2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)
Mexico City, Mexico November 10-12, 2021

A Resampling Approach for the Data-Based Optimization of Nanosensors†

Patricia G. López-Cárdenas
Research Laboratory on Optimal Design, Devices and Advanced Materials, Department of Mathematics and Physics ITESO
Tlaquepaque, Jalisco, Mexico
plopez@iteso.mx

Emmanuel Alcalá
Research Laboratory on Optimal Design, Devices and Advanced Materials, Department of Mathematics and Physics ITESO
Tlaquepaque, Jalisco, Mexico
jaime.alcala@iteso.mx

Juan Diego Sánchez-Torres
Research Laboratory on Optimal Design, Devices and Advanced Materials, Department of Mathematics and Physics ITESO
Tlaquepaque, Jalisco, Mexico
dsanchez@iteso.mx

Elsie Araujo*
Research Laboratory on Optimal Design, Devices and Advanced Materials, Department of Mathematics and Physics ITESO
Tlaquepaque, Jalisco, Mexico
elsie@iteso.mx

***Abstract*—Nanostructured sensors are a promising alternative for substance detection from biological samples in food and clinical industries. Linear regression is often used for the design and calibration, relating the sensor response with known concentrations of the substance to be detected. The linear regression slope is then used to estimate the sensor sensitivity and the limits of detection and quantification. These estimates, however, are reported without measures of uncertainty, like confidence intervals. Here we used nonparametric bootstrapping to compute the estimates and their confidence intervals using the data from cyclic voltammetry and compare two classes of sensors.**

***Index Terms*—nanostructured sensors, bootstrap, resampling, sensitivity**

I. INTRODUCTION

The need for more sensitive and selective non-enzymatic sensors has led to the development of nanostructure-based sensors, on which the manipulation of the geometry and area have given promising results with electrochemical methods [1], [2]. However, accurate designs require knowing most of the behavior present in the sensors, leading to complex and complicated models. Using those representations often presents drawbacks for computation and the need for model validation. Therefore, it is common to consider other approaches that rely directly on the experimental or simulated data obtained from the systems inputs, outputs, and tunable design variables. Some examples are the use of the design of experiments to improve device verification at the industrial level [3], machine learning for sensor analysis [4] and calibration [5], and the application of kriging metamodeling [6], [7]. Therefore, in the experimental literature, the sensor's performance is often summarized in several statistics, like the slope of the regression line relating the sensor response with a known concentration of the substance (for example, hydrogen peroxide, H_2O_2, as in this study), as well as the associated p-value [8], [9]. However, several current guidelines recommend reporting not just the point estimate but also the interval estimate, as the confidence interval [10], [11] for assessing the accuracy of the estimation.

The aim of this work is twofold. First, we illustrate the computationally intensive procedure of bootstrapping, a distribution-free technique, to obtain confidence intervals for several estimates that summarize a sensor performance [12]. Second, we show the strength of interval estimation using the bootstrap, giving more precise results without parametric assumptions. For this purpose, we used data obtained with self-supported nanowire arrays sensors using cyclic voltammetry (CV), an electrochemical method that measures the electron transfer by oxidation-reduction reactions under the application of linearly ramped potentials. We tested the sensor response with concentrations of H_2O_2 of 0, 1.27, 2.53, 3.79, and 5.04 mM and used two types of metals for the sensors, Nickel (Ni) and Gold (Au), with two different geometries, nanowires (NW) and Planar. We used the planar geometry for comparison because it represents the more extreme case of the no-nanostructured sensor [13].

†The authors thank COECYTJAL for partially financial support by project FODECIJAL 8248-2019 and ITESO by the Research Support Fund FAI-2018. Also, P.G.L.C thanks CONACYT for Ph.D. scholarship No. 751908, and E.Alcalá thanks COECYTJAL for a postdoctoral fellowship throughout the project FODECIJAL 8248-2019.

* Corresponding author.

