

Analysis of Macronutrients in Soil Using Impedimetric Multisensor Arrays

Maria Luisa Braunger, Mario Popolin Neto, Dmitry Kirsanov, Igor Fier, Lucas R. Amaral, Flavio M. Shimizu, Daniel S. Correa, Fernando V. Paulovich, Andrey Legin, Osvaldo N. Oliveira Jr., and Antonio Riul Jr.*



Cite This: *ACS Omega* 2024, 9, 33949–33958



Read Online

ACCESS |



Metrics & More

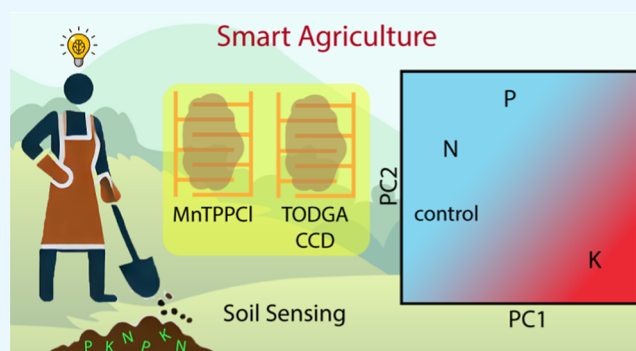


Article Recommendations



Supporting Information

ABSTRACT: The need to increase food production to address the world population growth can only be fulfilled with precision agriculture strategies to increase crop yield with minimal expansion of the cultivated area. One example is site-specific fertilization based on accurate monitoring of soil nutrient levels, which can be made more cost-effective using sensors. This study developed an impedimetric multisensor array using ion-selective membranes to analyze soil samples enriched with macronutrients (N, P, and K), which is compared with another array based on layer-by-layer films. The results obtained from both devices are analyzed with multidimensional projection techniques and machine learning methods, where a decision tree model algorithm chooses the calibrations (best frequencies and sensors). The multicalibration space method indicates that both devices effectively distinguished all soil samples tested, with the ion-selective membrane setup presenting a higher sensitivity to K content. These findings pave the way for more environmentally friendly and efficient agricultural practices, facilitating the mapping of cropping areas for precise fertilizer application and optimized crop yield.



INTRODUCTION

The projected global population of 9.8 billion by 2050 necessitates a 70% increase in food production.¹ To achieve this sustainably, innovative strategies are required to enhance agricultural yield without expanding arable land.^{2,3} Precision agriculture, a key aspect of smart agriculture, emerges as a crucial technique that enables increased productivity through the site-specific application of fertilizers, guided by precise soil nutrient monitoring.^{4,5} Traditionally, soil fertility is assessed by collecting multiple samples from fields for chemical analysis in laboratories, a process that can be expensive and often leads to insufficient sampling, thus failing to accurately represent the spatial variability of soil nutrients.⁶ Recent advancements in soil sensing technologies underscore a shift toward noninvasive and real-time monitoring methods, crucial for both precision and smart agriculture.^{7–12}

Recent advancements in smart agriculture emphasize the integration of advanced sensor technologies and the use of machine learning algorithms to enhance the precision and efficiency of farming practices.^{7,8} Wearable sensors developed for applications like leaf moisture monitoring employ capacitive measurements to provide real-time data on plant health, demonstrating significant advances in noninvasive agricultural monitoring.^{9,10} The integration of internet of things and artificial intelligence technologies in agriculture not

only improves productivity but also contributes to sustainable farming practices by minimizing resource wastage and environmental impact. Moreover, the deployment of multi-sensor arrays has proven their versatility across various agricultural applications, including the precise detection of soil nutrients and environmental contaminants, boosting a more rational and rapid decision-making action in the future.^{11–13}

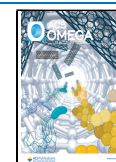
Multisensor arrays, particularly effective in analyzing complex fluids, transform raw data into recognizable patterns through advanced computational and statistical methods.^{14–16} These arrays have proven their utility in various sectors, including food and beverage quality control,^{17–20} and the detection of pollutants and biomedical markers.^{13,21–23} Specifically, potentiometric and impedimetric multisensor arrays have shown great promise in soil analysis.^{24–28} For instance, an impedimetric multisensor array based on layer-by-layer (LbL) coatings showed good distinction of soil samples

Received: May 10, 2024

Revised: July 8, 2024

Accepted: July 17, 2024

Published: July 25, 2024



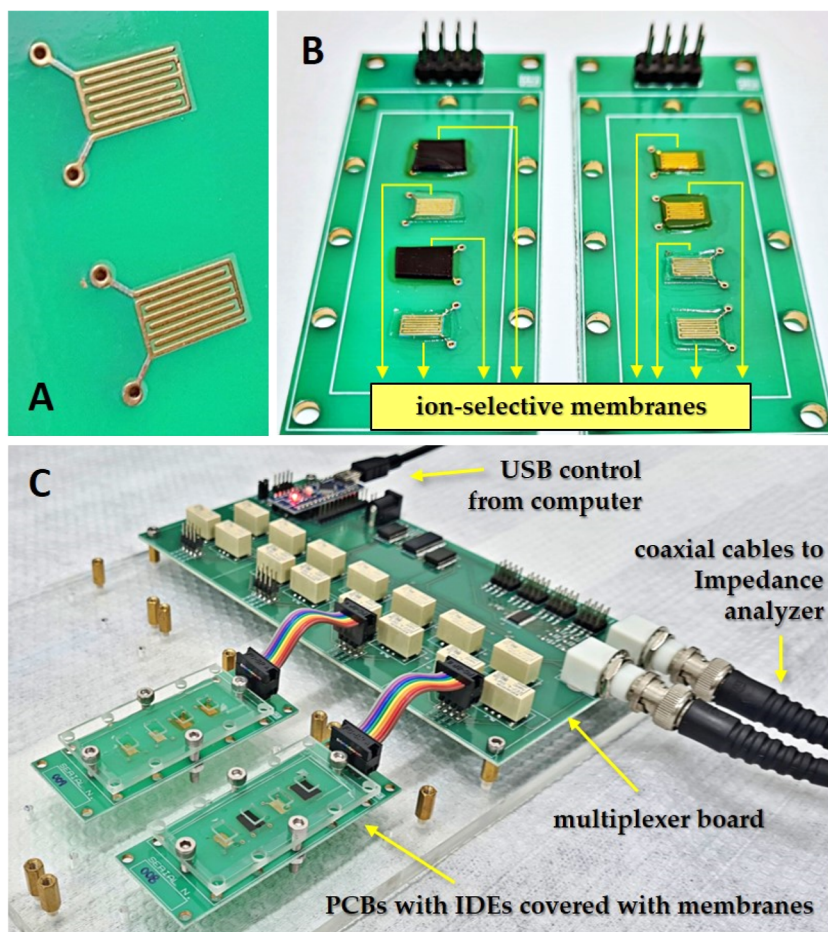


Figure 1. Impedimetric multisensor array: (a) close-up of the bare IDEs before the adhesion of ion-selective membranes. (b) Two PCBs with IDEs covered by eight distinct ion-selective membranes, allowing for simultaneous detection of multiple analytes. (c) The multisensor array setup consists of sensing units connected to a digitally controlled analog multiplexing circuit, enabling precise and efficient impedance measurements. The IDEs have four pairs of digits having 5 mm length, 0.2 mm width, and 0.2 mm spacing.

enriched with nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg) and sulfur (S), all of them easily dispersed in water with raw impedance data processed by distinct multidimensional projection techniques for improved information visualization.²⁶ In a similar context, Khaydukova et al. applied a potentiometric multisensor array based on cross-sensitive plasticized polymeric membranes to determine the NPK content in aqueous soil dispersions, with the main advantage of a fast-measuring procedure (only 8 min for data acquisition).²⁵ These nutrients, crucial for plant growth, are routinely applied each cropping season.

