Microfluidic-Electrochemical Sensor utilizing Statistical Modeling for Enhanced Nitrate Detection in Surface Water towards Environmental Monitoring

Sai Kiran Mani^{‡a}, Revati Kadolkar^{‡a,b}, Tithi Prajapati^{a,d}, Preety Ahuja^a, Mesha Shajahan^{a,d}, JungHun Lee^{a,b}, Michael Tolosa^a, Mary McWilliams^c, Claire Welty^{b,c}, Douglas D. Frey^{a,b}, Venkatesh Srinivasan^a, Sanjeev Kumar Ujjain^{*a}, Govind Rao^{*a,b}

^a Center for Advanced Sensor Technology, Department of Chemical, Biochemical and Environmental Engineering, University of Maryland Baltimore County, Baltimore, MD 21250, USA

^b Department of Chemical, Biochemical and Environmental Engineering, University of Maryland Baltimore County, Baltimore, MD 21250, USA

^c Center for Urban Environmental Research and Education, University of Maryland Baltimore County, Baltimore, MD 21250, USA

^d Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD, 21250, USA

Supplementary Information

Materials and Methods:

Potassium Ferricyanide (K₃Fe(CN)₆, 99%), Potassium Chloride (KCl, 99%), Sodium Chloride (NaCl, 99%), Sodium Sulphate (Na₂SO₄, 99%), Sulfuric Acid (H₂SO₄, 99%), Phosphoric Acid (H₃PO₄, 99%), Polyaniline (emeraldine salt) composite (20 wt.% polyaniline on carbon black), aniline, ammonium persulphate, isopropyl alcohol (IPA) and ethanol were procured from Sigma - Aldrich. Nitrilotriacetic Acid (NTA) was obtained from Thermo Scientific, Nickel Chloride (NiCl₂.6H₂O) from CAROLINA, Phosphate Buffer Saline (PBS, 99%) from Merck, and 200 proof Absolute Ethanol from Fisher bioreagents. 1000 ppm Nitrate standard (NO₃⁻, 99%) was purchased from RICCA and was diluted to make nitrate solutions for analysis. All purchased chemicals were of analytical grade and used without further processing. All solutions for analysis were prepared in high performance liquid chromatography (HPLC) grade deionized water (DI).

An Airbrush Trigger Gun with 0.3 mm air hose (Outeo) was used for coating working electrodes. For the 75 mL macro-scale three-electrode cell setup, glassy carbon (r=1.5 mm) working electrode, silver-silver chloride (Ag-AgCl) reference electrode (stored in 3 M KCl) and platinum wire counter electrode were all obtained from CH Instruments (CHI) in Austin, Texas. High purity Minseal 2010-B industrial grade plain graphite sheet purchased from Sealwiz was used to make working electrodes for the microfluidic cassette. Screen-printed electrodes were procured from CAULYS comprising of Ag-AgCl reference electrode with a conducting carbon counter electrode. These electrodes measure 12 x 34 mm in size and have a thickness of 0.3 mm which is ideal for the inner channel of our microfluidic cassette, developed at the Center for Advanced Sensor technology (CAST).

Clarex thin polymethyl methacrylate (PMMA) sheets of 0.5 mm and 1 mm were purchased from Astra products for making the microfluidic cassette. CO₂ laser engraver and cutter from Universal Laser Systems was used for fabrication of the cassette. A domestic microwave oven was purchased from Emerson (MW8987W) with a power rating of 1.35 kW and an operating frequency of 2.45 GHz. pH and specific conductance of the field water samples were measured with a Thermo Scientific Orion Star A329 portable meter.

Synthesis of Ni@Pani/C Sensing Material:

A commercial composite of Polyaniline emeraldine salt with 20 wt% polyaniline on carbon black (designated as Pani/C(20)) was used for synthesizing 65 wt% polyaniline on carbon black (designated as Pani/C (65)) through an in-situ chemical polymerization method with varying weight ratios of aniline. In each synthesis set, purified aniline was introduced into a 100 mg suspension of Pani/C(20) and subjected to sonication for approximately 3 hours.