978-1-6654-0029-9/21/$31.00 ©2021 IEEE

II. Preliminaries

The bootstrap belongs to a class of statistical tools called resampling methods. These methods share three basic steps [14]–[16]:

1) Repeatedly (re)sampling (with or without replacement) from a dataset.
2) Computing a statistic from the new sample (a subset of the original).
3) Compute the summary statistics (for example, measures of variation).

The resampling methods are often conceptually more straightforward and accurate than the asymptotic approximation methods and require fewer assumptions [14]. Other resampling methods include permutation test for hypothesis testing, cross-validation for model assessment, and model selection based on the model performance [16], [17].

In this work, we used the bootstrap to obtain measures of accuracy of estimates, i.e., interval estimation [10], with linear regression. We start by briefly describing the sensor, the cyclic voltammetry data obtention, and transformation.

A. Data acquisition

Fig. 1 diagrammatically shows the experimental setup for the fabrication of the sensor using restrictive nanoporous membranes (panel A) by the electrodeposition of Au and Ni ions [13], [18], as well as the data acquisition process using cyclic voltammetry (panel B). The response (current intensity) was transformed to current density (I density, in mA/cm^2), dividing the current by the sensor area. In panel C, we also show representative CV data from the Au NW sensor.

We are interested in the sensitivity of the sensor, which we define as the change in current density by a unit increase in H_2O_2 concentration, with dimensions mA/(mM·cm^2). Because the sensitivity critically depends on the potential applied [19], we identified the potential V^* at which is maximized. Panel C shows this process as follows: in the reduction peak region of the CV (approximately from -0.5 V to 0 V), we interpolated I density for each i-th potential with a step size of 0.0005 V. Hence, for each $V[i]$ we obtained five pairs (I density$_{V[i]}$, concentration$_k$) for the $k = 5$ concentrations, and estimated the sensitivity of the sensor with the slope of the linear regression β in

$$I \text{ density}_{V[i]} = \alpha_{V[i]} + \beta_{V[i]} \times \text{H}_2\text{O}_2 \text{ concentration}$$

The potential at which the sensitivity is maximized is therefore $\underset{V}{\operatorname{argmax}}\, \beta = V^*$.

III. Nonparametric bootstrapping for linear regression

The idea of the bootstrap is to simulate the sampling distribution variation, which arises when we repeatedly draw samples from a population, by resampling from the original sample. The resulting variation is an approximation of the variation of the sampling distribution [12].