In this study, we explored the capabilities of novel impedimetric multisensor arrays for distinguishing soil samples enriched with essential macronutrients. Using ion-selective membranes and LbL films, we demonstrated the arrays' ability to effectively classify these samples through simple water dispersion and advanced statistical techniques. The machine learning analysis, particularly decision tree (DT) models, was employed to enhance the classification accuracy, indicating that both devices effectively distinguished all soil samples tested, with the ion-selective membrane setup presenting a higher sensitivity to K content. The findings indicate the potential for these sensor arrays to contribute significantly to in situ soil analysis, paving the way for more precise and affordable nutrient monitoring in support of precision agriculture.

MATERIALS AND METHODS

The impedimetric multisensor array used in this study is not a commercial product, but a custom-built setup. The core of the device is based on four collinear, gold-plated interdigitated electrodes (IDEs), embedded on a printed circuit board, fabricated by TEC-CI Circuitos Impressos (São Paulo-SP, Brazil) according to the authors' instruction.²⁹ Figure 1 displays photos of the bare IDEs, the multisensor array, and the setup for impedance measurements. The multisensor array comprises eight sensing units made with ion-selective membranes, five anionic and three cationic membranes, labeled here from S1 to S8, with full description displayed in Table S1 of the Supporting Information. The membranes consist of plasticized polymeric sensor membranes containing poly(vinyl chloride) (PVC) as a polymer, dioctyl sebacate or *o*-nitrophenyloctyl ether (NPOE) as a plasticizer, and various membrane-active compounds listed in Table S1.^{30,31} The membranes were synthesized using a conventional procedure, where the weighed amounts of all components were dissolved in freshly distilled tetrahydrofuran (THF), and the resultant solutions were poured into flat-bottom Teflon beakers to evaporate the solvent. The resulting membranes were cut in $\sim(6 \times 4)$ mm² and adhered to the IDEs using THF.

A multisensor array having LbL films³² as sensing units is used for comparison, with a picture of the device shown in

Figure S1. Full details on the fabrication of this microfluidic device are given elsewhere.²⁹ Briefly, copper phthalocyanine-3,4',4'',4'''-tetrasulfonic acid tetrasodium salt (CuTsPc), montmorillonite clay (MMt-K), poly(3,4-ethylenedioxythiophene)–poly(styrenesulfonate) (PEDOT:PSS), and poly(diallyldimethylammonium chloride) solution (PDDA) were purchased from Sigma-Aldrich and used as received as cationic and anionic layers in LbL assemblies. These materials are simply dispersed in deionized water and the LbL films are fabricated by sequentially alternating their aqueous dispersions onto the IDEs to form 10 bilayers of PDDA/CuTsPc, PDDA/MMt-K, and PDDA/PEDOT:PSS. Three electrodes modified with LbL films and one bare IDE comprise four distinct sensing units for the setup configuration.

Impedance measurements are conducted in a Solartron 1260A impedance/gain-phase analyzer, applying a 25 mV AC voltage over 1–10⁶ Hz frequency range. To facilitate the routing of analog signals between the sensing units and impedance analyzer, a new homemade multiplexer was developed based on previous ones.^{29,33} The software-controlled multiplexer utilized a cascading relay network to enable fully automated and customizable experimental sequences of measurements. In a separate electronic assembly, a digitally controlled analog multiplexing system was designed and built to independently route the signal of up to 16 IDEs (4 PCBs containing 4 IDEs each) to a pair of coaxial BNC connectors. This unit features an inexpensive Arduino Nano microcontroller module, communicating via USB to a PC unit. The Arduino powers an array of small-signal relays through a Darlington driver integrated circuit (ULN2003) from Texas Instruments.

Before employing the sensor arrays for soil analysis, which comprise much more complex samples, we put both in a proof-of-concept test, aiming to distinguish electrolytes and non-electrolytes solutions representing the five basic tastes relevant to human gustatory perception: sweet, salty, sour, bitter, and umami. To this end, we analyze 1 mM aqueous solutions of sucrose (C₁₂H₂₂O₁₁), sodium chloride (NaCl), hydrochloric acid (HCl), caffeine (C₈H₁₀N₄O₂), and L-glutamic acid (C₅H₉NO₄). This molar concentration is close to, or even lower than, some human taste thresholds,^{34–37} with an exception for bitterness that can be perceived at μ M concentrations by the biological counterpart.³⁵ Both impedimetric multisensor arrays successfully respond to all basic tastes, encouraging us to evaluate their performance in more complex analytes, such as soil samples.

Soil with low natural fertility from Campinas (SP/Brazil), referred here as the “control”, is divided into distinct containers and individually enriched with N, P, and K, using NH₄NO₃, NH₄H₂PO₄, and KCl fertilizers, respectively, as described elsewhere.^{38,39} The soil nutrients have been previously analyzed with conventional wet-chemical analysis, with the results presented in Table S2.⁴⁰ Wet-chemical analyses encompass the estimation of K and P availability using the ion-exchange resin extraction method, and total N is estimated through the Kjeldahl method, as previously described.⁴¹ Note that the soil samples naturally contain N, P and K nutrients in addition to others not mentioned here. Once again, control samples are those not enriched, and named “N”, “P”, and “K” refer to samples enriched with these nutrients, respectively.

The following procedure is adopted for impedance measurements. Initially, control, and N, P, and K enriched samples are

dispersed in water at 100 mg/mL. They are sequentially sonicated 30 min in an ultrasound bath, and then left resting for 1 h before analysis. The resulting suspensions are diluted three times to yield four portions of each soil sample at 1, 10, 50, and 100 mg/mL. We need 1 mL for three independent sets of measurements for each sample using the LbL-based array, and 6 mL for triplicates using the ion-selective membranes array. The smaller volume required by the LbL multisensor array is due to the microfluidic integration enabled in that setup. To assess possible cross-contamination before evaluating a new soil sample, the response of a control analyte (distilled water) is evaluated after each soil measurement. The acquisition time for each analyte is \sim 5 min for the LbL-based array (4 sensing units) and \sim 10 min for the ion-selective membranes array (8 sensing units). The analytes are evaluated at 5 mL/h in the microfluidic device containing LbL films as sensing units, with the flow rate chosen in previous work, with the best discrimination reached without wasting unnecessary amounts of sample.²⁹ An example of the time expended in this analysis using the multiplexer, with 8 sensing units based on ion-selective membranes, an analyte was measured in triplicate and with three independent measurements in \sim 30 min. Consequently, the entire process for the triplicates in the 16 soil samples studied here (control, N, P and K dispersed in four distinct concentrations in water) can be completed in \sim 8 h. This is in stark contrast to individual devices without a multiplexing system, which would require more than a single working day to accomplish the same task with the same soil samples.^{26,28}