The solution was then stirred in an ice bath at a temperature of 0–4°C. Maintaining that temperature, a pre-cooled solution of APS (ammonium persulfate) with a molar ratio of 1:4 aniline to APS was added dropwise into the above solution under vigorous stirring conditions overnight. The polymerized product denoted as Pani/C(65) was collected after centrifugation, washed with water and alcohol, and dried under vacuum. Subsequently, NiCl₂ and NTA were separately dissolved in alcohol and mixed in a molar ratio of 1:1 with a fixed content of Pani/C(65) under constant stirring over 5 hours, as described previously by Lori et al.^{S1} The resulting product, Ni@Pani/C, was collected through centrifugation, washed with alcohol, and dried under vacuum overnight.

Physicochemical Characterization of Ni@Pani/C:

Powdered samples were analyzed with Hitachi HT7800 120 kV microscope to conduct transmission electron microscopy (TEM). FEI Nova NanoSEM 450 field-emission scanning electron microscope, also equipped with EDS and EBSD capabilities (Oxford Aztec) was used for studying morphology and elemental analysis (Energy Dispersive X-ray analysis) of Ni@Pani/C. Powdered samples were analyzed with FESEM on a conductive carbon tape.



Figure S1. (A) Scanning electron microscopy (SEM) micrograph of Ni@Pani/C for surface morphology analysis; (B) Energy dispersive spectroscopy (EDX) images exhibiting elemental mapping of carbon, oxygen, nitrogen and nickel; (C) Transmission electron microscopy (TEM) image of Ni@Pani/C.

Fabrication of the Sensing Electrode:

For fabrication of the sensing electrodes, 5 mg of sample was sonicated with 10 wt% PVA in isopropyl alcohol for 30 minutes. The resulting suspension was deposited on a polished glassy carbon surface, followed by thermal treatment at 50°C for 1 hour. This was used as a working electrode in the three-electrode electrochemical setup. For microfluidic sensor, the fabrication process for working electrodes followed a similar protocol, as previously reported. ⁵² Initially, 10 mg of sample was sonicated with 10 wt% PVA binder in isopropyl alcohol for 30 minutes. The obtained suspension was spray-coated on 1 cm² graphite plates. Drying in an oven at 50°C for 1 hour ensured a secure adhesion of the active material on the graphite plates. To determine the amount of active material loaded onto each electrode, the plates were weighed before and after the spraying and drying processes respectively. Loading of the active material was found to be in a range of 1.1-1.5 mg/cm².



Figure S2. A conventional three-electrode system (75 mL) used for optimization of electrochemical sensing of nitrate using Ni@Pani/C nanocomposite material.

Fabrication of the Microfluidic Cassette:

CorelDraw was used to design 2D drawings of multiple layers of the microfluidic cassette. These designs were printed on PMMA sheets using a >9.3 µm CO₂ laser engraver with 10–75 watt power and 2.0 lens module.⁵³⁻⁵⁴ Laser settings in terms of power and speed were altered for cutting, engraving, and avoiding burned patches on PMMA. For optimization studies, 0.5-2 mm PMMA sheets were used to form layers of the microfluidic cassette. They were cleaned with 100% absolute ethanol (200 proof) and lint free tissues. After cleaning, the layers were aligned together by spraying ethanol between them, and further loaded into a custom-designed microwave-safe vise to hold the layers together under immense pressure.⁵⁵ This vise was placed in a domestic microwave for 2 minutes to ensure heating and binding of the layers into a cassette.

The fabricated microfluidic cassette is very sturdy and holds a volume of 500 μ L of electrolyte between the two slots for the insertion of the electrodes. The sockets or wells are designed in such a way that after the insertion of the coated working electrode and screen-printed electrode, their active surfaces face each other with the bulk of the solution between them in a vertical well. The cassette is made of PMMA material since it is durable, chemically inert to various buffers and pH conditions, and stable across a wide temperature range, thereby ensuring robustness in performance without degradation of the cassette.