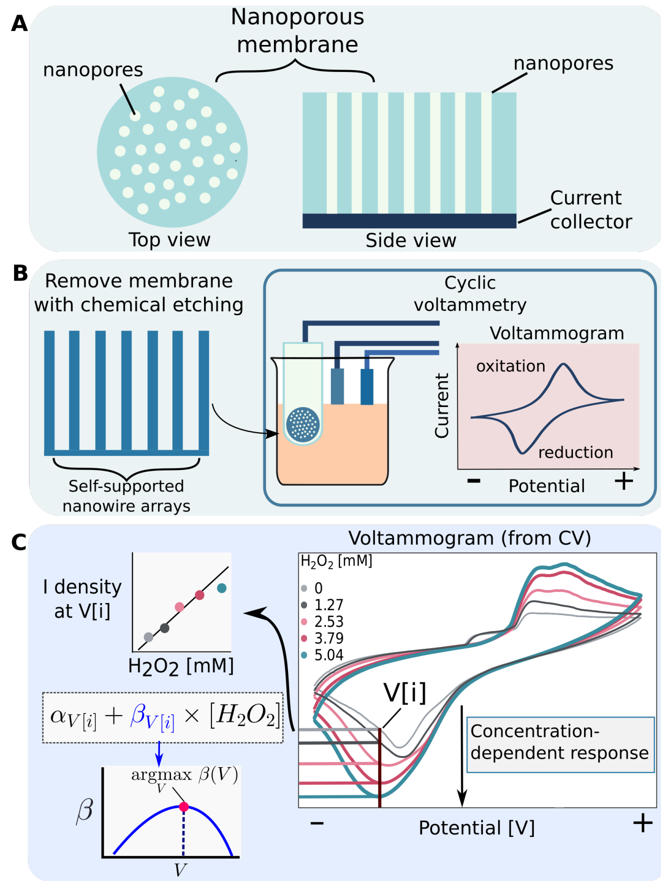

Fig. 1. Data acquisition and transformation. **A:** Top and side views of the restrictive nanoporous membrane used for electrodeposition. **B:** After electodeposition, self-supported nanowire arrays are released by chemical etching. **C:** Computation of the sensing potential at which the sensitivity is maximized, using representative data form Au NW.

To illustrate how to compute the percentile confidence intervals for linear regression, let us suppose we have a sample X of size n of ordered pairs $(x_i, y_i), i = 1, 2, ..., n$ (bivariate observations). Let X_b^* be the b-th resample of X, and (x_i^*, y_i^*) a bivariate observation from X^*. We sample X^* from X with replacement; that is, we allow the pair (x_i^*, y_i^*) to appear more than once (or none). Then, we compute the statistic of interest from (x_i^*, y_i^*) (for example, the linear regression slope) and store the b-th result. We repeat this process B times, and the 95 % CI can be obtained with the 0.025 and 0.975 quantiles of the bootstrap sample of size B [20].

Fig. 2 shows the regression lines computed with 250 resamples from the original data for both classes of sensors (NW and Planar) and metals (Ni and Au). The shaded areas show the 95 % CI computed with an asymptotic approximation based on the t-statistic with the expression

$$\text{CI}_{95\%} = \left[\beta - s_\beta t_{n-2}^*,\ \beta + s_\beta t_{n-2}^*\right]$$

On which β is the sensor sensitivity, s_β is the standard error of β, and t_{n-2}^* is the t-statistic obtained, with $n-2$ degrees of freedom. Note that in Fig. 2, regression lines with the bootstrap samples do not necessarily cover the asymptotically approximated CI. This means that different conclusions can be obtained with both methods. Indeed, with equation III, the 95 % CI of sensitivity for Ni NW is [1.79, 4.25]; for Ni Planar is [-0.001, 0.002], both wider than the obtained with the bootstrap (Table I), and for the Ni Planar, the CI includes zero, which can be interpreted as non-significant sensitivity. Similarly, for the Au NW [4.12, 6.27] and, more dramatically, for the Au Planar [0.045, 0.22] (compare them with Table I).

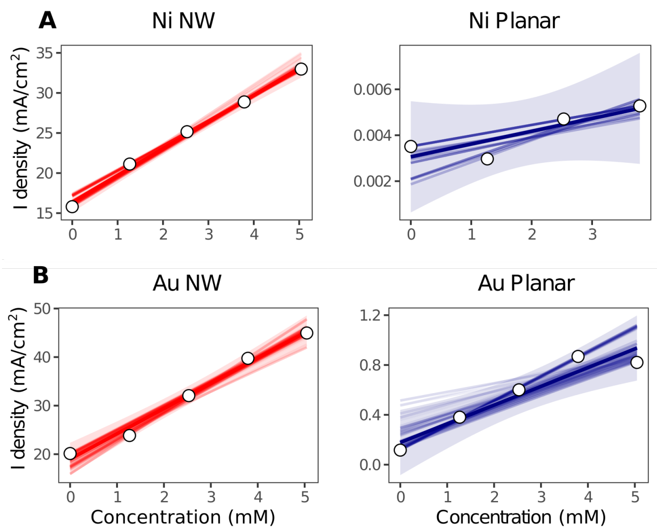

Fig. 2. Linear regressions relating current intensity with H_2O_2 concentrations. **A:** Regression lines, with 250 bootstrap samples, for Ni NW (red) and Planar (blue). **B:** Regression lines for Au NW (red) and Au Planar (blue), also with 250 bootstrap samples. In both cases, shaded areas show the 95 % CI computed with asymptotic approximation methods with III

TABLE I
PERCENTILE BOOTSTRAP 95 % CIS FOR SENSITIVITY, LOD AND LOQ.

