Machine Learning-Assisted Calibration of Hg^{2+} Sensors Based on Carbon Nanotube Field-Effect Transistors

Long Bian,[†] Zhunheng Wang,[‡] David L. White,[†] Alexander Star[† §*]

[†] Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

[‡]Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

[§]Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

ABSTRACT:

Nanomaterial-based electronic sensors have demonstrated ultra-low detection limits, down to parts-per-billion (ppb) or parts-per-trillion (ppt) concentrations. However, these extreme sensitivities also make them susceptible to signal saturation at higher concentrations and restrict their usage primarily to low concentrations. Here, we report machine learning techniques to create a calibration method for carbon nanotube-based field-effect transistor (FET) devices. We started with linear regression, followed by regression splines to capture the non-linearity in the data. Further improvements in model performance were obtained with regression trees. Finally we lowered the model variance and further boosted the model performance by introducing random forest. The resulting performance as measured by R^2 was estimated to be 0.8260 using out-of-bag error. The methodology avoids saturation and extends the dynamic range of the nanosensors up to 12 orders of magnitude in analyte concentrations. Further investigations of the sensing mechanism include analysis of feature importance in each of the model we tested. Functionalized nanosensors demonstrate selective detection of Hg^{2+} ions with detection limits $10^{-14.36 \pm 0.78}$ M, and maintain

calibration to concentrations as high as 1 mM. Application of machine learning techniques to investigate which features in the FET signal maximally correlate with concentration changes provide valuable insight into the carbon nanotube sensing mechanism and assist in the rational design of future nanosensors.

KEYWORDS: carbon nanotubes • graphene • nanosensor calibration • multivariable regression • molecular recognition • machine learning

1. Introduction

Sensors are devices that measure physical and chemical changes in the local environment, i.e., presence of analyte, and convert those changes into readable signals (Amjadi et al. 2016; Bandodkar et al. 2019; Bariya et al. 2018; Kim et al. 2019). Sensors have been widely applied in medical, industrial, and environmental fields. Mercury, which has been increasing in concentration for the past 15 years (Liu et al. 2018), is an important analyte due to its notorious toxicity at relatively low concentrations to humans and ecosystems (Rice et al. 2014). In addition, the Hg^{2+} is easily methylated to organic mercury which has higher bioavailability and toxicity (Bernhoft 2012). According to the data published by FDA, during 1990-2012, the mercury concentration in commercial seafood, for example, tuna and swordfish, were already as high as around 2 and 3 mg/L respectively. All of these considerations make it desirable for the development of mercury sensors that are ultrasensitive to not only trace amount of mercury but also towards higher mercury levels. Various Hg^{2+}sensors have been developed using a wide variety of analytical methods in recent years (Table S1). Although some of them demonstrate superior sensitivity and selectivity towards Hg^{2+}, narrow dynamic ranges restrict their usage in many applications.

Nanomaterials offer unique opportunities in design and fabrication of ultrasensitive sensors for detection of mercury ions due to the high surface to bulk ratio and good electronic properties (Jiang et al. 2015). Among the vast array of available nanomaterials, single-walled carbon nanotubes (SWCNTs) have attracted substantial attention in fabrication of chemical sensors due to their high sensitivity (Schroeder et al. 2019; Zhao et al. 2020). Chemical functionalization of SWCNTs can further improve their sensitivity and selectivity. For example, gold nanoparticle (AuNP) decorated sensor was employed to detect biomolecules (Michael et al. 2017). While as-synthesized SWCNTs are mixtures of metallic and semiconducting nanotubes have been successfully used in fabrication of chemical sensors, semiconductor enriched SWCNTs (sc-SWCNTs) have been shown to improve FET device characteristics and the resulting sensor performance (Hwang et al. 2019; Shao et al. 2019). FET devices based on SWCNTs can provide more information about the nature of the chemical interaction with analyte compared to two probe resistive measurements (Heller et al. 2008). Combining FET characteristics with statistical learning methods leverages the interactions between sensor and analytes to further optimize the sensor system.

Inclusion of multiple characteristics of sensor signal is especially useful in diagnosing sensor behavior and mechanism. Statistical machine learning is a powerful tool to handle the multivariable modeling and interpretation. Combining biosensors with machine learning leverages the interactions between sensor and analytes to further optimize the sensor system (Cui et al. 2020; Ha et al. 2020). Statistical machine learning algorithms can be roughly divided into supervised learning, such as classification and regression, and unsupervised learning. In sensor research, a classification task usually refers to analyte detection, linear discriminant analysis (LDA) was used to detect liver fibrosis with polymer sensor arrays (Peveler et al. 2018), and support vector machine

(SVM) was employed as a classification model to identify purine derivatives in the mixture with metal nanoparticle-decorated SWCNT FET devices (Bian et al. 2019). Regression modeling, in the context of sensors, establish quantitative sensitivity in response to a concentration of analyte. Neural network modeling was employed along with copper nanoparticles functionalized polymer film to quantify maleic hydrazide (Sheng et al. 2019). Relevance vector machine enables the MXene nanoshuttles sensor to determine the ultra-trace amount of carbendazim in tea and rice (Zhu et al. 2021). Unsupervised learning such as principal component analysis (PCA) is also frequently used in chemical research to reduce noise and visualize the high dimensional data (Star et al. 2006).