Electrochemical Experiments for Sensing:

Initial optimization of the electrochemical experiments was performed in a conventional three-electrode cell with a CHI760E / CHI604D workstation. Prior to performing experiments, nitrogen (N₂) gas was purged for 30 minutes to get rid of the interferences caused by dissolved oxygen (DO). The electrical circuit entailed an electron flow from the reducing Pt counter electrode (anode half-cell) to the oxidizing Ni@Pani/C sensing electrode (cathode half-cell). On application of an electromotive force (emf) or cell potential (Ecell), the signal generated due to electron exchange at the electrode/electrolyte interface was analyzed with reference to the AgCl reference electrode. Electrochemical experiments such as cyclic voltammetry (CV) studies were carried out reversibly between -0.8 V to 0.8 V over 5 mV/s to 200 mV/s cycles to understand the oxidation state and pseudocapacitive properties of Ni@PaniC.⁵⁶

To study the ionic and electronic conduction across the electrolyte/electrode surface, electrochemical impedance studies (EIS) were performed at a voltage of -0.45 V and frequency range of 1 MHz – 0.1 Hz. To understand the effect of diffusion, Warburg impedance coefficient was obtained by simulating Nyquist plots in an EIS spectrum analyzer (EISSA).

Charge transfer resistance (R_{CT}) and corresponding solution resistance (R_S) values were fitted to an equivalent Randles Cell circuit in EISSA where a double layer Capacitance was in parallel, and together they were in series with R_S and Warburg impedance. Powell algorithm, 300 iterations, $n1 \approx 1$ and $P1 = 1e^{-5}$ to $1e^{-7}$ parameters were chosen to get a good fit. Various electrolyte compositions were tested, amongst which 0.1 M KCl + 50 mM PBS + 100 mM K₃Fe(CN)₆ was chosen for the micro-fluidic electrochemical studies. Sensing behavior was also assessed by adjusting the pH of the electrolytes at 4.5, 6 and 7 by using 0.1 wt% (V/V) phosphoric acid. Nitrate sensing was performed through successive addition of aliquots in the 500 µL microfluidic setup.

A wide concentration range of 0.6-50 ppm nitrate in electrolyte was chosen to maximize the ES \leftrightarrow PN redox transition of Pani upon application of redox potential, thereby resulting in a higher diffusion of nitrate ions within the electrode surface. ⁵⁶ An amperometry stabilization study was also performed to determine the sensor's response time. Selectivity of the electrode towards different ions (sulphate, acetate, bicarbonate, ammonium); stability in sensing at different temperatures (5–60°C); and consistency in sensing for a month (26 days) were also evaluated.



Figure S3. (A) CV plots of Ni@Pani/C at different scan rates (5 mV/s to 200 mV/s). Redox peak currents were monitored with an increase in scan rates; (B) Variation of redox peak currents versus scan rate (5 mV/s to 200 mV/s).



Figure S4. (A) CV plots of Ni@Pani/C nanocomposite at 50 mV/s and different pHs of the electrolyte: 4.5, 6.0 and 7.0; (B) CV plots of Ni@Pani/C nanocomposite at 50 mV/s in different electrolytes compositions.

Surface Water Testing:

To verify the applicability of our sensor in field water testing, samples containing nitrate were collected from streams in the Baltimore region, Maryland, USA. Sample sites were chosen based on previous work done by two of the co-authors (CW, MM) in the area.⁵⁷ Grab samples were collected from Gwynns Run, Dead Run, and Alexander Avenue tributaries of the Gwynns Falls (Table S1). 500 mL high density polyethylene (HDPE) Nalgene bottles were rinsed 3 times with stream water; submerged in the stream for sampling; capped underwater to avoid head space; and kept on ice during the sampling period. pH and specific conductance of the samples were measured with a Thermo Scientific Orion Star A329 portable meter. Samples were continuously kept on ice in the laboratory and filtered within the first 24-hours of the collection time using a 0.45 µm glass microfiber filter. In cases where samples could not be filtered immediately, they were frozen at -4°C and then analyzed within 24 hours after thawing. A commercially available standard optical and absorbance-based device, the Submersible UV Nitrate Analyser (SUNA V2) from Sea-Bird Scientific, was utilized for nitrate detection, as well as for generating a comparative analysis with the developed CAST microfluidic (CAST-MF) electrochemical sensor.



