Electrochimica Acta 491 (2024) 144304

Contents lists available at ScienceDirect

Electrochimica Acta

journal homepage: www.journals.elsevier.com/electrochimica-acta

Automatic simulation of electrochemical sensors by machine learning for drugs quantification

Lin Du [a,*], Yann Thoma [b], Francesca Rodino [a], Sandro Carrara [a]

[a] Bio/CMOS Interfaces Laboratory (BCI), École Polytechnique Fédérale de Lausanne, 1011 Lausanne, Switzerland
[b] School of Engineering and Management Vaud, HES-SO University of Applied Sciences and Arts Western Switzerland, 1401 Yverdon-les-Bains, Switzerland

ARTICLE INFO

Keywords:
Therapeutic drug monitoring
Machine learning
Cyclic voltammogram
Simulated dataset generation
Drug concentration quantification

ABSTRACT

Multiple drug concentrations concurrent detection and quantification based on electrochemical sensors are of great importance for therapeutic drug monitoring (TDM) and the development of personalized therapy. Cyclic voltammogram (CV) results obtained by electrochemical sensors can be used to offer quantitative information about drug concentrations. Several approaches have been proposed for single-concentration quantification based on CV results and machine learning (ML) models with lower training difficulty. However, insufficient measured dataset hinders the application of diverse large-scale ML algorithms in this field. A new method for automatic parameter estimation by ML of measured CV samples is here illustrated with the aim to generate a large simulated dataset for generalized drug concentration quantification model training in this paper. We present an ML-based approach that combines k-means clustering, polynomial regression, and Gaussian Mixture Model (GMM), which automates parameter estimation and simulation using peak detection, baseline subtraction, and peak Gaussian fitting with a small number of measurements. Large simulation datasets constructed on the basis of the estimation results open the possibility of training ML models for more generalized drug concentration quantification. The simulated dataset is processed to assess the efficiency of the proposed method. The Mean Average Percentage Error (MAPE) was 0.32% for etoposide (ETO) and 4.78% for methotrexate (MTX).

1. Introduction

Oncology is currently one of the main fields of clinical research, with chemotherapy being a commonly prescribed treatment method [1]. Compared with each single therapeutic agent, therapeutic cocktails (i.e. chemotherapeutic therapies incorporating multiple drugs) are more effective and efficient. However, many drugs (e.g., methotrexate, cyclophosphamide, ifosfamide, and etoposide) in chemotherapeutic therapies show non-negligible inter-individual and intra-individual variability with the combined impact of different genetic predisposition, variable surrounding environment and dynamic daily metabolism [2]. Successful application of personalized therapy is of great importance in addressing the problem of poor efficacy or risk of toxicity of current pharmacological treatments by optimizing the therapy for different patients with different characteristics [3]. Specifically, therapeutic drug monitoring (TDM) [4] of chemotherapeutic drugs, i.e. the individualized adjustment of drug dosage based on drug concentrations detection and quantification, aims at preventing toxicity due to overexposure and low treatment efficacy due to insufficiency, and is a significant component to achieve clinical practice.

Electrochemical sensors [5], with the characteristic of high sensitivity and low cost [6], have been employed in monitoring chemotherapeutic drugs. Especially, the cyclic voltammograms (CV) (Fig. 1) obtained by cyclic voltammetry techniques can be exploited to analyze drug quantitative information by the corresponding peculiar characterization of distinct redox peaks [7,8]. Namely, the effective application of CV results obtained by drug with different concentration levels open the possibility to track and predict the corresponding variation and thus become a reliable measurement data source for TDM implementation. However, several anticancer drugs have strong interactions. They are not only observed in patient metabolism, but also happen on electrochemical sensors, which reduce the selectivity of electrochemical methods in discriminating on drug cocktails and make subsequent processing suffer the risk of errors. Consequently, the quantification of drug concentrations obtained by CV results becomes more difficult to be acquired, requiring drug interactions to be taken into consideration [9]. Also, the nonlinear characteristics between CV results and corresponding drug concentrations need to be extracted using appropriate machine learning (ML) algorithms.