The impedance raw data presents high dimensionality ($n \geq 124$) for the multisensor arrays, so we underwent dimensionality reduction using Principal Component Analysis (PCA),⁴² which involved projecting the multidimensional data into a 2D Euclidean space for easier interpretation and visualization. To assess the quality of the resulting data discrimination, we applied the silhouette coefficient (SC)⁴⁴ computed as a value $s(i)$ for each object i that measures its similarity to its cluster (cohesion), when compared to other clusters (separation). We also calculated the SC⁴⁴ for the entire plot, which is the average of $s(i)$ for all objects in the data set. Positive values of SC are interpreted as strong (0.71–1.0), reasonable (0.51–0.70), weak (0.26–0.50), and no discrimination (≤ 0.25).⁴⁴ Based on the SC scores, we used the k -means algorithm⁴⁵ to identify the number (k) of nonoverlapping clusters in an unbiased manner. PCA and SC scores are complementary approaches in data analysis, with the first being a dimensionality reduction technique largely used to identify patterns and structures in the data, while SC provides a value indicating how well an object is allocated in a cluster (it assesses the quality of clustering). The raw impedance data are processed using the Orange Data Mining software to perform PCA, compute SC scores, and execute k -means clustering.⁴⁶

Classification models from machine learning methods combined with visual representations are employed for analyzing the data from the multisensor arrays. For each multisensor array a multidimensional calibration space (MCS)¹⁶ is produced by taking the impedance magnitude value (Ohms) in different frequencies. Thus, for distinguishing soil samples categorized into 4 classes (control, K, N, and P), one MCS is created for the multisensor array made with LbL films, and another for the ion-selective membrane-based array. DT models^{47–49} are used for creating such calibrations, using ExMatrix for interpretability.⁵⁰ A model selection experiment is

conducted via KFold Cross-Validation⁵¹ to choose the hyperparameter combinations that maximize DT models' performance. Then, the hyperparameters holding the highest average performance from the KFold Cross-Validation are applied to create the final DT model using all data. To obtain the deployed (final) DT model's average performance, a Nested KFold Cross-Validation is conducted.^{52,53} In this procedure, two KFold Cross-Validation loops (inner and outer) are executed. The inner loop assesses model hyperparameter tuning (model selection),^{48,49} whereas the outer loop involves model performance estimation.^{52,53} Optimistic (overestimation) and biased performance measures are typical issues on small data sets, and they can be mitigated using a Nested KFold Cross-Validation.^{52,53} The MCS provides interpretability, unveiling associations among the impedance magnitude values and classes (e.g., soil samples category), and the Nested KFold Cross-Validation method delivers a more truthful performance estimation.⁴³

RESULTS AND DISCUSSION

Both multisensor arrays can easily distinguish solutions responsible for relevant basic tastes (sweet, salty, sour, bitter, and umami) by processing the impedance raw data in PCA score plots. Some examples of the impedance spectra obtained are shown in Figure S2. Figure 2 shows the PCA plot for the

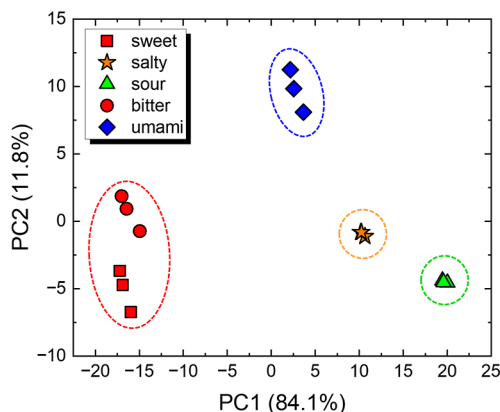


Figure 2. PCA score plots from impedance data in three independent sets of measurements evaluating basic tastes relevant to human gustative perception (sweet, salty, sour, bitter, and umami) at 1 mM obtained with the multisensor array based on ion-selective membranes.

multisensor array made with ion-selective membranes, while the PCA plot corresponding to the LbL-based array is shown in Figure S3 in the Supporting Information. The sum of the explained variance for the first two principal components yielded 97.0 and 95.9% of the total information for the data sets obtained with the devices made with LbL films and ion-selective membranes, respectively. The different shapes in the markers of the scatter plot correspond to the sort of sample, while the different colors indicate the clustering suggested by the *k*-means method considering the highest SC score. Even though one can visually distinguish all five taste samples with both setups, the *k*-means clustering indicates four groups maximizing the discrimination ($SC = 0.789$) for the ion-selective membrane-based device, with the algorithm combining sweetness (sucrose) and bitterness (caffeine) in a single cluster (red area). As the membranes are ion-selective, it is

anticipated that they perform well with electrolytes but not as effectively with nonelectrolytes. This expectation is confirmed by these results, where the PCA clusters caffeine and sucrose together, while distinctly separating the other tastes. In contrast, due to their nonion-specific nature, the LbL-based array demonstrates a superior ability to distinguish between all five taste samples, including sucrose and caffeine, reflected in higher SC scores and distinct clustering in the PCA plots (Figure S3). When five clusters are considered, the SC values are 0.823 and 0.722 for the arrays based on LbL films and ion-selective membranes, respectively. Despite the decrease in the SC score from 0.789 to 0.722 when five clusters are considered for the ion-selective membrane-based data set, this score remains above 0.7, indicating robust discrimination within the clusters.⁴⁴

The analysis of impedance data for soil samples facilitated the distinction between control, N, P, and K samples across all tested concentrations (1, 10, 50, and 100 mg/mL), as depicted in the PCA score plots. Figure 3 illustrates the PCA score plot

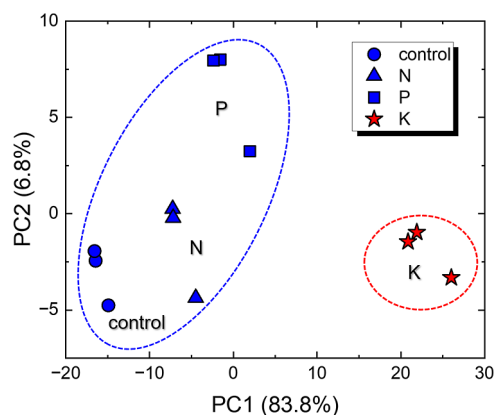


Figure 3. PCA score plots of the impedance data in three independent sets of measurements to evaluate soil samples individually enriched with N, P, and K. Data sets obtained from the evaluation of the soil samples dispersed in water at 100 mg/mL with the multisensor array made with ion-selective membranes. The PCA score plots for soil samples at 1, 10, and 50 mg/mL presented a similar pattern and are shown in the Supporting Information (Figure S4).

for the 100 mg/mL data set obtained with the ion-selective membrane array, with additional plots for the other concentrations and the LbL-based array presented in Figures S4 and S5, respectively. The PCA score plots for the ion-selective membrane array enabled a visual distinction between the control sample and those enriched with N, P, and K in Figure 3. However, *k*-means clustering suggests that maximal discrimination is achieved in two clusters: one combining the control with N and P samples (blue area), and another exclusively containing the K-enriched sample (red area). The results indicate that the ion-selective membranes exhibit particular sensitivity to potassium, likely due to its prevalent ionic state which enhances its mobility and solubility, thereby facilitating detection at the electrode/electrolyte interface. This sensitivity could be instrumental for the selective detection of potassium in soil samples, offering a significant step toward the precise quantification of this essential macronutrient.