Linear regression is a widely used statistical learning technique to calibrate sensors and determine their dynamic ranges, due to ease of implementation and interpretability of the results (Bratko 1997). However, as the relationship between the concentration and sensor characteristics take on more complicated functional dependence, linear regression diminishes in its efficacy. In general, many physical processes may be inadequately described in simple linear relationships over large ranges, ultimately requiring other types of functional correspondence to be established between input variables and output characteristics (Varmuza and Filzmoser 2009). While nanosensors have been demonstrated to possess ultra-low detection limits; ultrasensitive sensors typically have small dynamic ranges and are inherently more susceptible to saturation at higher concentrations of analytes (An et al. 2013; Wang et al. 2019). As a result, higher concentration ranges represent challenges during sensor calibration.

Herein, we demonstrate a machine learning-based methodology to tackle the nanosensor response saturation. Functionalized sc-SWCNTs FET based mercury ion (Hg^{2+}) sensors serve as an illustrative example of machine learning assistance in solving sensor response saturation. Our

device actually has two adsorption layers for analyte Hg^{2+} :1) the Hg^{2+} receptor thymine layer, captures the Hg^{2+} to generate the signals; 2) the other one is the gold nanoparticle layer, which can adsorb Hg^{2+} non-selectively. The low limit of detection is driven by sensor with receptors, when all receptors are saturated, detection of higher concentration is evident by the interaction with AuNPs, which was confirmed by control experiment using the sensors without receptors. The mercury sensor detects concentrations as low as 10 fM. But it exhibits reduced sensitivity in the expected region for concentrations higher than 10^{-8} M, the FET curves kept changing in a complicated, non-discernable way. To uncover the patterns buried under the unusual signals, we constructed a series of statistical models ranging from linear regression and non-linear model like regression spline to more flexible non-parametric models like trees and ensembles. Machine learning-based calibration improves dynamic range of the device to detect concentration ranging from 10^{-3} to 10^{-14} M. Construction of the statistical model also enables additional information regarding the sensing mechanism to be extracted from the dataset. This further insight highlights the advantage of statistical models in not only improving existing sensors but also providing guidance on how to improve individual sensor chemistry.

2. Results and Discussion

2.1. Sensor characterization, feature definition, calibration plot, saturation

Figure 1a shows a schematic illustration AuNPs/Cys/Thy-AcOH decorated sc-SWCNTs FET device for Hg^{2+} ion detection. FET device fabrication was conducted as follows: sc-SWCNTs were deposited from a toluene solution onto silicon device with prefabricated electrodes via dielectrophoresis (DEP). Gold nanoparticles were subsequently grown on sc-SWCNTs by electrodeposition (Shao et al. 2019). Cysteine (Cys) molecules were incubated with devices

attaching to gold nanoparticles via the Au-sulfur interaction. A similar synthetic approach was recently employed in hydroxyl radical label-free sensing using graphene-based FET devices, that had low nanomolar detection limit (Wang et al. 2019). Thymine-1-acetic acid (Thy-AcOH) molecules were coupled to the amine group of Cys via EDC/NHS mediated coupling. A scanning electron microscopy (SEM) micrograph of functionalized sc-SWCNTs is shown in Figure 1b. The size distribution of functionalized Au nanoparticles is 76.18 ± 19.43 nm (n = 26). Selectivity test (Figure S1b) and limit of detection $10^{-14.36 \pm 0.78}$ M were discussed in details in supplementary information.

Establishing the linear range of response for the devices was accomplished by taking the NTFET measurements in water as a baseline. We incubated devices with increasing concentrations of mercury solutions (mercury nitrate) from 10^{-14} M to 10^{-9} M to the sensors for a given time. Mercury solutions were washed off with the final FET measurements being taken in water to mitigate the effect of ionic strength on the measurement. NTFET curves with increasing concentrations of mercury ions are shown in Figure S1c. Among the FET features previously defined (Bian et al. 2019), the feature Var4 (normalized conductance change at –0.4 V) shows the most robust linearity with lower ranges of concentration as compared with using other features we define in Figure 1c. As the concentration of mercury further increases (Figure S1d), the changes of the FET curves begin to deviate from linearity (saturation). For example, the curves for each concentration start crossing each other, making discernment of patterns in increasing concentrations more difficult.

