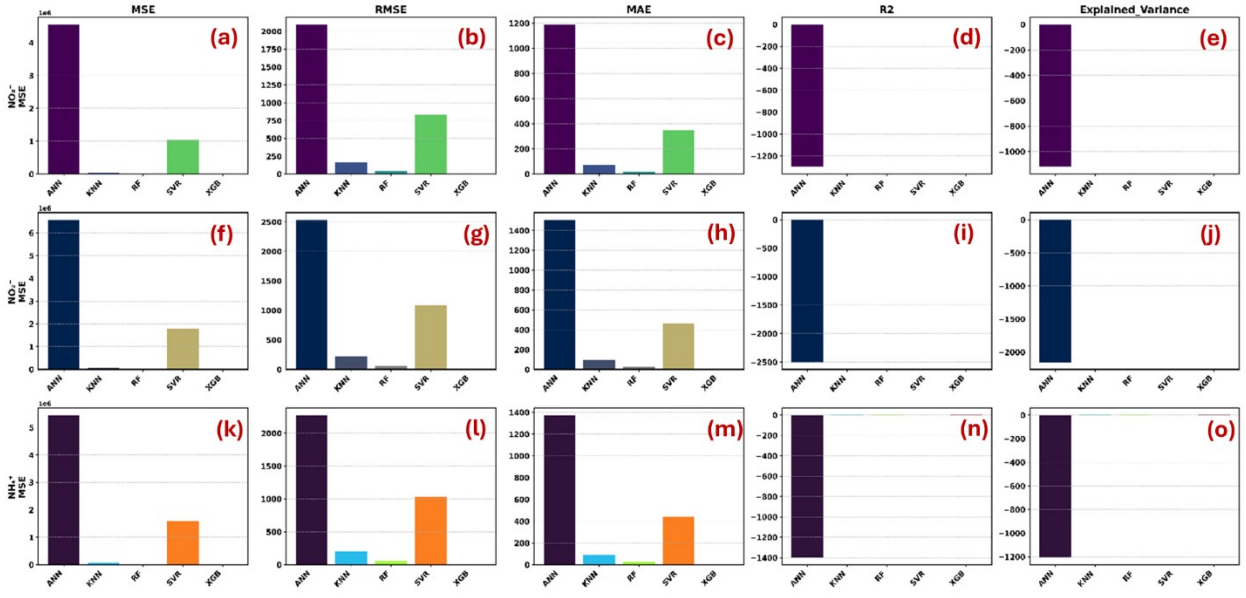

Figure S8 Bar diagram representation of Mean Evaluation Metrics for Different Machine Learning Models for **(a-e)** NO_3^-, **(f-j)** NO_2^- and **(k-o)** NH_4^+ ions.

S6. ***Extraction of XGBoost split values:***

In tree-based models such as XGBoost, each non-leaf node corresponds to a split threshold on the input variable. In this study, the input was frequency, and each split value, denoted as f_s, represented a critical frequency at which the model detected significant changes in the impedance response. Extracting these split values therefore allowed the conversion of XGBoost from a black-box predictor into an interpretable tool for feature generation.

For each analyte (NO_3^-, NO_2^-, and NH_4^+), split values were extracted separately for the four impedance parameters: the real part (Z_r), imaginary part (Z_i), amplitude (|Z|), and phase (ϕ). The resulting datasets can be formally expressed as **eqn. S23:**

$$D_{Zr} = \{(f_s, C)\};\ D_{Zi} = \{(f_s, C)\};\ D_{|Z|} = \{(f_s, C)\};\ D_{\phi} = \{(f_s, C)\} \qquad \text{(S23)}$$

Here, D denotes the structured dataset extracted from the XGBoost split values, f_s are the frequency thresholds identified at non-leaf nodes, and C represents the concentration of the analyte ((NO_3^-, NO_2^-, and NH_4^+) Each dataset was compiled into CSV format, with frequency thresholds as the independent variable and concentration as the dependent variable.

S7. ***Final concentration prediction:***

After dimensionality reduction, the first six principal components (PC_1-PC_4) were used as the predictive descriptor vector $X = [PC_1 \ldots\ldots PC_4]^T$. To capture both global and local nonlinear dependencies between impedance-derived features and analyte concentration, a two-stage nonlinear stacking ensemble was adopted.