In contrast, the PCA plots for the LbL-based device (Figure S5) show that all soil samples are visually distinguishable across aliquots of 1, 10, 50, and 100 mg/mL. This is expected,

Table 1. Number of Clusters Found With the *k*-Means Method, SC Scores, and Contributions from PC1 and PC2 in the Impedance Data for Soil Samples Dispersed in Water at 1, 10, 50, and 100 mg/mL, Obtained with the LbL-Based and Ion-Selective Membrane-Based Arrays

	LbL films				ion-selective membranes			
data sets (mg/mL)	1	10	50	100	1	10	50	100
number of clusters (<i>k</i>)	2	4	4	4	2	2	2	2
silhouette coefficient	0.585	0.694	0.810	0.841	0.401	0.505	0.634	0.624
PC1 (%)	72.3	88.5	88.0	85.7	66.1	79.6	83.8	83.8
PC2 (%)	14.7	7.3	8.0	7.8	13.0	11.2	7.9	6.8
PC1 + PC2 (%)	87.0	95.8	96.0	93.5	79.1	90.8	91.7	90.6

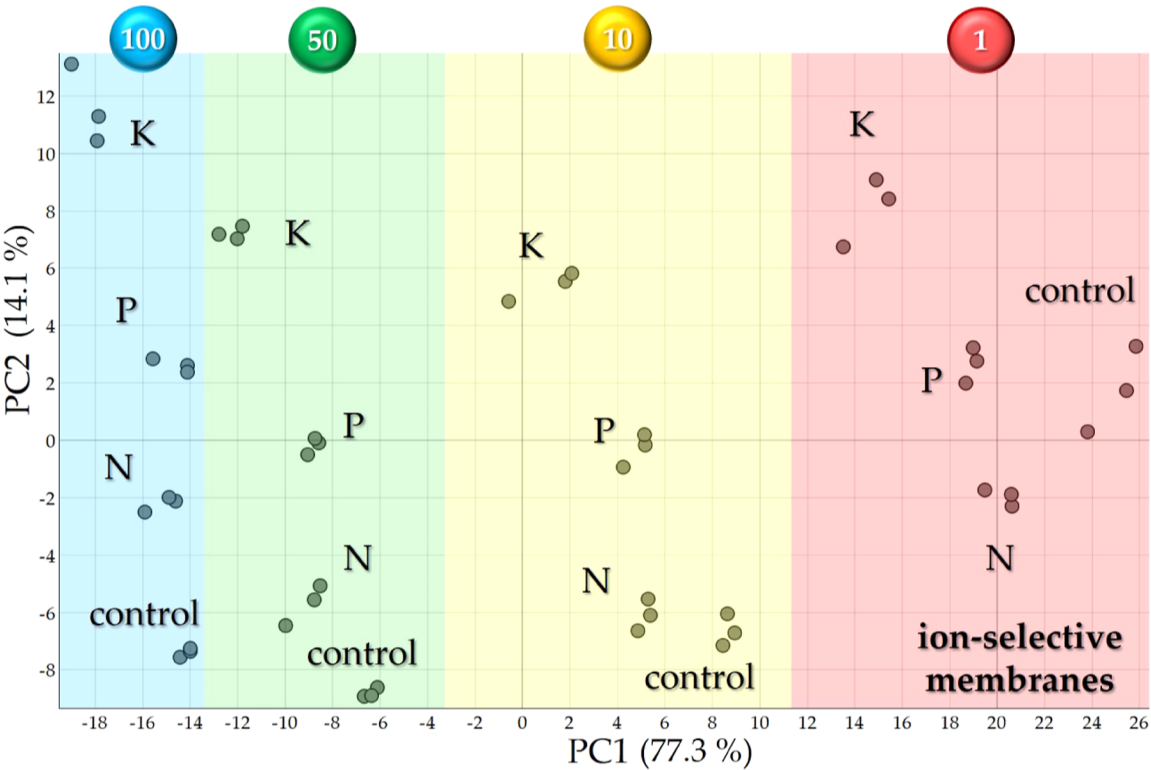


Figure 4. PCA score plots of the impedance data of control, N, P, and K soil samples dispersed in water at 1 (red), 10 (yellow), 50 (green), and 100 (blue) mg/mL obtained with the multisensor array made with ion-selective membranes.

as the LbL-based array, unlike the ion-selective membrane array, does not employ materials that are selectively sensitive to specific ions. Specifically, for the 1 mg/mL aliquot, *k*-means clustering and SC scores reveal that the control and P samples cluster together, while N and K form a separate cluster. This pattern, previously observed by Americo da Silva et al. using an LbL-based sensor array, underscores the differential sensitivity to N and K over P.²⁸ Due to the straightforward sample preparation method, involving only the dispersion of soil in water without any chemical pretreatments, phosphorus tends to remain bound within soil colloids as complex compounds. This results in its poor solubility and minimal contribution to the differentiation in PCA plots, which explains the proximity of P-enriched samples to the control.

The method of simple dispersion in water, while aiding rapid and noncomplex field measurements, does not fully address the detection of nitrogen and phosphorus in their varied and complex forms within the soil. Future improvements should incorporate more sophisticated extraction techniques to enhance detection capabilities for nitrogen and phosphorus, as well as other nutrients, recognizing that the current

approach may limit the accuracy and sensitivity particularly for nitrogen, given its presence in hydrolyzed and insoluble forms. The forthcoming enhancements to the extraction methods will be crucial to optimize the sensor arrays' performance and to ensure more reliable, sensitive measurements of soil nutrients.

Table 1 presents the results with the two multisensor arrays for the control, N, P, and K soil samples dispersed in water at 1, 10, 50, 100 mg/mL, including the number of clusters suggested by the *k*-means method, the SC values used to find the best clustering in each case, and the contributions from the first (PC1) and second (PC2) principal components. Regardless of the soil sample dilution rate, most information is concentrated on PC1, above 66% for all data sets and higher for the data obtained with the multisensor array made with LbL films. We observed a trend in the data sets from the ion-selective membrane array, where the contribution of PC1 increased in proportion to the concentration of the soil aliquots, as detailed in Table 1. However, caution is warranted, as more than one physical or chemical variable may be influencing a principal component.