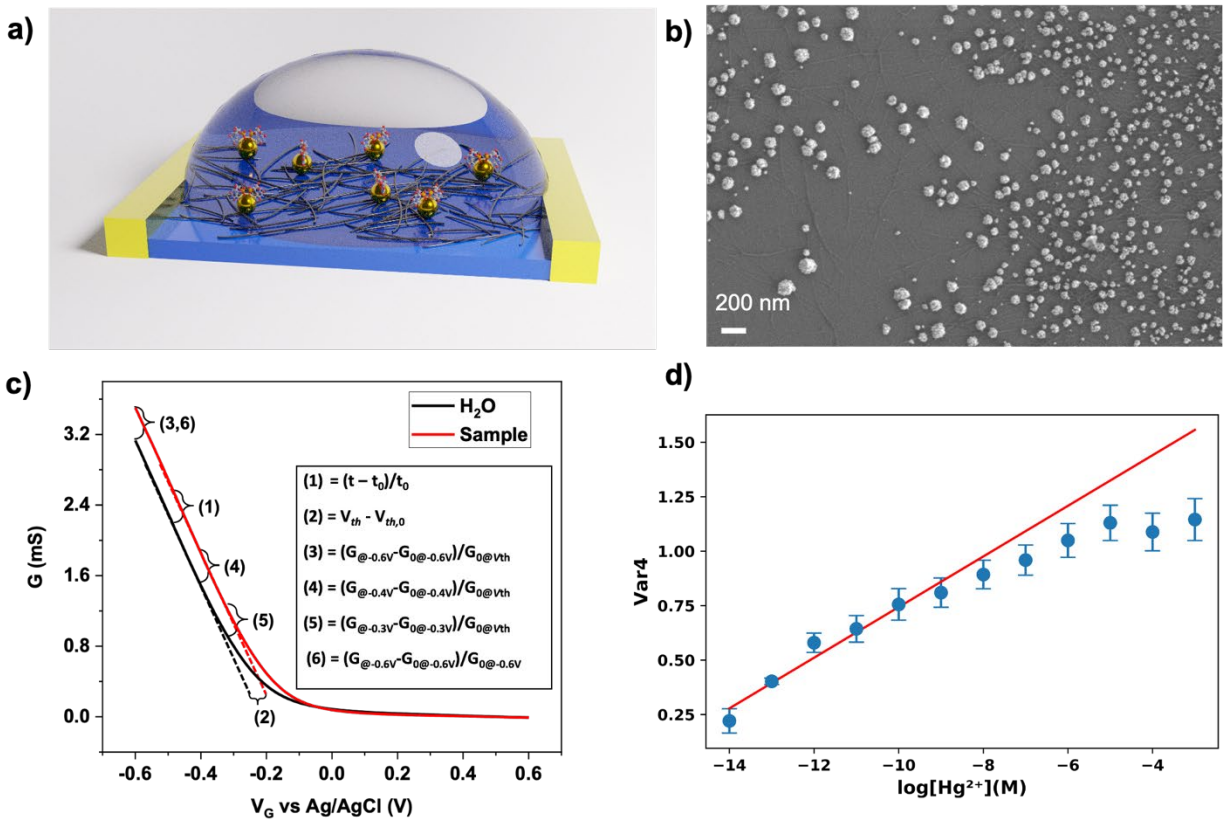


Figure 1. a) Schematic diagram for the composition of sensing material and liquid-gated carbon nanotube field-effect transistor (NTFET). Semiconductor enriched ingle-walled carbon nanotubes (sc-SWCNTs) were decorated with gold nanoparticles (Au NPs) and functionalized with cysteine (Cys) and thymine-1-acetic acid (Thy) for Hg^{2+} detection. b) Scanning electron microscope (SEM) image of a network of functionalized sc-SWCNTs. c) Definitions of the selected six features (Var1-6) and corresponding equations. d) Scatter plot of sensor response (Var4 was defined in panel c) for whole concentration range (blue scatters with error bars), error bar is from four devices on a single chip; Red straight line, ideal calibration plot for concentration range (10^{-14} M to 10^{-9} M).

The semi-log scatter plot (Figure 1d) using the data from four devices on a single chip shows a clear linear trend at lower concentration as expected, but the devices start showing decreased sensitivity for higher concentrations (10^{-8} - 10^{-3} M). The saturation behavior is

identified with lack of statistically significant changes in device response as the analyte concentration increased. We ran the one-way ANOVA test to compare the signal of each concentration with their neighbor's (Table S2). For example, for the comparison between the signals (normalized conductance at –0.4 V) of 10^{-8} M and 10^{-7} M, the p-value is around 0.15, which is much higher than the p-value 0.05. p-value 0.15 means it is 85% confident that the two sets of data are different. Increasing inability to distinguish between adjacent concentrations make the sensor unreliable at higher concentrations. Therefore, in Figure 1d, even though it seems like there exists a straight line from 10^{-12} to 10^{-5} from a single chip with four devices, considering the variance of the relative conductance between the groups in the whole dataset, there is no significant difference between the groups 10^{-8} M to 10^{-3} M. For example, given a relative conductance around 1, it's hard to tell whether it belongs to 10^{-8} M or 10^{-7} M.

Developing more robust device response to mercury ion concentrations correlations can be achieved through multivariate modeling. NTFET curves contains a wealth of information beyond simple comparisons of conductance value at a single gate voltage. The six defined features (Figure 1c), which accurately capture changes in FET transfer characteristics, span the predominant mechanisms identified in NTFET biosensors including electrostatic gating, charge transfer effects, changes in scattering sites, charge-carrier mobility, and modulation of Schottky barriers (Heller et al. 2008).

Since the lower concentration range shows a clear linear trend, we therefore fit the data with a linear model ($R^2 = 0.9974$) for concentration range (10^{-14} M to 10^{-9} M), as employing more flexible models on a clear linear pattern will unnecessarily capture the noise in the data, and potentially cause overfitting. For concentration range (10^{-8} M to 10^{-3} M), however, linear model no longer applies and more flexible machine learning needed to discover the pattern. For the

saturation region, we utilize the selected six features from six concentrations measured across 12 devices as input variables (Var1-6), the log concentrations as the responses (y). We are estimating a robust function that can best describe the relationship between response and variables, that is $\hat{y}_i = \hat{f}(x_i)$ such that the residual sum of square (RSS) $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is minimized.

2.2. Statistical modeling for sensor saturation

2.2.1. Linear regression and feature importance analysis

We started with a single variable linear regression on only one variable Var4, to observe how much model performance improves by including more information in the model. We use adjusted R^2 (R^2_{adj}) instead of R^2 as the metric to evaluate the model performance, and to compare between different models, as R^2_{adj} takes the dimension of the data into account to avoid the curse of dimensionality (Trunk 1979). The R^2_{adj} of single variable linear regression was 0.4225. After we included all six features to fit a multi-variable linear model, the R^2_{adj} was improved to 0.5816. Even though the linear model improves around 40%, the R^2_{adj} for multi-variable linear model is still improvable. To explore the causes of the poor model performance, we plotted the diagnostic plot. We can get a sense of what are missing in the model from the plot. Ideally, if our estimated model is close to the true model, a flat line close to 0 is expected, as the residuals are expected to be 0. However, instead of a flat line close to 0, we get curves in both single variable and multivariable linear models (Figure S6a and 6b), indicative of non-linearity in the data that linear model cannot capture.