Figure S5. Observed and fitted data using EISSA circuit for analyzing the Nyquist plot. Inset shows equivalent circuit and components used for fitting: R1 corresponds to R_s, R2 to R_{CT}, CPE1 to the double layer Capacitance and Ws1 to Warburg coefficient. Applying a potential of -0.45V during EIS ensures that the system is actively undergoing the redox process, thereby enabling us to probe the interfacial electron transfer kinetics and electrochemical processes accurately.



Figure S6. (A) Nitrate (1 ppm) sensing response of the CAST-MF sensor with long term storage (26 days). (B) Performance of the PMMA cassette over a period of two months (67 days). Both experiments were conducted at room temperature and at pH 6.0.

Statistical Model:

An enhanced alternative to the traditional calibration curve was developed in MATLAB R2024a programming language (Mathworks, Massachusetts, U.S.A.) using multiple regression to obtain a 3D statistical model. Our objective was to understand the linearity of mathematical correlation between nitrate concentration (ppm) in water samples and their respective R_{CT} values extracted from the microfluidic sensor's electrochemical response.

A training data set for the model (Table S2) was prepared from dilutions of the collected samples mentioned in Table S1 to create a wide range of nitrate concentrations. Their $R_s(\Omega)$ and $R_{CT}(\Omega)$ values were obtained from EIS Nyquist plots of the microfluidic sensor's experiments. These R_s and R_{CT} values (Ω) were chosen as independent variables denoted by ' x_1 ' and ' x_2 ' in Equation 1. The concentration of nitrate (ppm) in field water samples measured by our microfluidic sensor was the dependent variable in Equation 1, denoted by 'Y'.

 $Y = \beta_0 + \beta_1 . x_1 + \beta_2 . x_2 + \beta_3 . \{f(x_1, x_2)\} \quad \dots \mbox{ Equation 1}$

First, the model computes regression coefficients denoted by ' β s' and then creates a design matrix denoted by 'X' to estimate these coefficients.^{S8-S9} As observed in the code^{S8-S9}, each row of the design matrix represents an existing datapoint from the data set and each column represents a feature. The design matrix consisted of four parts:

- Column vector of ones with the same number of rows as x₁.
- Column vector of x₁.
- Column vector of x₂.
- Column vector of a second order non-linear term as a function of x₁ and / or x₂.

Table S1: Details of the chosen surface water streams and locations near Baltimore in Maryland, U.S.A for sample collection.

Sample		Latitude	Longitude	Date Time		Specific	Nitrate -	
Number	Sampling Site			(GMT-05)	рН	Conductance	Nitrogen	Nitrate
						(µS/cm)	(mg N/L)	(ppm)
							(SUNA)	
Sample 1	Dead Run at Franklintown	39.311222°	-76.716639°	2023-04-27	8.0	988.7	1.06	4.43
				11:30:00				
Sample 2	Gwynns Run	39.276201°	-76.652614°	2023-04-27	7.76	580.1	1.98	8.27
				11:35:00				
Sample 3	Dead Run Unnamed	39.290121°	-76.746359°	2023-04-27	7.62	748.4	2.86	11.95
	Tributary at Alexander			12:17:00				
	Avenue							
Sample 4	Dead Run Unnamed	39.329612°	-76.748300°	2023-04-27	6.55	338.1	0.23	0.96
	Tributary near Rutherford Road			13:07:00				
Sample 5	Dead Run Unnamed	39.297585°	-76.731318°	2023-04-27	7.00	411.1	2.53	10.57
	Tributary near Kent Avenue			13:13:00				
Sample 6	Gwynns Run	39.276201°	-76.652614°	2023-07-20	7.87	522.6	1.87	7.82
				13:28:00				
Sample 7	Dead Run Unnamed	39.290121°	-76.746359°	2023-07-20	7.67	756.0	2.80	11.7
	Tributary at Alexander			14:05:00				
	Avenue							
Sample 8	Dead Run Unnamed	39.297585°	-76.731318°	2023-07-20	6.97	551.0	2.28	9.53
	Tributary near Kent Avenue			14:25:00				
						1		

Our goal was to screen various second order non-linear terms to identify the most effective one for enhancing the goodness of our fit. An element-wise product of x_1 and $x_2(x_1.x_2)$, their individual quadratic functions (x_1^2 and x_2^2), a square root function ($\sqrt[3]{x_1.x_2}$), a cube root function ($\sqrt[3]{x_1.x_2}$), and some mixed functions ($x_1\sqrt{x_2}$ and $x_2\sqrt{x_1}$) were tested separately to determine a single best non-linear term (Table S3, Figure S7 and Figure 8). Once identified, this non-linear term was used as the fourth column vector.