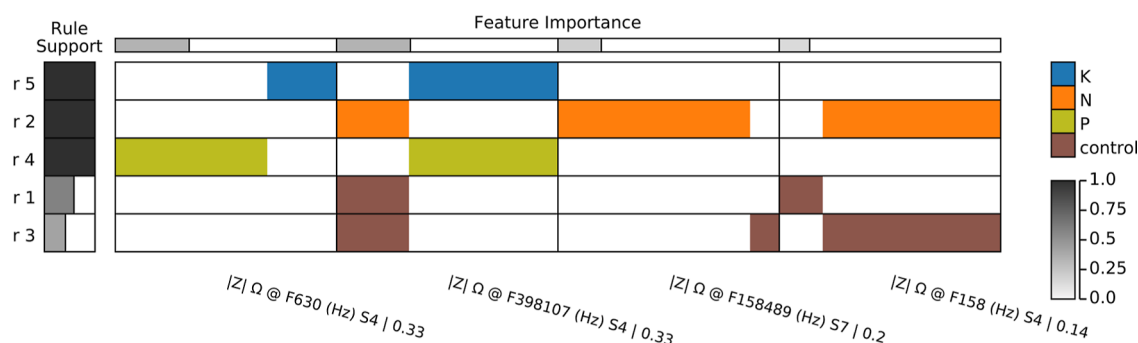


Figure 5. MCS for the multisensor array based on ion-selective membranes. The space has 4 dimensions corresponding to 4 frequencies (features) selected among the 248 available (F1 to F10⁶ Hz for 8 sensors). Rules with maximum support are found, e.g., rules r5, r2, and r4 (first to the third row), distinguishing all samples from classes K (blue), N (orange), and P (olive), respectively. Only two sensors in the multisensor array are used, namely sensor S4 at the frequencies F630, F398107, and F158 (first, second, and fourth columns) and sensor S7 at frequency F158489 (third column). The average accuracy estimated for the MCS via the DT model is 91.5%.

Figure 4 displays all data sets obtained with the multisensor array made with ion-selective membranes for the soil samples at four concentrations, featuring a concentration trend in PC1 with increasing concentrations positioned from right to left. Red, yellow, green, and blue areas in the graphs are related to soil samples prepared at 1, 10, 50, and 100 mg/mL, respectively. A similar trend was observed by Americo da Silva et al. for sandy and clayey soils, with PCA plots indicating a PC1 trend related to soil fertility.²⁸ As clearly seen in Figure 4 it is easy to discriminate all samples from the ion-selective membrane array, while the visual analysis of the data from the LbL-based array is not straightforward as a strong contribution from PC2 results in a boomerang-like pattern (Figure S6). The boomerang-shaped pattern observed in Figure S6 demonstrates a nonlinear response in the PC1 vs. PC2 plot across different concentrations of soil samples dispersed in water. This distinct pattern indicates that the LbL-based array achieves optimal clustering performance at certain concentration thresholds, thereby optimizing the detection efficacy for this specific set of soil samples and sensor configuration. For instance, the LbL-array fails to distinguish all soil macronutrients at 1 mg/mL using the *k*-means clustering method (Figure S5a). Additionally, the sample enriched with K at 50 mg/mL clusters in the same region as the 100 mg/mL samples (Figure S6). These results emphasize that the 10 mg/mL concentration offers optimal clustering performance for the LbL-array.

One of the most relevant challenges in sensor arrays is calibration, which can be troublesome when one of the sensing units of the array must be replaced. This can be partially addressed by using the concept of MCS¹⁶ which is based on machine learning methods rules, providing prediction power along with interpretability (i.e., explainability). In addition to assigning the class (e.g., soil samples category) in a forecasting task, these calibrations allow for establishing relationships between the classes and the impedance values and specific sensing units. For distinguishing soil samples into 4 classes (control, N, P, and K), two MCS are created from the impedance modulus spectra, one for the LbL-based device and the other for the ion-selective membranes setup. Each sensing unit provides the impedance magnitude at 31 frequencies (F1 to F10⁶ Hz), resulting in 124 frequencies for the LbL-based array comprising 4 sensing units (bare IDE, PDDA/CuTsPc, PDDA/MMt-K, and PDDA/PEDOT:PSS), and 248 frequencies for the ion-selective membrane array that integrates 8

sensing units (named S1 to S8, whose compositions are given in Table S1 of the Supporting Information). The calibrations (MCS) are generated by DT models for classifying samples among control, N, P, and K (target variable - soil categories) classes,^{48,49} with models' hyperparameters and accuracy (performance) obtained via the Nested KFold Cross-Validation procedure.^{52,53} The estimated average accuracy—i.e., model's performance in classifying samples among classes (control, N, P, and K)—is 91.5% for the ion-selective membrane array, while a 87% accuracy is achieved for the LbL-based array. These averages are not statistically different (*p*-value of 0.52578); therefore, based on accuracy both devices perform similarly.

The MCS can be visually represented and interpreted using the ExMatrix approach.⁵⁰ In this representation, logic rules derived from DT models are displayed in a matrix-like visual metaphor, where rows represent rules, columns signify features, and cells indicate rule predicates. These cells define the range of impedance magnitude values for different classes, such as soil sample categories, which are displayed using color coding. Figure 5 illustrates the ExMatrix representation for the MCS developed using the ion-selective membrane-based multisensor array, with the corresponding MCS for the LbL films array shown in Figure S7. Within the ExMatrix, rules (rows) are organized by class and support level, while features (columns) are ranked by importance. The MCS for the ion-selective membrane array includes four selected frequencies from the 248 available, spanning from F1 to F10⁶ Hz across eight sensors. Notably, data from only two sensing units, S4 and S7, are utilized among the eight available. The ion-selective capturing mechanisms of these membranes leverage specific ion-exchange interactions, where active compounds within the membranes bind selectively to target ions in the sample. Frequencies F630, F398107, and F158 emanate from the S4 sensing unit, which incorporates Mn(III) tetraphenylporphyrin chloride (MnTPPCL) for anion exchange, particularly effective with chloride ions from KCl, and tetradodecylammonium bromide that adjusts membrane surface properties affecting ion transport. The frequency F158489 originates from the S7 sensing unit, employing tetraoctyl diglycolamide (TODGA) for its strong cation-exchange capacity with potassium ions, enhanced by chlorinated cobalt dicarbollide, which further aids in selectively capturing potassium. This dual-component strategy in S7 enhances its sensitivity and selectivity for

potassium, showcasing the intricate balance of membrane composition and its impact on device performance.

For the LbL-based array, the MCS has 13 dimensions, i.e. thirteen selected frequencies among the 124 (F1 to F10⁶ Hz for 4 sensing units). At least two frequencies (feature/column) from each sensing unit (bare IDE, PDDA/CuTsPc, PDDA/MMt-K, and PDDA/PEDOT:PSS) are used. The rules in the MCS for ion-selective membranes are more generic (higher support) than the rules for the LbL-based array. Indeed, rules r5, r2, and r4 (first to the third row in Figure 5) have maximum support, i.e. they permit distinction of all soil samples from classes K (blue), N (orange), and P (olive), respectively. Ranges of impedance values from such rules can be interpreted regarding class associations. Taking rule r5 (first row), for example, class K (blue) is associated with soil samples containing high impedance (right-most rectangle) at the F630 frequency (first column), and medium to high values (center-to-right rectangle) at the F398107 frequency (second column), both for sensing unit S4. No rules from the MCS obtained with the LbL array have maximum support and that explains a certain data complexity to acquire rules distinguishing samples among their 4 possible classes.