Sensor Class	Ni Estimates[a]		
	β mA/(mM·cm^2)	LOD mM	LOQ mM
NW	3.327 [3.073, 3.698]	0.3 [0.04, 0.542]	0.4 [0.053, 0.723]
Planar	0.001 [0.0004, 0.001]	2.629 [0.06, 5.283]	3.505 [0.08, 7.05]
Sensor Class	**Au Estimates**		
	β mA/(mM·cm^2)	LOD mM	LOQ mM
NW	5.287 [4.891, 6.299]	0.551 [0.067, 0.913]	0.734 [0.089, 1.217]
Planar	0.157 [0.105, 0.198]	1.348 [0.114, 2.878]	1.798 [0.152, 3.837]

[a]The 95% CI are shown in square brackets.

Besides point and interval estimates of β, Table I shows the estimates of LOD and LOQ with the 95% confidence intervals computed with the bootstrap. Confidence intervals that do not contain zero are significant; that is, the estimates are significantly different from zero. We can see that the LOD and LOQ 95 % CIs are wider for the Planar class of sensors. This can be stressed as follows: the estimates of LOD and LOQ for NW sensors are lower so that a smaller quantity of H_2O_2 can be quantified with a high degree of confidence [21]; moreover, because the 95 % CIs of the NW sensors are narrower than the CIs of the Planar sensors, their LOD and LOQ estimates are more accurate.

IV. CONCLUSION

The class of high sensitivity and selective non-enzymatic sensors treated here crucially depends on the working conditions of the electrochemical response. This is because the reaction potential, which the electrochemical methods exploit, can be different for different substances. That means that if the sensing potential is not correctly settled, the simultaneous oxidation of other substances can interfere with the measurements [1], [22]. Therefore, for sensitivity and selectivity, better estimation methods are necessary to confidently separate the signal from the noise coming from interfering substances when the sensor is developed and calibrated, whenever possible.

Here we illustrated the use of distribution-free methods like the nonparametric bootstrap for point and interval estimation, methods that the statistical literature highly recommends as a complement or an alternative for the more used asymptotic approximation methods, like the t-tests [10], [12]. We also showed how the 95 % CI could give additional information for assessing the reliability and precision of point estimates, information that can and should be used for calibration. For example, although the point estimates for the LOD and LOQ in the Au NW sensor are lower than those of the corresponding Planar sensors, their 95 % CIs overlaps, which suggests no differences [23].

ACKNOWLEDGMENT

The authors thank Sofía Vilchis Sánchez and Lourdes Susana Martínez García for helping us in data recollection, and Eleazar Benítez Nuñez for his constant support.

REFERENCES

[1] L. Lin, S. Weng, Y. Zheng, X. Liu, S. Ying, F. Chen, and D. You, "Bimetallic ptau alloy nanomaterials for nonenzymatic selective glucose sensing at low potential," *Journal of Electroanalytical Chemistry*, vol. 865, p. 114147, 2020.

[2] E. Araujo, A. Encinas, Y. Velázquez-Galván, J. M. Martínez-Huerta, G. Hamoir, E. Ferain, and L. Piraux, "Artificially modified magnetic anisotropy in interconnected nanowire networks," *Nanoscale*, vol. 7, no. 4, pp. 1485–1490, 2015.

[3] A. Viveros-Wacher and J. E. Rayas-Sánchez, "Eye diagram optimization based on design of experiments (doe) to accelerate industrial testing of high speed links," in *2016 IEEE MTT-S Latin America Microwave Conference (LAMC)*, 2016, pp. 1–3.