The variable importance for linear model can be interpreted using p-value for each variable. p-value represents the probability of the null hypothesis is true. In linear regression, the null hypothesis is the coefficient of that variable equals zero. Table S3 summarizes the single variable

and multivariable linear model, and the p-values for each variable. Linear model indicates that Var1, Var3, and Var4 are significant to the model with 95% confidence, as the p-values are 0.0551, 0.0338, and 0.0503. The reason why both Var3 and Var4 are significant is that Var3 and Var4 are highly correlated to each other (correlation coefficient = 0.7516), meaning the Var3 might mask the importance of Var4 in linear model, and causes it to have a larger p-value.

2.2.2. regression splines and feature importance analysis

The poor performance of the linear model is caused by the inherent non-linearity in the data. To solve this problem, we considered polynomial terms, but the non-linearity is not just limited to fitting the whole dataset with polynomial order. A more flexible model, the regression spline, puts K knots into the whole data space, which ends up fitting K+1 cubic polynomials, for each of the piecewise polynomials, constraints are imposed to ensure the curve across the data space is smooth. A larger K will end up fitting a more flexible model, whereas K=0 will give a simple cubic polynomial. As we expected, by choosing optimal Ks for each variable via cross-validation, the R^2_{adj} for spline model is 0.8517. While taking non-linearity into account and including more variables drastically improve the model performance, the model still has room for improvement. As for the regression spline, it still assumes the splines are additive and solves for a coefficient for each spline similar to linear model.

To explore the spline model and gain some insights regarding the variable importance, as the model performance was boosted by 46% from linear model (from 0.5816 to 0.8517). To evaluate the variable importance for regression spline, we performed a permutation test. Variable significance in the model is evaluated by breaking the connections between that variable and the response. We then reconstruct the model with the permuted variable, and get the model

performance (R^2_{adj}). We iterate the process for 1000 times, and get the distributions of the R^2_{adj} for the models with permuted variable. If the original R^2_{adj} is from the distribution of the R^2_{adj} from permuted variable, then the null hypothesis is true, which is there is no significant connection between variable and the model performance. Otherwise, the alternative hypothesis is true, that is, variable is significant to the model.

The permutation test results for all 6 variables are summarized in Figure 2. The 95% confidence intervals for the permuted R^2_{adj} for each of the variable were summarized in Table S4. As we can see, the original R^2_{adj} is from the distribution of the permuted models for Var1, Var5, and Var6 with 95% confidence interval, which indicates these three variables are not significant to the spline model. The noticeable difference regarding the feature significance from linear model is that in this regression spline model, Var2 (the threshold voltage change) turns out to be significant, which shows Var2 is expressing the non-linearity in the data, which was able to be captured in the model. Both Var3 and Var4 are still significant in the regression spline due to the high correlation.

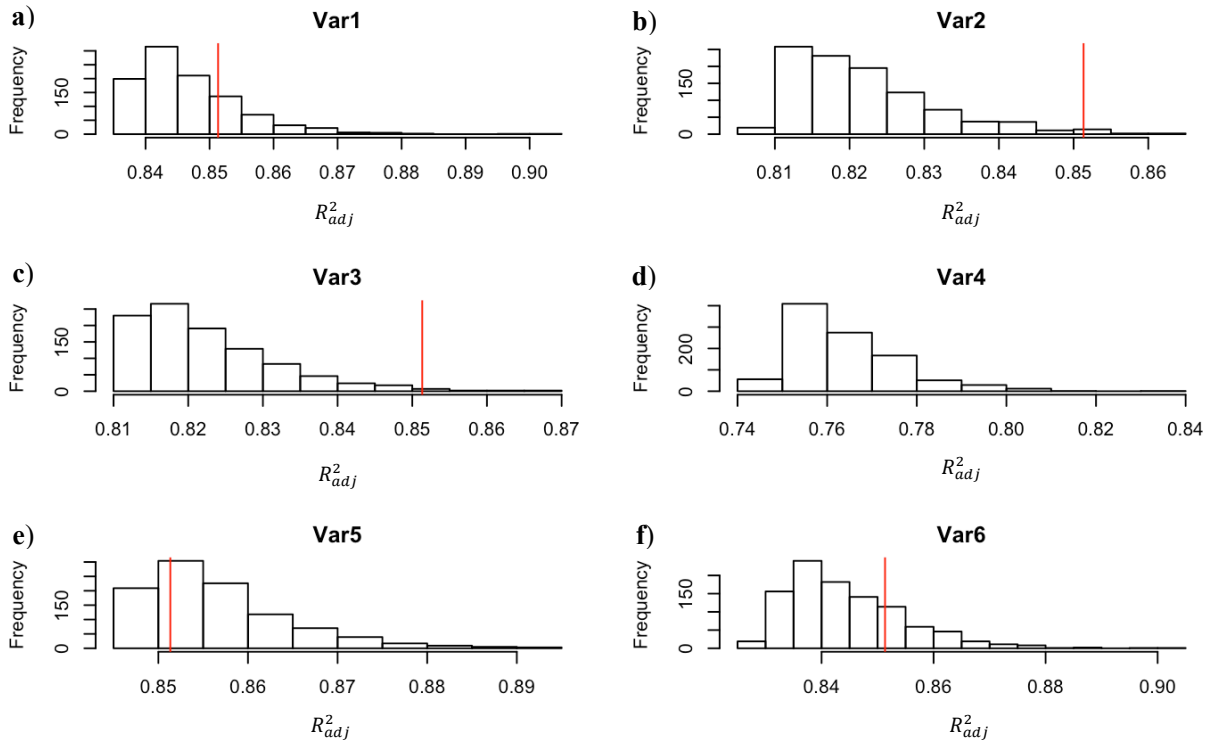


Figure 2. The distribution of the R^2_{adj} of models with permuted variables a) Var1; b) Var2; c) Var3; d) Var4; e) Var5; f) Var6. The red vertical line represents the original R^2_{adj} = 0.8516, which is out of x-scale in panel d.