In a nutshell, the calibrations obtained for the two multisensor arrays are typical examples of the potential of the MCS approach, relying on interpreting the machine learning model instead of only inspecting quantitative measurements. Both calibrations have similar performance (e.g., accuracy), but the MCS for the ion-selective membrane-based array (Figure 5) is simpler (4 dimensions) and more generic (rules with high support) than the MCS for the LbL-based device (13 dimensions and not having high support rules—Figure S7). Based on the calibrations created, the data (impedance at different frequencies) produced by the ion-selective membrane-based array present a better fingerprint of the soil samples among 4 classes/categories (control, K, N, and P). It is worth mentioning that frequencies and sensors employed in both calibrations (Figures 5 and S7) are chosen by DT model inference algorithm. Nevertheless, it is not possible to state unambiguously that other frequencies and sensor selections could not produce equivalent models' performance.

CONCLUSIONS

In conclusion, this study introduces a novel impedimetric multisensor array featuring ion-selective membranes that successfully discriminates between soil samples enriched with various macronutrients. The simplicity of the sample preparation—i.e., mere dispersion in water—eliminates the need for complex preprocessing, enhancing the practicality of soil analysis. The distinct profiles of these samples are evident from visual PCA and further supported by comparative results from another array utilizing LbL films. Utilization of *k*-means clustering on the impedance data reveals that although the LbL film-based array can statistically cluster and discriminate all samples, the ion-selective multisensor array exhibits exceptional sensitivity to K-enriched samples, clearly differentiating it from samples enriched with N and phosphorus P.

The calibration results from both sensor arrays highlight the versatility of the MCS, which prioritizes the interpretation of machine learning models instead of solely inspecting quantitative measurements. Although both arrays show comparable performance metrics, such as accuracy, the MCS for the ion-selective membrane-based array is notably simpler

and more universal than that for the LbL film-based array. Specifically, the ion-selective sensor array, utilizing data from sensors S4 and S7—which are designed to allow ion exchange with K⁺ and/or Cl[−]—provides a more definitive fingerprint of the soil samples across four categories (control, N, P, and K). These findings align with the observations from both PCA and *k*-means analysis.

While this array does not yet enable specific nutrient recognition, its heightened sensitivity toward K-enriched samples is a significant step toward selective detection, which is essential for identifying individual macronutrients in soil. These promising results open avenues for further research into developing new multisensor arrays for in situ soil analysis. Such technology promises a quick, economical method for supporting precision agriculture by simply diluting soil samples in water.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.4c04452>.

Table S1: Composition of the sensor array based on ion-selective membranes. Table S2: Soil composition regarding nitrogen, phosphorus, and potassium availability, according to wet-chemical analysis. Figure S1: Multisensor array based on LbL films and microfluidic device. Figure S2: Representative examples of impedance magnitude spectra obtained for the multisensor array based on 8 ion-selective membranes for NaCl and sucrose analytes. Figure S3: PCA score plots from impedance data in three independent sets of measurements evaluating basic tastes relevant to human gustative perception (sweet, salty, sour, bitter, and umami) at 1 mM obtained with the multisensory array based on LbL films. Figure S4: PCA score plots of the impedance data in three independent sets of measurements to evaluate soil samples individually enriched with N, P, and K. The data sets are obtained from the evaluation of the soil samples dispersed in water at (A) 1, (B) 10, (C) 50 and (D) 100 mg/mL with the multisensory array based on ion-selective membranes. Figure S5: PCA score plots of the impedance data in three independent sets of measurements to evaluate soil samples individually enriched with N, P, and K. The data sets are obtained from the evaluation of the soil samples dispersed in water at (A) 1, (B) 10, (C) 50 and (D) 100 mg/mL with the multisensory array based on LbL films. Figure S6: PCA score plots of the impedance data of control, N, P, and K soil samples dispersed in water at 1 (red), 10 (yellow), 50 (green), and 100 (blue) mg/mL obtained with the multisensory array based on LbL films. Colored areas are just a guide to the eye, once soil sample enriched with K at 50 mg/mL is in the blue area merged with the aliquots at 100 mg/mL (highlighted in the green ellipse). Figure S7: MCS for the multisensor array based on LbL films. The space has 13 dimensions corresponding to 13 frequencies (features) selected among the 124 available (F1 to F10⁶ Hz for 4 sensors). Rules with maximum support are not found, which may indicate a certain data complexity, and at least two frequencies from each sensor employed in the multisensor array (bare IDE, PDDA/CuTsPc, PDDA/MMt-

K, and PDDA/PEDOT:PSS) are used. The average accuracy estimated for the MCS via the DT model is 87% (PDF)

AUTHOR INFORMATION

Corresponding Author

Antonio Riul Jr. – Instituto de Física “Gleb Wataghin” (IFGW), Universidade Estadual de Campinas—UNICAMP, Campinas 13083-859 São Paulo, Brazil; orcid.org/0000-0002-9760-1851; Email: riul@unicamp.br

Authors

Maria Luisa Braunger – Instituto de Física “Gleb Wataghin” (IFGW), Universidade Estadual de Campinas—UNICAMP, Campinas 13083-859 São Paulo, Brazil; orcid.org/0000-0002-6794-5362

Mario Popolin Neto – Federal Institute of São Paulo—IFSP, Araraquara 14804-296 São Paulo, Brazil

Dmitry Kirsanov – Institute of Chemistry, Mendeleev Center, St. Petersburg State University, St. Petersburg 199034, Russia; Laboratory of Artificial Sensory Systems, ITMO University, St. Petersburg 197101, Russia

Igor Fier – Quantum Design Latin America, Campinas 13080-655 São Paulo, Brazil

Lucas R. Amaral – School of Agricultural Engineering (FEAGRI), University of Campinas—UNICAMP, Campinas 13083-875 São Paulo, Brazil

Flavio M. Shimizu – Instituto de Física “Gleb Wataghin” (IFGW), Universidade Estadual de Campinas—UNICAMP, Campinas 13083-859 São Paulo, Brazil

Daniel S. Correa – Nanotechnology National Laboratory for Agriculture (LNNA), Embrapa Instrumentação, São Carlos 13560-970 São Paulo, Brazil; orcid.org/0000-0002-5592-0627

Fernando V. Paulovich – Department of Mathematics and Computer Science, Eindhoven University of Technology (TU/e), Eindhoven 5600 MB, The Netherlands

Andrey Legin – Institute of Chemistry, Mendeleev Center, St. Petersburg State University, St. Petersburg 199034, Russia; Laboratory of Artificial Sensory Systems, ITMO University, St. Petersburg 197101, Russia

Oswaldo N. Oliveira Jr. – São Carlos Institute of Physics (IFSC), University of São Paulo—USP, São Carlos 13566-590 São Paulo, Brazil; orcid.org/0000-0002-5399-5860

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.4c04452>

Funding

The Article Processing Charge for the publication of this research was funded by the Coordination for the Improvement of Higher Education Personnel - CAPES (ROR identifier: 00x0ma614).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by FAPESP (no 2015/14836-9, 2018/22214-6), CAPES, CNPq (308570/2018-9, 308943/2021-0, 310745/2022-5, BRICS-STI 29/2017 no 442196/2017-2, 402676/2021-1, 403359/2023-6), INEO (FAPESP/CNPq/INCT no 2014/50869-6), RFBR BRICS project #18-

53-80010. Additionally, we would like to thank the University of Ljubljana (Slovenia) for the Orange Data Mining software.