[4] J. C. M. Gomes, L. C. Souza, and L. C. Oliveira, "SmartSPR sensor: Machine learning approaches to create intelligent surface plasmon based sensors," *Biosensors and Bioelectronics*, vol. 172, p. 112760, 2021.

[5] Z. Geng, F. Yang, M. Li, and N. Wu, "A bootstrapping-based statistical procedure for multivariate calibration of sensor arrays," *Sensors and Actuators B: Chemical*, vol. 188, pp. 440–453, 2013.

[6] W. C. M. Van Beers and J. P. C. Kleijnen, "Customized sequential designs for random simulation experiments: Kriging metamodeling and bootstrapping," *European journal of operational research*, vol. 186, no. 3, pp. 1099–1113, 2008.

[7] A. P. Kyprioti, J. Zhang, and A. A. Taflanidis, "Adaptive design of experiments for global kriging metamodeling through cross-validation information," *Structural and Multidisciplinary Optimization*, vol. 62, no. 3, pp. 1135–1157, 2020.

[8] M.-Y. Hua, H.-C. Chen, C.-K. Chuang, R.-Y. Tsai, J.-L. Jeng, H.-W. Yang, and Y.-T. Chern, "The intrinsic redox reactions of polyamic acid derivatives and their application in hydrogen peroxide sensor," *Biomaterials*, vol. 32, no. 21, pp. 4885–4895, 2011.

[9] X. Cui, Z. Li, Y. Yang, W. Zhang, and Q. Wang, "Low-potential sensitive hydrogen peroxide detection based on nanotubular TiO_2 and platinum composite electrode," *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis*, vol. 20, no. 9, pp. 970–975, Apr. 2008.

[10] R. J. Calin-Jageman and G. Cumming, "The new statistics for better science: Ask how much, how uncertain, and what else is known," *The American Statistician*, vol. 73, no. sup1, pp. 271–280, 2019.

[11] G. Cumming, *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge, 2013.

[12] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Statistical science*, pp. 54–75, 1986.

[13] P. G. López-Cárdenas, E. Alcalá, J. D. Sánchez-Torres, and E. Araujo, "Enhancing the sensitivity of a class of sensors: A data-based engineering approach," in *2021 IEEE 21th International Conference on Nanotechnology (IEEE-NANO)*, 2021, pp. 221–224.

[14] P. I. Good, *Introduction to statistics through resampling methods and R*. John Wiley & Sons, 2013.

[15] L. M. Chihara and T. C. Hesterberg, *Mathematical statistics with resampling and R*. John Wiley & Sons, 2018.

[16] B. Efron and T. Hastie, *Computer age statistical inference*. Cambridge University Press, 2016, vol. 5.

[17] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.

[18] F. Nasirpouri, *Electrodeposition of nanostructured materials*. Springer, 2017, vol. 62.

[19] M. M. Rahman, M. M. Hussain, and A. M. Asiri, "d-glucose sensor based on $ZnO-V_2O_5$ NRs by an enzyme-free electrochemical approach," *RSC advances*, vol. 9, no. 54, pp. 31 670–31 682, Oct. 2019.

[20] L. Wasserman, *All of statistics: a concise course in statistical inference*. Springer Science+Business Media, 2004.

[21] D. A. Armbruster and T. Pry, "Limit of blank, limit of detection and limit of quantitation," *The clinical biochemist reviews*, vol. 29, no. Suppl 1, p. S49, Aug. 2008.

[22] M. Ghanei-Motlagh, M. A. Taher, M. Fayazi, M. Baghayeri, and A. Hosseinifar, "Non-enzymatic amperometric sensing of hydrogen peroxide based on vanadium pentoxide nanostructures," *Journal of The Electrochemical Society*, vol. 166, no. 6, p. B367, 2019.

[23] G. Cumming and S. Finch, "Inference by eye: confidence intervals and how to read pictures of data." *American psychologist*, vol. 60, no. 2, p. 170, 2005.