The 'regress' function then performs statistical regression to estimate coefficients using a least squares method. Essentially, it predicts a new set of nitrate values by drawing mathematical correlations between the existing data from the training set. These predicted nitrate concentration values were plotted against the test nitrate concentration values to obtain a good fit (Figure S7 and Figure 8).

Table S2: Training data set for the statistical model comprising of $R_S(\Omega)$, $R_{CT}(\Omega)$ and measured nitrate concentration (ppm) values for various stream water samples obtained from the CAST-MF microfluidic sensor. Samples mentioned in Table S1 were diluted to obtain a wide range of nitrate concentrations (ppm) for conditioning the model.

Nitrate Concentration measured with CAST-MF Microfluidic Sensor (ppm) after performing dilutions of the samples mentioned in Table S1	Sample Number	Solution Resistance (Rs) (Ω)	Charge Transfer Resistance (Rcr) (Ω)
0	Sample 4	4.41	3.28
0	Sample 7	17.41	3.54
0.6	Sample 6	17.65	3.55
1	Sample 6	17.79	6.23
1	Sample 5	17.1	8.65
3	Sample 8	17.49	10.93
4	Sample 2	16.67	11.66
5	Sample 2	15.52	15.53
5	Sample 3	20.03	23.72
6	Sample 3	17.93	28.77

It was observed that incorporation of the R_{CT} quadratic term (x_2^2) resulted in the best fitting of the model's predictions to the training data set (Figure S8) with the highest R^2 value of 0.9330 (Table S3). However, when this function was evaluated for the test data set, it showed a large absolute error in its predictions (±1.7 ppm) resulting in a case of overfitting. Upon examining all functions for the test data set containing randomly chosen stream water samples, the single best non-linear term was determined to be the cross-product function ($x_1.x_2$) with an RMSE=0.6558. The efficiency of the model increased by 20% when non-linear regression was compared against a traditional linear regression (RMSE=0.8246). The R^2 value also increased from 0.8578 (linear: left) to 0.9101 (non-linear: right) (Figure S7) proving to be acceptable for this first-generation sensor.^{S10-12} In the future studies, additional non-linear terms shall also be incorporated along with more data points to reduce training error. For this study, a cross-product term ($x_1.x_2$) was chosen as the fourth column in Equation 1 to prepare the 3D prediction plot as described below. The output vector denoted by 'b' in the code contains regression coefficients. Next, equally spaced vectors are created to generate a 2D grid with the 'meshgrid' function, followed by a vector of 'YFIT' grid points using the electrochemically measured nitrate concentrations. The meshgrid output and YFIT are further utilized to visualize a 3D mesh plot (surface plot), with colors varying from yellow to pink based on values specified by the Z-axis (Figure S9).

To assess the accuracy of our model's predictions, the test data set (Table 1) was prepared from four samples selected randomly from Table S1 and further diluted blindly to obtain unknown nitrate concentrations (ppm). Their extrapolated nitrate concentration values were compared against the experimentally measured values from the CAST-MF sensor. Two of four such extrapolations have been highlighted in Figure S10. A maximum standard deviation (SD) of ±0.6 ppm in absolute terms was observed in the model's predictions (Table 1) which can be further reduced by analyzing a greater number of samples obtained from each of the water streams being tested. Overall, we have shown the development of a low-cost and portable sensor that can eventually be used outside of the laboratory by untrained operators due to the simplicity of our predictive model. Unlike the other complicated data analyses reported previously in nitrate sensors⁵¹⁰⁻¹², our model requires a single-point entry of input parameters and can provide the estimated nitrate concentrations in water samples to the operator within seconds.