* Corresponding author.
E-mail address: lin.du@epfl.ch (L. Du).

https://doi.org/10.1016/j.electacta.2024.144304
Received 13 October 2023; Received in revised form 18 April 2024; Accepted 19 April 2024
Available online 20 April 2024

Various ML-based methodologies have been used to achieve one single drug concentration quantification based on CV results. For example, fouling phenomenon compensation and continuous concentration monitoring of propofol are achieved by support vector classifier (SVC) with an RBF kernel [10]. As well, the detection of caffeine and chlorogenic acid has been obtained by using Artificial Neural Network (ANN) to extract information directly by cyclic voltammetries with electrochemical sensors [11]. The ANN-based algorithm also takes advantage of the selected features to classify the voltammograms for determining the amount of potassium ferricyanide in the solution [12]. Similarly, we can then imagine to apply ML algorithms to extract information about anti-cancer compounds in data from electrochemical sensors. However, when the predicted target contains concentrations of multiple drugs with interactions between the drugs, the first challenge before selecting a suitable predictive model to achieve the quantification is to build a dataset that can be used to train the model. The literature also presents methods to predict drug cocktail effects when the possible number of measurements is fairly limited [13]. Nevertheless, when the number of experiments increases exponentially with the number of drugs and concentrations, effective datasets for training and validating numerous alternative ML models are still difficult to be obtained physically. A reasonable simulated dataset capable of modelling the CV results containing the drugs interaction for ML-based model training is therefore necessary. A mathematical estimation has been demonstrated in the literature [9], which can be used as the foundation to build large simulated datasets for ML model training. However, multiple parameters in this equation are affected by experimental setup and environment, which cannot be settled down accurately in absence of the information about these external variables.

This paper introduces and discusses a method to achieve automatic parameter estimation of measured CV samples by ML algorithms to build the foundation to train ML systems for generalized drug concentration quantification. Our method achieves automatic transformation from small sets of CV result samples to large simulation datasets through parameters estimation based on ML algorithms, which can be used for training and validation of ML models for concurrent quantification in multi-drug concentrations with interaction. The CV results obtained by Etoposide (ETO) and Methotrexate (MTX) with interaction are used for model validation.

2. Machine learning for CV results processing

Section 2.1 details the equipment used in the experiment for data collection about drugs detection, along with the procedure and obtained CV results. Section 2.2 shows the flowchart of the proposed model based on ML algorithms for CV results processing.

2.1. Data collection

ETO [14] and MTX [15] are the pharmaceutical anti-cancer agents used in this paper as model drugs. They are usually categorized as a topoisomerase II inhibitor and a folate antagonist, respectively, and are commonly employed in chemotherapy. Simultaneous detection of ETO and MTX using single electrochemical sensors based on carbon-based working electrodes would result in a chemical interaction [16]. Specifically, the presence of a higher concentration of ETO inhibits the oxidation process of MTX, which can be observed in the corresponding CV results. To evaluate the interaction between these two drugs, testing samples in different concentration groups were prepared. With the MTX concentration fixed at 40 μM, the concentration of ETO was increased from 5 μM to 30 μM. Owing to their limited water solubility, ETO and MTX were dissolved in dimethyl sulfoxide (DMSO) first. Phosphate-Buffered Saline (PBS) 0.1M worked as a supporting electrolyte, and the pH was adjusted to 7.4 by the addition of 0.1M NaOH. All of the aforementioned chemicals were purchased from Sigma-Aldrich®.

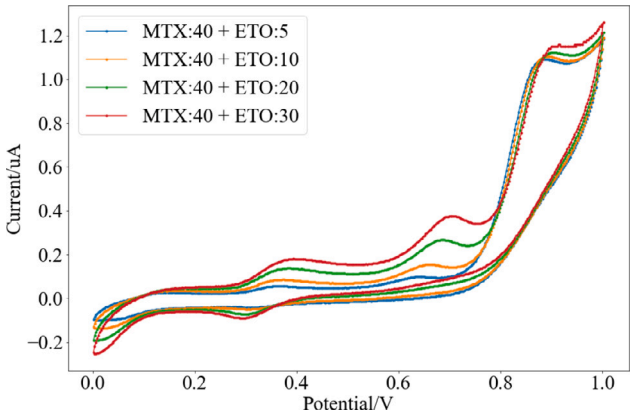


Fig. 1. CV results from ETO detection in the range of concentrations from 5 μM to 30 μM, in samples also containing 40 μM of MTX.

The screen-printed carbon electrodes (SPCEs) DRP-11L purchased from DropSens are used in the experiment without modifications. They are constructed by a carbon working electrode (circular shape, 4 mm diameter), a carbon counter electrode, and an Ag/AgCl pseudo-reference electrode. Under aerobic conditions, the CV results, whose potential was swept between 0 and 1 V with a scan rate of 20 mV/s, were obtained by drop-casting 100 μL of the testing sample, containing the drugs in PBS 0.1M, on the SPCE.

Fig. 1 presents the published measured CV results obtained from the drug cocktail consisting of ETO and MTX [16]. The figure shows the three anodic peaks on the oxidation side and one cathodic peak on the reduction side, for a drug cocktail with a fixed MTX concentration of 40 μM and an increasing concentration of ETO from 5 μM to 30 μM. Such voltammogram corresponds to the input of the model proposed to accomplish validation.