LIST OF ABBREVIATIONS AND SYMBOLS

AC	alternating current
Ca	calcium
CuTsPc	copper phthalocyanine-3,4',4'',4'''-tetra-sulfonic acid tetrasodium salt
DC	direct current
DOS	dioctyl sebacate
DT	decision trees
HCl	hydrochloric acid
IDE	interdigitated electrode
IDMAP	interactive document mapping
K	potassium
LbL	layer-by-layer
MCS	multidimensional calibration space
Mg	magnesium
MMt-K	montmorillonite clay
N	nitrogen
NaCl	sodium chloride
NPOE	<i>o</i> -nitrophenyloctyl ether
P	phosphorus
PCA	principal component analysis
PCB	printed circuit board
PC1	first principal component
PC2	second principal component
PDDA	poly(diallyldimethylammonium chloride) solution
PDDA/CuTsPc	layer-by-layer film of PDDA and CuTsPc
PDDA/MMt-K	layer-by-layer film of PDDA and MMt-K
PDDA/PEDOT:PSS	layer-by-layer film of PDDA and PEDOT:PSS
PEDOT:PSS	poly(3,4-ethylenedioxythiophene)–poly(styrenesulfonate)
PVC	poly(vinyl chloride)
r1 to r15	logic rules represented from 1 to 15 rows in a matrix
S	sulfur
SC	silhouette coefficient
<i>s</i> (<i>i</i>)	silhouette value
S1 to S8	ion-selective membranes from 1 to 8, according to compounds presented in Table S1
THF	tetrahydrofuran

REFERENCES

- (1) United Nations. *World Population Prospects 2019: Highlights; Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report*; UNLibrary: New York, 2019.
- (2) FAO. *The Future of Food and Agriculture: Trends and Challenges*; Food and Agriculture Organization of the United Nations: Rome, 2017.
- (3) World Wildlife. *Living Planet Report 2020 - Bending the Curve of Biodiversity Loss*; Almond, R. E. A., Grooten, M.; Petersen, T., Eds.; WWF International: Gland, 2020.
- (4) Nadporozhskaya, M.; Kovsh, N.; Paolesse, R.; Lvova, L. Recent Advances in Chemical Sensors for Soil Analysis: A Review. *Chemosensors* **2022**, *10* (1), 35.
- (5) Adamchuk, V. I.; Hummel, J. W.; Morgan, M. T.; Upadhyaya, S. K. On-the-Go Soil Sensors for Precision Agriculture. *Comput. Electron Agric* **2004**, *44* (1), 71–91.