2.2.3. Regression trees and feature importance analysis

As discussed above, the regression spline still has limitations. To overcome the limitations of spline model, we fit the data with a regression tree algorithm. Regression tree(James et al. 2014) is a more flexible, non-parametric model which does not rely on the parameters (coefficients) used in regression splines and linear model, and it does not assume the additive properties of the variables. Each split node of the tree gives a pair of half planes (R_1 and R_2), $R_1(j, s) = \{X \mid X_j < s\}$and $R_2(j, s) = \{X \mid X_j \geq s\}$, j is the index of the variable, s the spot where the split happens. The goal is to find the value of j and s, such as the sum of the RSS in both half planes are minimized. The advantages of regression tree (Figure 3a) include it does not rely on the assumptions discussed

above and it is highly interpretable, but the disadvantage is it tends to suffer from overfitting. As we expected, the R^2_{adj} of regression tree model is improved to 0.9247. The variable importance in regression tree was computed as Var4>Var5>Var2>Var3>Var1> Var6, the ranking of feature importance is consistent with regression splines, except for Var5. The inconsistency arises from how variable importance was computed (using RPart) as the weighted sum goodness of splits when it was a primary variable and surrogate variable, and Var4 and Var5 are highly correlated (r = 0.8771).

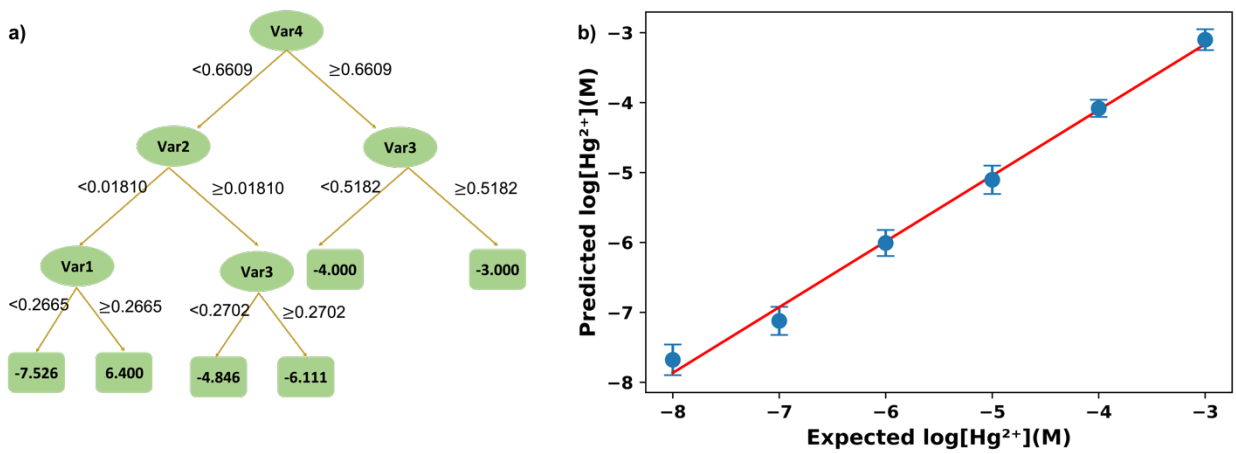


Figure 3. a) The structure of the regression tree. b) The comparison between the average of predicted value for each concentration by random forest and the expected concentrations, error bars are the standard deviations of the predicted values.

2.2.4. Random forest and feature importance analysis

We are trying to boost the model performance as close as possible to 1, even though performance of the tree model is adequate. To improve the model is instead of using one single tree, we can utilize more trees to form a random forest. Random forest (Breiman 2001) has a lower bias (susceptibility to idiosyncrasies of a particular dataset) because it forces each tree to be grown. It also lowers the variance of the whole model, as the prediction is the average of predictions from

each tree. In addition, the model variance is further reduced by decorrelating each of the tree in the forest. To prove the efficiency of random forest modeling, we averaged the predicted value of each concentration, and plot them against the expected concentrations (Figure 3b). The R^2 of the plot is 0.9974 with slope 0.9402, which not only shows the average predicted values are highly correlated to the expected values, but also demonstrated the average predicted values very close to the expected values. To further prove the saturation problem is solved, we ran the one-way ANOVA test on the prediction from random forest regression (Table S2). For example, the p-value is <0.01 for the comparison between the signals from 10^{-8} M and 10^{-7} M, a substantial improvement over the single variate prediction. The random forest model has an R^2_{adj} 0.9660, which is the best model we achieved so far, and the raw predicted values can be found in Figure S7. The model generalizability was evaluated using out-of-bag (OOB) error. The OOB is around 0.5 for 500 trees (Figure S8), and the corresponding R^2 0.8260.

Beyond accurately modeling the concentration, the structure of the model itself can be explored to identify important contributing variables to elucidate the mechanism and provide guidance for future improvements. Variable importance for random forest (Table 1) are evaluated using two metrics, out-of-bag error (OOB Err), and node purity (James et al. 2014). Both of the metrics show the consistent ranking: the largest contribution to the regression is Var4 (the conductance change at –0.4 V), next largest contribution is Var2 (change of threshold voltage). Making the model as parsimonious as possible is important to limit potential sources of error. By combining only these two features, we obtained the R^2 and adjusted R^2 0.9359 and 0.9292. The random forest regression allows us to accurately predict the concentration in the previously saturated range, while also identifying features are important to sensing. Therefore, our sensor and calibration model not only have an ultra-low detection limit (10^{-14} M), but also have a wide

dynamic range, which exceeds detection limits and dynamic ranges of the mercury sensors as summarized in Table S1.