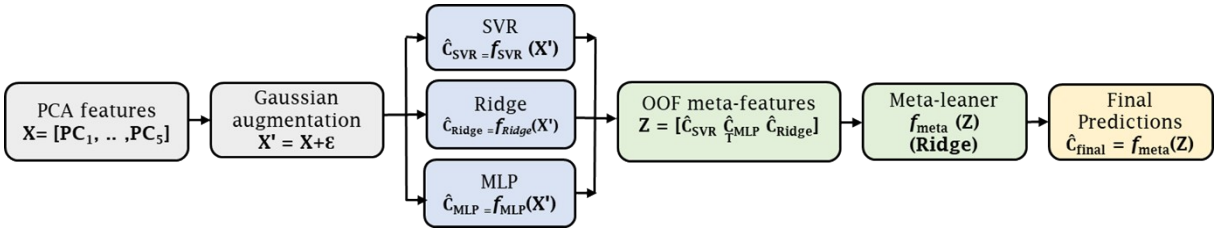

Figure S9 Schematic of the two-stage nonlinear stacking ensemble used for final concentration prediction.

The overall architecture is summarized in **Figure S9.** PCA features (PC_1-PC_4) are perturbed by small Gaussian noise for augmentation, then passed to three base learners (SVR, MLP, Ridge). The out-of-fold (OOF) base predictions form the meta-feature vector, which is input to the meta-learner to produce the final concentration estimate $\hat{C}_{final}$.

S7.1. Stage I: ***Base-learner modelling:***

Three base regressors were trained independently using Group K-Fold cross-validation (groups based on concentration levels) to avoid data leakage between concentration groups. Each base

learner is represented as a mapping from the PCA feature vector X to an intermediate predicted concentration $\hat{C}_i$.The three base mappings are written explicitly as **eqn. S24 – S26**

$$\hat{C}_{SVR} = f_{SVR}(X) = \sum_{j=1}^{n} \alpha_j K(X_j X) + b \quad \text{(S24)}$$

$$\hat{C}_{MLP} = f_{MLP}(X) = \sigma_2(W_2\, \sigma_1(W_1 X + b_1) + b_2) \quad \text{(S25)}$$

$$\hat{C}_{Ride} = f_{Ridge}(X) = X\, \beta_{Ridge} + b \quad \text{(S26)}$$

where, $\hat{C}_{SVR}, \hat{C}_{MLP}, \hat{C}_{Ridge}$: out-of-fold predicted concentrations produced by the SVR, MLP, and Ridge base learners, respectively.

$f_{SVR}, f_{MLP}, f_{Ridge}$: mapping functions implemented by the respective base regressors.

$K(X_j X)$: kernel function (radial-basis function used in this work) between training exemplar X_j and input X.

β : Ridge regression coefficient.

α_j, b : learned dual coefficients and intercept in the SVR representation.

W_1, W_2: weight matrices of the first and second layers in the MLP; b_1, b_2 their corresponding biases.

σ_1, σ_2: nonlinear activation functions (e.g., ReLU, tanh) used in the hidden and output layers.

$T_k(X)$: k^{th} regression tree in the gradient-boosting ensemble; $\gamma_k$its scalar weight.

To improve smoothness and generalization for the meta stage, a small Gaussian perturbation (**eqn. S27**) was applied in the PCA space prior to model training:

$$X' = X + \varepsilon \text{ ;} \qquad \varepsilon \sim N(0,\sigma^2 \mathrm{I}); \quad \sigma \approx 0.002 \qquad \text{(S27)}$$

where X' is augmented PCA vector after small Gaussian perturbation. $\varepsilon \sim N(0,\sigma^2 \mathrm{I})$ is the Gaussian noise vector used for PCA-space augmentation, scaled relative to the magnitude of the PCA components. I denotes the identity matrix. Augmented samples were used only for training; final model evaluation was performed on the original (non-augmented) measurements (see **Section 3.7**).

The out-of-fold (OOF) predictions from each base learner for a given sample form the meta-feature vector Z,(**eqn. S28**) is,

$$Z = [\hat{C}_{SVR}, \hat{C}_{MLP}, \hat{C}_{Ridge}]^T \quad \text{(S28)}$$

$Z = [\hat{C}_{SVR}, \hat{C}_{MLP}, \hat{C}_{Ridge}]^T$ is the meta-feature vector of OOF base predictions used to train the meta learner.

S7.2. Stage II: *Meta-learner integration:*

The meta-learner f_{meta} integrates the base learners' responses to produce the final concentration ($\hat{C}_{final}$) estimate using **eqn. S29**:

$$\hat{C}_{final} = f_{meta}(Z) \quad \text{(S29)}$$

$f_{meta}(.)\ is\ the$ meta-learner function (nonlinear mapping from Z to ˆCfinal), implemented as Ridge regression in the present work.

The above formulation describes a physically interpretable and statistically robust mapping from impedance derived PCA features to analyte concentration values. **Figure S9** provides a schematic overview of the pathway described by **eqn. S24– S29**.