- (6) Ferguson, R. B.; Hergert, G. W. *Soil Sampling for Precision Agriculture*; Institute of Agriculture and Natural Resources, 2009.
- (7) Leukel, J.; Zimpel, T.; Stumpe, C. Machine Learning Technology for Early Prediction of Grain Yield at the Field Scale: A Systematic Review. *Comput. Electron. Agric.* **2023**, *207*, 107721.
- (8) Mazuryk, J.; Klepacka, K.; Kutner, W.; Sharma, P. S. Glyphosate Separating and Sensing for Precision Agriculture and Environmental Protection in the Era of Smart Materials. *Environ. Sci. Technol.* **2023**, *57*, 9898–9924.
- (9) Peng, B.; Liu, X.; Yao, Y.; Ping, J.; Ying, Y. A Wearable and Capacitive Sensor for Leaf Moisture Status Monitoring. *Biosens. Bioelectron.* **2024**, *245*, 115804.
- (10) Zhou, S.; Zhou, J.; Pan, Y.; Wu, Q.; Ping, J. Wearable Electrochemical Sensors for Plant Small-Molecule Detection. *Trends Plant Sci.* **2024**, *29*, 219–231.
- (11) Lo Presti, D.; Di Tocco, J.; Massaroni, C.; Cimini, S.; De Gara, L.; Singh, S.; Raucci, A.; Manganiello, G.; Woo, S. L.; Schena, E.; Cinti, S. Current Understanding, Challenges and Perspective on Portable Systems Applied to Plant Monitoring and Precision Agriculture. *Biosens. Bioelectron.* **2023**, *222*, 115005.
- (12) Wesoly, M.; Przewodowski, W.; Ciosek-Skibińska, P. Electronic Noses and Electronic Tongues for the Agricultural Purposes. *TrAC, Trends Anal. Chem.* **2023**, *164*, 117082.
- (13) Cetó, X.; Valle, M. del. Electronic Tongue Applications for Wastewater and Soil Analysis. *iScience* **2022**, *25* (5), 104304.
- (14) Tahara, Y.; Toko, K. Electronic Tongues-a Review. *IEEE Sens. J.* **2013**, *13* (8), 3001–3011.
- (15) Podrazka, M.; Baczynska, E.; Kundys, M.; Jelen, P. S.; Nery, E. W. Electronic Tongue — A Tool for All Tastes? *Biosensors* **2017**, *8* (1), 3.
- (16) Neto, M. P.; Soares, A. C.; Oliveira, O. N.; Paulovich, F. V. Machine Learning Used to Create a Multidimensional Calibration Space for Sensing and Biosensing Data. *Bull. Chem. Soc. Jpn.* **2021**, *94* (5), 1553–1562.
- (17) Legin, A.; Rudnitskaya, A.; Vlasov, Y.; Di Natale, C.; Davide, F.; D'Amico, A. Tasting of Beverages Using an Electronic Tongue. *Sens. Actuators, B* **1997**, *44* (1–3), 291–296.
- (18) Winquist, F.; Bjorklund, R.; Krantz-Rülcker, C.; Lundström, I.; Östergren, K.; Skoglund, T. An Electronic Tongue in the Dairy Industry. *Sens. Actuators, B* **2005**, *111–112*, 299–304.
- (19) Blanco, C. A.; De La Fuente, R.; Caballero, I.; Rodríguez-Méndez, M. L. Beer Discrimination Using a Portable Electronic Tongue Based on Screen-Printed Electrodes. *J. Food Eng.* **2015**, *157*, 57–62.
- (20) Liu, Y.; Wu, X.; Tahara, Y.; Ikezaki, H.; Toko, K. A Quantitative Method for Acesulfame K Using the Taste Sensor. *Sensors* **2020**, *20* (2), 400.
- (21) Facure, M. H. M.; Mercante, L. A.; Mattoso, L. H. C.; Correa, D. S. Detection of Trace Levels of Organophosphate Pesticides Using an Electronic Tongue Based on Graphene Hybrid Nanocomposites. *Talanta* **2017**, *167*, 59–66.
- (22) Kirsanov, D.; Correa, D. S.; Gaal, G.; Riul, A., Jr.; Braunger, M. L.; Shimizu, F. M.; Oliveira, O. N., Jr.; Liang, T.; Wan, H.; Wang, P.; Oleneva, E.; Legin, A. Electronic Tongues for Inedible Media. *Sensors* **2019**, *19* (23), 5113.
- (23) Herrera-Chacón, A.; Torabi, F.; Faridbod, F.; Ghasemi, J. B.; González-Colabuig, A.; Del Valle, M. Voltammetric Electronic Tongue for the Simultaneous Determination of Three Benzodiazepines. *Sensors* **2019**, *19* (22), 5002.
- (24) Mimendia, A.; Gutiérrez, J. M.; Alcañiz, J. M.; del Valle, M. Discrimination of Soils and Assessment of Soil Fertility Using Information from an Ion Selective Electrodes Array and Artificial Neural Networks. *Clean* **2014**, *42* (12), 1808–1815.
- (25) Khaydukova, M.; Kirsanov, D.; Sarkar, S.; Mukherjee, S.; Ashina, J.; Bhattacharyya, N.; Chanda, S.; Bandyopadhyay, R.; Legin, A. One Shot Evaluation of NPK in Soils by “Electronic Tongue”. *Comput. Electron. Agric.* **2021**, *186*, 106208.
- (26) Braunger, M. L.; Shimizu, F. M.; Jimenez, M. J. M.; Amaral, L. R.; Piazzetta, M. H. de O.; Gobbi, A. L.; Magalhães, P.; Rodrigues, V.; Oliveira, O. N.; Riul, A. Microfluidic Electronic Tongue Applied to Soil Analysis. *Chemosensors* **2017**, *5* (2), 14.
- (27) Gaál, G.; da Silva, T. A.; Gaál, V.; Hensel, R. C.; Amaral, L. R.; Rodrigues, V.; Riul, A. 3D Printed E-Tongue. *Front. Chem.* **2018**, *6*, 1–8.
- (28) da Silva, T. A.; Braunger, M. L.; Neris Coutinho, M. A.; Rios do Amaral, L.; Rodrigues, V.; Riul, A. 3D-Printed Graphene Electrodes Applied in an Impedimetric Electronic Tongue for Soil Analysis. *Chemosensors* **2019**, *7* (4), 50.
- (29) Braunger, M. L.; Fier, I.; Shimizu, F. M.; de Barros, A.; Rodrigues, V.; Riul, A. Influence of the Flow Rate in an Automated Microfluidic Electronic Tongue Tested for Sucralose Differentiation. *Sensors* **2020**, *20* (21), 6194.
- (30) Kirsanov, D. O.; Legin, A. V.; Kulikova, A. P.; Pol'shin, E. N.; Vlasov, Y. G. Polymeric Sensors for Determination of Anions of Organic Acids. *Russ. J. Appl. Chem.* **2007**, *80* (5), 799–804.
- (31) Legin, A. V.; Babain, V. A.; Kirsanov, D. O.; Mednova, O. V. Cross-Sensitive Rare Earth Metal Ion Sensors Based on Extraction Systems. *Sens. Actuators, B* **2008**, *131* (1), 29–36.
- (32) Richardson, J. J.; Björnalm, M.; Caruso, F. Technology-Driven Layer-by-Layer Assembly of Nanofilms. *Science* **2015**, *348* (6233), 2491.
- (33) Braunger, M.; Fier, I.; Rodrigues, V.; Arratia, P.; Riul, A., Jr. Microfluidic Mixer with Automated Electrode Switching for Sensing Applications. *Chemosensors* **2020**, *8* (1), 13.
- (34) Schiffman, S. S.; Crumbliss, A. L.; Warwick, Z. S.; Graham, B. G. Thresholds for Sodium Salts in Young and Elderly Human Subjects: Correlation with Molar Conductivity of Anion. *Chem. Senses* **1990**, *15* (6), 671–678.
- (35) Heath, T. P.; Melichar, J. K.; Nutt, D. J.; Donaldson, L. F. Human Taste Thresholds Are Modulated by Serotonin and Noradrenaline. *J. Neurosci.* **2006**, *26* (49), 12664–12671.
- (36) Shigemura, N.; Shirosaki, S.; Sanematsu, K.; Yoshida, R.; Ninomiya, Y. Genetic and Molecular Basis of Individual Differences in Human Umami Taste Perception. *PLoS One* **2009**, *4* (8), No. e6717.
- (37) Mouillot, T.; Barthet, S.; Janin, L.; Creteau, C.; Devilliers, H.; Brindisi, M. C.; Penicaud, L.; Leloup, C.; Brondel, L.; Jacquin-Piques, A. Taste Perception and Cerebral Activity in the Human Gustatory Cortex Induced by Glucose, Fructose, and Sucrose Solutions. *Chem. Senses* **2019**, *44* (7), 435–447.
- (38) Hiroyuki Higa, B.; Rios do Amaral, L. Preparação de Amostras de Solo Para Espectroscopia e Seu Uso No Mapeamento Dos Parâmetros de Fertilidade Do Solo Em Áreas Cultivadas Com Cana-de-Açúcar. In *XXIV Congresso de Iniciação Científica da UNICAMP* 2016, 51131..
- (39) Higa, B. H.; Rios do Amaral, L. Preparação de Amostras de Solo Para Espectroscopia e Quantificação de Macronutrientes Disponíveis No Solo. In *Congresso Brasileiro de Agricultura de Precisão*, 2016.
- (40) Higa, B. H.; Rios do Amaral, L. Preparação de Amostras de Solo Para Espectroscopia e Seu Uso No Mapeamento Dos Parâmetros de Fertilidade Do Solo Em Áreas Cultivadas Com Cana-de-Açúcar. In *XXIV Congresso de Iniciação Científica da UNICAMP*, 2016.
- (41) Coutinho, M. A. N.; Alari, F. D. O.; Ferreira, M. M. C.; Amaral, L. R. do. Influence of Soil Sample Preparation on the Quantification of NPK Content via Spectroscopy. *Geoderma* **2019**, *338*, 401–409.
- (42) Jolliffe, I. T.; Cadima, J. Principal component analysis: a review and recent developments. *Phys. Eng. Sci.* **2016**, *374* (2065), 20150202.
- (43) Rousseeuw, P. J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
- (44) Struyf, A.; Hubert, M.; Rousseeuw, P. Clustering in an Object-Oriented Environment. *J. Stat. Softw.* **1996**, *1* (4), 1–30.
- (45) Wu, J. Advances in K-Means Clustering. In *Springer Theses*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; Vol. 13..
- (46) *Orange Data Mining*. University of Ljubljana, 2024. <https://orange.biolab.si/>.
- (47) Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification And Regression Trees*; Routledge, 2017; ..

- (48) Tan, P.-N.; Steinbach, M.; Kumar, V. *Introduction to Data Mining*; Pearson Addison-Wesley: Boston, 2006.
- (49) James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer Texts in Statistics; Springer New York: New York, NY, 2013; Vol. 103..
- (50) Neto, M. P.; Paulovich, F. V. Explainable Matrix - Visualization for Global and Local Interpretability of Random Forest Classification Ensembles. *IEEE Trans Vis Comput. Graph* **2021**, 27 (2), 1427–1437.
- (51) Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *International Joint Conference of Artificial Intelligence* 1995.
- (52) Varma, S.; Simon, R. Bias in Error Estimation When Using Cross-Validation for Model Selection. *BMC Bioinf.* **2006**, 7, 91.
- (53) Tsamardinos, I.; Rakhshani, A.; Lagani, V. Performance-Estimation Properties of Cross-Validation-Based Protocols with Simultaneous Hyper-Parameter Optimization. *Int. J. Artif. Intell. Tool.* **2015**, 24 (05), 1540023.