So far, regarding the feature importance, we observe that, the variable Var2 becomes significant, as the model become more flexible and capture the non-linearity in the data. And variable Var4 is significant regardless the model flexibility. From the feature importance for higher concentration range (Table 1), in addition to the feature Var4, which we initially used to calibrate the concentration, one additional feature – the threshold voltage shift – has substantial contribution. Adding threshold voltage shift feature corrected the saturation at higher concentration range. In the feature importance ranking, features Var3 through Var6 are the relative conductance changes, among them, Var4 and Var5 are most correlated with the response with correlation coefficients r equal to 0.6569 and 0.7074. They correspond to an electrostatic gating effect, with the curves shift horizontally as more analytes were added (Figure S1c). Var1 and Var2 describe the curve bending, they are relative transconductance and threshold voltage change respectively, Var2 is more correlated to the response (r = 0.6352). The curve bending corresponds to the Schottky barrier mechanism. The Schottky barrier exists on the interfaces of AuNPs and SWCNTs, excess Hg^{2+} adsorbed onto AuNPs alters the Schottky barrier, which is shown by crossing of FET curves in Figure S1d.

2.2.5. Rationale for the selected models for low and high concentration calibration

We adapted a linear regression at low concentrations, because we already knew the relationship was linear, approaching a linear function with linear model is a better choice than using flexible model like random forest. In this situation, the unbiased and interpretable nature of a linear model were preferable to random forest.

At higher concentrations, the linear model is not capable to capture the complexity in the data (Figure S6). Linear regression model has higher mean square error (MSE) because the hypothesized true model is non-linear. Approximating the true model using linear regression would induce bias, in addition to the inherent variance in the linear model. Random forest is a very flexible model that captures the complexity within the data. Additionally, due to the large number and decorrelation of the trees in the forest, random forest has low variance. Random forest has a lower MSE, and better performance.

2.3. Feature importance and possible sensing mechanism

Improvement in sensing correlation leads us to assume threshold voltage shift was caused by the excess of Hg^{2+} adsorbed on the device when the receptor on the sensor is saturated by Hg^{2+}. Previous research has shown that Hg^{2+} tends to bind onto AuNPs surface, as it forms favorable Au-Hg NPs after applying the bias (Hezard et al. 2012; Li et al. 2014; Ojea-Jiménez et al. 2012; Ratner and Mandler 2015). Therefore, we could assume that the excess of Hg^{2+} absorbed onto AuNPs during the FET measurements.

Table 1. The variable importance for random forest model.

Variable	OOB Err	Node Purity
Var1	0.1122	7.1682
Var2	0.9971	48.0462
Var3	0.3087	14.7186

Var4	1.7097	66.2755
Var5	0.5366	26.0602
Var6	0.3281	14.9293

To test this hypothesis, we fabricated the Au NPs decorated sc-SWCNTs nanotube devices as control, measured sensor response for high concentration of the mercury analyte. From random forest analysis, we can see the threshold voltage change is the most prominent feature (Table 2 control sensor) in order to fit the data. From the correlation analysis (Table 3), the correlation coefficient between threshold voltage shift and concentration is significantly higher than other features. These analyses lead us to conclude that excess mercury cations bound to Au NPs induce threshold voltage shifts. From a mechanistic viewpoint, threshold voltage shift is induced by charge transfer (Heller et al. 2008), as direct contact between Hg^{2+} and Au NP transfers the charge more efficiently.

Table 2. The variable importance for random forest model based on the control device[a].

Variable	OOB Err	Node Purity
Var1	–0.0639	27.6138
Var2	0.7086	46.8411
Var3	0.1547	28.1684

Var4	0.1461	25.0609
Var5	0.2683	30.6461
Var6	0.1217	34.2395

[a] The control device has only sc-SWCNTs and Au NPs.

Table 3. Correlation coefficient for each of feature to the concentration.

Sensors	1	2	3	4	5	6
Control[a]	0.0113	–0.2251	–0.0264	0.0743	0.0603	–0.0192

[a] The control device has only sc-SWCNTs and Au NPs.

3. Conclusion

In conclusion, we demonstrated applying statistical learning models to successfully uncover the hidden, unexpected pattern in the saturated sensor signals, the calibrated saturated region has R^2 0.8260. The method is especially useful for biosensors that have multivariable signals, as signal change might not just happen at the expected region. Being able to mine the information outside the traditional signal region, not only allows us to utilize the sensor to its full potential, but also help us to investigate the sensing mechanism after saturation. The calibrated sensor is capable of detecting mercury cations as low as $10^{-14.36 \pm 0.78}$ M as well as concentrations as high as 10^{-3} M, which exceeds the performance of the previously published mercury sensors (Table S1).

Corresponding Author

*E-mail: astar@pitt.edu

Acknowledgements

We thank Characterization Facility (NFCF) at the University of Pittsburgh for access to the SEM instrumentation. We thank Dr. Jianfu Ding and Dr. Patrick Malenfant at the National Research Council of Canada for providing high-purity semiconducting single-walled carbon nanotubes.