2.2. Model flowchart

Compared with the single application in the quantification [10,12,17], different ML algorithms are used in the different steps of the whole procedure in this paper for CV results processing to achieve both large simulation dataset construction and multiple drug concentrations quantification.

The flowchart of the proposed method to achieve automatic simulation of electrochemical sensors by ML for drugs quantification is demonstrated in Fig. 2. Leveraging on “Parameter Extractor”, which consists of peak detection, baseline subtraction, and peak gaussian fitting with different ML algorithms implementation, the parameters related to the mathematical estimation of measured CV results can be obtained automatically and are fed to the “Dataset Generator” for large simulated dataset generation. The quantification section is achieved by the “Convertor” trained by simulated dataset for quantifying multiple drug concentrations. The proposed method provides a powerful approach for dataset construction based on small dataset of measured samples, which opens the possibility to achieve ML model training and concentration quantification with the limitation of measured data.

3. Methods

In this section, the three main steps of the proposed method based on ML algorithms for automatic simulation are comprehensively described: peak detection (Section 3.1), baseline subtraction (Section 3.2) and peak Gaussian fitting (Section 3.3). Section 3.4 presents the process to achieve simulated dataset generation based on the proposed method.

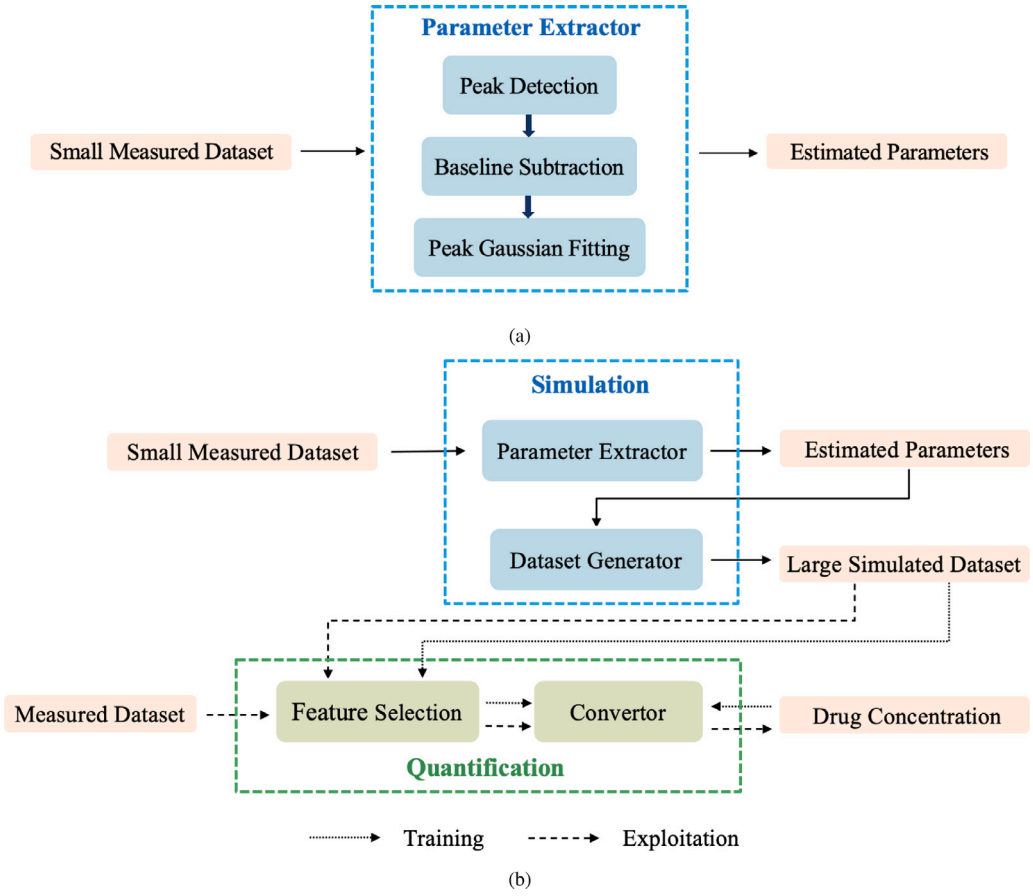


Fig. 2. The flowchart of the proposed method. "Parameter Extractor" (a) is constructed by peak detection, baseline subtraction, and peak Gaussian fitting for parameter estimation. The complete process (b) to build a generalized concentration quantification model based on several measured CV samples, can be divided into simulation and quantification, where the "Parameter Extractor" works as the main component in the simulation. The small measured dataset is used by the "Parameter Extractor" and the "Dataset Generator" for parameter estimation and large simulated dataset generation for model training. Both the measured dataset and the aforementioned simulated data can be exploited as the input of the quantification of drug concentration.