Table S3: Comparative assessment of the model's performance for regression analysis between predicted (model output)versus test nitrate concentrations (Table S1) using various second order non-linear terms.

Equation	Incorporated Second Order Term {f(x1, x2)}		RMSE Value	R ² Value	Goodness of Fit: Based on R ² Value	Absolute Error in Predicted Nitrate Concentrations	Quality of Fitting for the Test Data Set	Overall Precision Rank
$Y = \beta_0 + \beta_{1.x_1} + \beta_{2.x_2}$	No Second Terr (Linear Reg	l Order n ression)	0.8246	0.8578	8	±0.7 ppm	Fits Well	3
$Y = \beta_0 + \beta_{1,x_1} + \beta_{2,x_2} + \beta_{3} \cdot \{f(x_1, x_2)\}$	Cross Product	X1.X2	0.6558	0.9101	4	±0.6 ppm	Fits Well	1
	Quadratic	X1 ²	0.6514	0.9113	3	±>10 ppm	Unac- ceptable	_
		x2 ²	0.5661	0.9330	1	±1.7 ppm	Overfits	2
	Square Root	$\sqrt{x_1.x_2}$	0.6558	0.8969	5	±>10 ppm	Unac- ceptable	_
	Cube Root	$\sqrt[3]{x_1^2.x_2^2}$	0.8018	0.8656	6	±>10 ppm	Unac- ceptable	-
	Mixed	$x_1\sqrt{x_2}$	0.8150	0.8611	7	±>10 ppm	Unac- ceptable	_
		$x_2\sqrt{x_1}$	0.6502	0.9116	2	Negative Values	Unac- ceptable	-



Figure S7: Plot of predicted nitrate concentrations by the statistical model (Y axis) versus test nitrate concentrations from experimental data in Table S2 (X axis). Linear regression analysis without a second order term (left). Multiple regression analysis with a cross-product term of $R_s(x_1)$ and $R_{CT}(x_2)$ values for reducing error in predictions (right).



Figure S8: Plots of multiple regression analysis between predicted nitrate concentrations by the statistical model (Y axis) versus test nitrate concentrations from experimental data in Table S2 (X axis). Various second order terms were tested: (A and B) individual quadratic function $(x_1^2 \text{ and } x_2^2)$, (C) square root function $(\sqrt[3]{x_1} \cdot x_2)$, (D) cube root function $(\sqrt[3]{x_1} \cdot x_2^2)$, and (E and F) mixed functions $(x_1\sqrt{x_2} \text{ and } x_2\sqrt{x_1})$.



Figure S9: Predicted nitrate concentrations showed a trend of strong dependence on R_{CT} values and weak dependence on R_S values.



Figure S10: Extrapolation of the predicted nitrate concentrations (Z) from experimental $R_s(X)$ and $R_{CT}(Y)$ values have been highlighted for two of the four test samples obtained from surface water steams in Maryland, U.S.A.

References:

[S1] Lori, J. A.; Morrin, A.; Killard, A. J.; Smyth, M. R. Development and Characterization of Nickel-NTA-Polyaniline Modified Electrodes. *Electroanalysis* **2006**, *18*, 1, 77-81.

[S2] Ujjain, S. K.; Ahuja, P.; Sharma, R. K. Facile Preparation of Graphene Nanoribbon/Cobalt Coordination Polymer Nanohybrid for Non-enzymatic H₂O₂ Sensing by Dual Transduction: Electrochemical and Fluorescence. J. Mater. Chem. B **2015**, *3*, 7614–7622.

[S3] Andar, A.; Hasan, M.-S.; Srinivasan, V.; Al-Adhami, M.; Gutierrez, E.; Burgenson, D.; Ge, X.; Tolosa, L.; Kostov, Y.; Rao, G. Wood Micro-fluidics. *Anal. Chem.* **2019**, *91*, 11004–11012.