Supporting Information

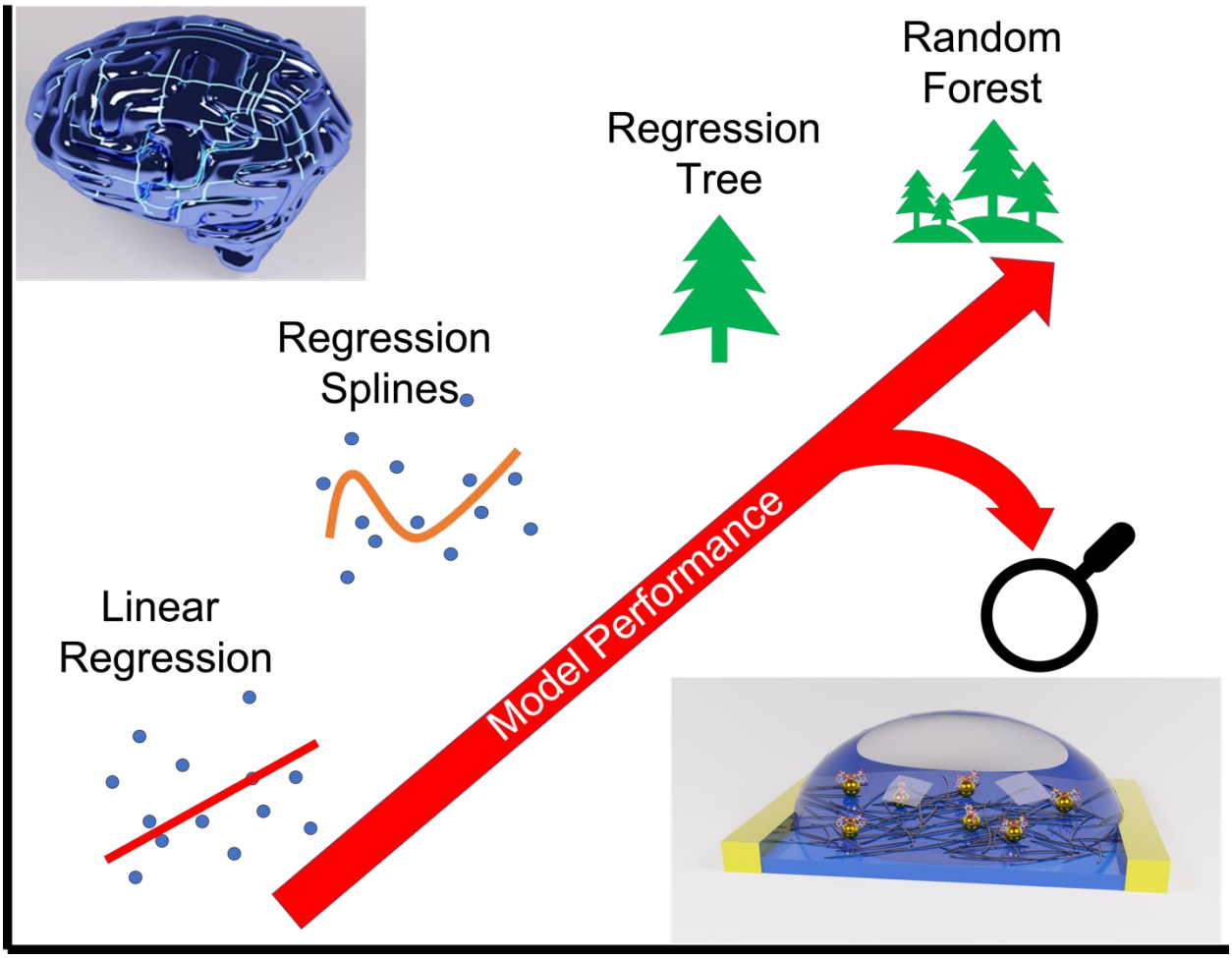
Supplementary material associated with this article can be found in the online version at

References:

Amjadi, M., Kyung, K.-U., Park, I., Sitti, M., 2016. Stretchable, skin-mountable, and wearable strain sensors and their potential applications: A review. Adv. Funct. Mater. 26(11), 1678-1698.
An, J.H., Park, S.J., Kwon, O.S., Bae, J., Jang, J., 2013. High-performance flexible graphene aptasensor for mercury detection in mussels. ACS Nano 7(12), 10563-10571.
Bandodkar, A.J., Jeang, W.J., Ghaffari, R., Rogers, J.A., 2019. Wearable sensors for biochemical sweat analysis. Annu. Rev. Anal. Chem. 12(1), 1-22.
Bariya, M., Nyein, H.Y.Y., Javey, A., 2018. Wearable sweat sensors. Nat. Electron. 1(3), 160-171.
Bernhoft, R.A., 2012. Mercury toxicity and treatment: A review of the literature. J. Environ. Public Health 2012, 460508.
Bian, L., Sorescu, D.C., Chen, L., White, D.L., Burkert, S.C., Khalifa, Y., Zhang, Z., Sejdic, E., Star, A., 2019. Machine-learning identification of the sensing descriptors relevant in molecular interactions with metal nanoparticle-decorated nanotube field-effect transistors. ACS Appl. Mater. Interfaces 11(1), 1219-1227.
Bratko, I., 1997. Machine learning: Between accuracy and interpretability. Learning, Networks and Statistics, pp. 163-177. Springer Vienna, Vienna.
Breiman, L., 2001. Random forests. Mach. Learn. 45(1), 5-32.
Cui, F., Yue, Y., Zhang, Y., Zhang, Z., Zhou, H.S., 2020. Advancing biosensors with machine learning. ACS Sens. 5(11), 3346-3364.
Ha, N., Xu, K., Ren, G., Mitchell, A., Ou, J.Z., 2020. Machine learning-enabled smart sensor systems. Adv. Intell. Syst. 2(9), 2000063.
Heller, I., Janssens, A.M., Männik, J., Minot, E.D., Lemay, S.G., Dekker, C., 2008. Identifying the mechanism of biosensing with carbon nanotube transistors. Nano Lett. 8(2), 591-595.
Hezard, T., Fajerwerg, K., Evrard, D., Collière, V., Behra, P., Gros, P., 2012. Gold nanoparticles electrodeposited on glassy carbon using cyclic voltammetry: Application to Hg(ii) trace analysis. J. Electroanal. Chem. 664, 46-52.
Hwang, S.I., Franconi, N.G., Rothfuss, M.A., Bocan, K.N., Bian, L., White, D.L., Burkert, S.C., Euler, R.W., Sopher, B.J., Vinay, M.L., Sejdic, E., Star, A., 2019. Tetrahydrocannabinol detection using semiconductor-enriched single-walled carbon nanotube chemiresistors. ACS Sens. 4(8), 2084-2093.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2014. An introduction to statistical learning: With applications in r. Springer Publishing Company, Incorporated.
Jiang, S., Cheng, R., Ng, R., Huang, Y., Duan, X., 2015. Highly sensitive detection of mercury(ii) ions with few-layer molybdenum disulfide. Nano Res. 8(1), 257-262.
Kim, J., Campbell, A.S., de Ávila, B.E.-F., Wang, J., 2019. Wearable biosensors for healthcare monitoring. Nat. Biotechnol. 37(4), 389-406.
Li, D., Li, J., Jia, X., Wang, E., 2014. Gold nanoparticles decorated carbon fiber mat as a novel sensing platform for sensitive detection of Hg(ii). Electrochem. Commun. 42, 30-33.
Liu, M., Du, P., Yu, C., He, Y., Zhang, H., Sun, X., Lin, H., Luo, Y., Xie, H., Guo, J., Tong, Y., Zhang, Q., Chen, L., Zhang, W., Li, X., Wang, X., 2018. Increases of total mercury and methylmercury releases from municipal sewage into environment in china and implications. Environmental Science & Technology 52(1), 124-134.

Michael, Z.P., Shao, W., Sorescu, D.C., Euler, R.W., Burkert, S.C., Star, A., 2017. Probing biomolecular interactions with gold nanoparticle-decorated single-walled carbon nanotubes. J. Phys. Chem. C 121(38), 20813-20820.
Ojea-Jiménez, I., López, X., Arbiol, J., Puntes, V., 2012. Citrate-coated gold nanoparticles as smart scavengers for mercury(ii) removal from polluted waters. ACS Nano 6(3), 2253-2260.
Peveler, W.J., Landis, R.F., Yazdani, M., Day, J.W., Modi, R., Carmalt, C.J., Rosenberg, W.M., Rotello, V.M., 2018. A rapid and robust diagnostic for liver fibrosis using a multichannel polymer sensor array. Adv. Mater. 30(28), 1800634.
Ratner, N., Mandler, D., 2015. Electrochemical detection of low concentrations of mercury in water using gold nanoparticles. Anal. Chem. 87(10), 5148-5155.
Rice, K.M., Walker, E.M., Wu, M., Gillette, C., Blough, E.R., 2014. Environmental mercury and its toxic effects. J Prev Med Public Health 47(2), 74-83.
Schroeder, V., Savagatrup, S., He, M., Lin, S., Swager, T.M., 2019. Carbon nanotube chemical sensors. Chem. Rev. 119(1), 599-663.
Shao, W., Burkert, S.C., White, D.L., Scott, V.L., Ding, J., Li, Z., Ouyang, J., Lapointe, F., Malenfant, P.R.L., Islam, K., Star, A., 2019. Probing Ca^{2+}-induced conformational change of calmodulin with gold nanoparticle-decorated single-walled carbon nanotube field-effect transistors. Nanoscale 11(28), 13397-13406.
Sheng, Y., Qian, W., Huang, J., Wu, B., Yang, J., Xue, T., Ge, Y., Wen, Y., 2019. Electrochemical detection combined with machine learning for intelligent sensing of maleic hydrazide by using carboxylated PEDOT modified with copper nanoparticles. Microchim. Acta 186(8), 543.
Star, A., Joshi, V., Skarupo, S., Thomas, D., Gabriel, J.-C.P., 2006. Gas sensor array based on metal-decorated carbon nanotubes. J. Phys. Chem. B 110(42), 21014-21020.
Trunk, G.V., 1979. A problem of dimensionality: A simple example. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-1(3), 306-307.
Varmuza, K., Filzmoser, P., 2009. Introduction to multivariate statistical analysis in chemometrics, 1st ed. CRC Press, Boca Raton.
Wang, Z., Yi, K., Lin, Q., Yang, L., Chen, X., Chen, H., Liu, Y., Wei, D., 2019. Free radical sensors based on inner-cutting graphene field-effect transistors. Nat. Comm. 10(1), 1544.
Zhao, M., Chen, Y., Wang, K., Zhang, Z., Streit, J.K., Fagan, J.A., Tang, J., Zheng, M., Yang, C., Zhu, Z., Sun, W., 2020. DNA-directed nanofabrication of high-performance carbon nanotube field-effect transistors. Science 368(6493), 878.
Zhu, X., Liu, P., Xue, T., Ge, Y., Ai, S., Sheng, Y., Wu, R., Xu, L., Tang, K., Wen, Y., 2021. A novel graphene-like titanium carbide MXene/Au–Ag nanoshuttles bifunctional nanosensor for electrochemical and SERS intelligent analysis of ultra-trace carbendazim coupled with machine learning. Ceram. Int. 47(1), 173-184.

TOC graphic:


Random
Forest
Regression
Tree
Regression
Splines
Linear
Regression
Model Performance