[S4] Hasan, M. S.; Borhani, S.; Ramamurthy, S. S.; Andar, A.; Ge, X.; Choa, F.-S.; Kostov, Y.; Rao, G. Microwave induced thermally assisted solvent-based bonding of biodegradable thermoplastics: an eco-friendly rapid approach for fabrication of microfluidic devices and analyte detection. *Sci. Rep.* **2022**, *12*, 16075.

[S5] Wang, N.; Wang, R. Q.; Zhu, Y. A Novel Ion Chromatography Cycling-Column-Switching System for the Determination of Low-Level Chlorate and Nitrite in High Salt Matrices. J. Hazard. Mater. **2012**, 235–236, 123–127.

[S6] Kosa, S. A. M.; Khan, A. N.; Ahmed, S.; Aslam, M.; Bawazir, W. A.; Hameed, A.; Soomro, M. T. Strategic Electrochemical Determina-tion of Nitrate over Polyaniline/Multi-Walled Carbon Nanotubes-Gum Arabic Architecture. *Nanomaterials* **2022**, *12*, 3542.

[S7] Welty, C.; Moore, J.; Bain, D. J.; Talebpour, M.; Kemper, J. T.; Groffman, P. M.; Duncan, J. M. Spatial heterogeneity and temporal stability of baseflow stream chemistry in an urban watershed. *Water Resour. Res.* **2023**, *59*, e2021WR031804.

[S8] Chatterjee, S.; Hadi, A. S. Influential Observations, High Leverage Points, and Outliers in Linear Regression. *Statistical Science*. **1986**, *1* (3), 379–393.

[S9] https://www.mathworks.com : https://www.mathworks.com/help/stats/regress.html?s_tid=doc_ta

[S10] Vakilian, K. A.; Massah, J. A Portable Nitrate Biosensing Device Using Electrochemistry and Spectroscopy. *IEEE Sens. J.* 2018, 18 (8), 3080-3089.

[S11] Vakilian, K. A.; Massah, J. A Portable Nitrate Biosensing Device Using Electrochemistry and Spectroscopy. *IEEE Sens. J.* 2018, 18 (8), 3080-3089.

[S12] Massah, J.; Vakilian, K. A. An Intelligent Portable Biosensor for Fast and Accurate Nitrate Determination Using Cyclic Voltammetry. *Biosyst. Eng.* **2019**, *177*, 49-58.

```
% Objective: To study the individual effects of solution resistance (Rs) and % charge transfer resistance (Rct) on nitrate concentration.
```

```
% Solution Resistance (ohm)
x1 = [4.41; 17.41; 17.65; 17.79; 17.1; 17.49; 16.67; 15.52; 20.03; 17.93];
```

```
% Charge Transfer Resistance (ohm)
x2 = [3.28; 3.54; 3.55; 6.23; 8.65; 10.93; 11.66; 15.53; 23.72; 28.77];
```

```
% Nitrate Concentration (ppm) (Y = aX1 + bX2 + c)
y = [0; 0; 0.6; 1; 1; 3; 4; 5; 5; 6];
```

```
% Non-Linear Regression: Identify the best second-order term that helps to reduce error in predictions.
```

```
% Create design matrix: X
X_cross = [ones(size(x1)) x1 x2 x1.*x2];
```

```
% Get offsets: coefficients in b
[b_cross,bint, r, rint, stats] = regress(y,X_cross);
```

```
% Get predicted nitrate concentration values: Y_cross
Y_cross = X_cross*b_cross;
```

```
% Plot residuals of nitrate concentration: Predicted v/s True
% Get min and max values from both arrays: y (true) and Y_cross (predicted)
[min_y, max_y] = bounds(y);
[min_Y_cross, max_Y_cross] = bounds(Y_cross);
[min_plot, max_plot] = bounds([min_y, max_y, min_Y_cross, max_Y_cross]);
```

```
% Plot Lines: best fit and zero-intercept regression
linearCoefficients = polyfit(y, Y_cross, 1);
y_line = polyval(linearCoefficients, y);
plot([min_plot:max_plot], [min_plot:max_plot], "LineStyle","--", "LineWidth",3,
"Color","black");
hold on;
plot(y, y_line, "LineStyle","-", "LineWidth",2, "Color","#2241f2");
hold on;
```

```
% Scatter Plot
scatter(y, Y_cross, 150, "MarkerEdgeColor", "#14abc9", "MarkerFaceColor", "#41575c");
hold off;
box on
grid on
ax = gca;
set(gca, 'gridlinestyle', '--', "gridlinewidth", 0.2)
ax.FontSize = 18;
ax.FontWeight = 'bold';
ax.LineWidth = 1.5;
```

```
% Set Axis Ticks, Limits and Labels
xticks([min_plot:max_plot])
yticks([min_plot:max_plot])
ylim([min_plot max_plot])
xlabel('Test Nitrate Concentration (ppm)', 'FontSize', 14, 'FontWeight', 'bold')
ylabel('Predicted Nitrate Concentration (ppm)', 'FontSize', 15, 'FontWeight',
'bold')
axis square
% Set Title and Legend Properties
title({['\fontsize{16}','Regression Analysis with a Cross Product Function'];
['\fontsize{15}','(R=0.95, R^{2}=0.91)']},'FontWeight','bold');
clear legend
legend('x=y','Fit','Data', 'Location','southeast');
```



Regression Analysis with a Cross Product Function

```
% RMSE Calculation and Statistics
RMSE = rmse(Y_cross,y)
```

```
RMSE = 0.6558
```

```
RMSE_weighted = rmse(Y_cross,y,Weights=y)
```

RMSE_weighted = 0.7786

disp("R^2 value with Cross Product Term (x1.*x2):")

R^2 value with Cross Product Term (x1.*x2):

<pre>stats(1)</pre>			
ans = 0.9101			
<pre>disp(' ');</pre>			

disp(' ');

```
% Various second-order functions of x1 and x2 were tested such as cross product,
% quadratic, square root, cube root and mixed functions. The cross product function
% provided a strong correlation between the predicted and tested values
% (RMSE=0.65, R^2=0.91), as well as a good fit for the test data set.
% Hence, chosen to prepare the 3D model as below.
% Create design matrix: X
X = [ones(size(x1)) x1 x2 (x1.*x2)];
% Get offsets: coefficients in b
[b, bint, r, rint, stats1] = regress(y,X);
% Create equal-spaced data points between x1 and x2 for mesh plot:
x1fit = min(x1):0.025:max(x1);
x2fit = min(x2):0.07:max(x2);
% Calculate grid points for mesh: 2D
[X1FIT,X2FIT] = meshgrid(x1fit,x2fit);
% Calculate grid points for mesh: 3D
VFIT = b(1) + b(2) * X1FIT + b(3) * X2FIT + b(4) * X1FIT. * X2FIT;
% Avoid mixing of figures:
hold off
% Create the mesh plot and provide color gradient
mesh(X1FIT,X2FIT,YFIT);
colorbar
colormap spring
oldcolormap = colormap;
colormap( flipud(oldcolormap) );
% Pause graph display
hold on
% Create a scatter plot of input data points: x1 (Rs) and x2 (Rct)
```

```
scatter3(x1, x2, y, 120, 'black', 'filled',"LineWidth", 4);
% Beautify plot
box on
ax = gca;
set(gca,'FontSize', 30, 'FontWeight', 'bold')
ax.FontSize = 14;
ax.LineWidth = 1.2;
ax.GridLineWidth = 1;
ax.ZGrid = 'on';
ax.XGrid = 'off';
ax.YGrid = 'off';
% Set axis labels
xlabel('R_{S} (Ω)', 'FontSize', 15, 'FontWeight', 'bold')
ylabel('R_{CT} (Ω)', 'FontSize', 15, 'FontWeight', 'bold')
zlabel('Nitrate Concentration (ppm)', 'FontSize', 14, 'FontWeight', 'bold')
% Generate background grid
grid on
% Set plot size for viewing
view(50,10)
% view([120.7 12.0])
% view([-0.3 90.0])
% Display final plot
hold off
```



% Display statistics: disp("R^2 value:")

R^2 value:

stats1(1)

ans = 0.9101

% References:

- % Mathworks (https://www.mathworks.com/)
- % Multiple Linear Regression: https://www.mathworks.com/help/stats/regress.html?
- s_tid=doc_ta