Table 1
Gradient (g) change analysis.

Type 1	Type 2	Start stage
①: $g > 0,\ g \uparrow$ ②: $g > 0,\ g \downarrow$ ③: $g < 0,\ g \downarrow$ ④: $g < 0,\ g \uparrow$	⑤: $g > 0,\ g \uparrow$ ⑥: $g > 0,\ g \downarrow$	⑦: $g > 0,\ g \downarrow$

↑ : Increasing Gradient.
↓ : Descending Gradient.

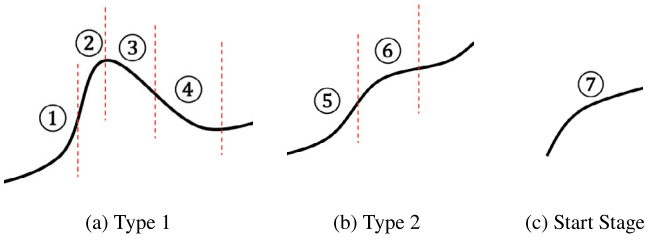


Fig. 3. Different types of peaks we found in CV results. Type 1 (a) first increases and then decreases. Type 2 (b) always shows increasing trend among its potential range. The CV results for both the oxidation and reduction sides contain the start stage (c) of increasing values but decreasing gradients.

3.1. Peak detection

Peak detection is based on the analysis of the gradient changes of the CV results in order to identify the key points for different types of peaks, namely, to find the corresponding potential to start the baseline subtraction and approximate potential range of each peak, thus transforming peak detection into a classification problem that can be solved using ML algorithms.

In detail, the shape of the anodic and cathodic peaks in the CV results caused by redox can be divided into two categories (3a, 3b), and the common characteristic of the two different peak types is that the initial stage and the cutoff stage include an increasing gradient phase (① and ⑤ in Table 1) and a descending gradient phase respectively (③ and ⑥ in Table 1). Simultaneously, the transition from the start stage (3) to any type of peak is essentially a transition from the gradient descending phase to the gradient ascending phase. In brief, peak detection of CV results is translated into an unsupervised classification problem by discriminating the articulation of gradient change.

After moving average-based denoising and gradient calculation, the dataset can be constructed for peak detection. Let $\boldsymbol{G} = \{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n\}$ with $\boldsymbol{g}_i \in \mathbb{R}^d$ denote the sequence including d adjacent gradient values of n samples as the feature vector. Subsequently, k-Means [18] is prescribed to achieve the aforementioned classification, which is a powerful unsupervised ML clustering algorithm with small computation cost and the requirement of very small number of measured samples. The objective of the proposed model is to construct a two-class classifier to separate the gradient increasing phase and gradient decreasing phase of input CV results for further peak detection. Let $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1$ with $\boldsymbol{\mu}_j \in \mathbb{R}^d$

denote two centroids for two target classes, which initial value are set as two arithmetic sequences with the same length as the feature vector in the dataset in order to eliminate the influence of randomly selected values on the accuracy of the classification results. Let $C_1, \ldots, C_n$ with $C_i \in \{0, 1\}$ denote the cluster assignment. The assignment step of the training process can be constructed as:

$$C_i = \arg\min_{j=0,1} \|\boldsymbol{g}_i - \boldsymbol{\mu}_j\|^2 \tag{1}$$

And the update step of the training process can be expressed as:

$$\boldsymbol{\mu}_j = \frac{\sum_{i=1}^{n} w_{ij}\boldsymbol{g}_i}{\sum_{i=1}^{n} w_{ij}} \tag{2}$$

where:

$$w_{ij} = \begin{cases} 1 & \text{if } C_i = j \\ 0 & \text{else} \end{cases} \tag{3}$$

$$J(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_n) = \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{g}_i - \boldsymbol{\mu}_{C_i}\|^2 \tag{4}$$

Therefore, the optimal classification is achieved by alternating the computation of the assignment step and the update step until convergence, namely, the cost function (Eq. (4)) reaches the minimum. Based on the classification results, two peak types and baseline fitting starting points can be identified, and then the cutoff position of the peaks of Type1 can be adjusted through gradient calculation results, thus achieving peak detection of CV results.

3.2. Baseline subtraction

Samples at the ends of the potential coverage range for each peak are selected for fitting the baseline for the oxidation side and reduction side respectively, namely, to use gradient descent to find the optimal parameter combination separately to minimize Mean Squared Error (MSE), as follows:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2N}\sum_{k=1}^{N}(h_{\boldsymbol{\theta}}(\boldsymbol{v}^k) - i^k)^2 \tag{5}$$

where h_θ is the objective polynomial, θ is the optimal parameter result for polynomial regression, N is the number of samples for baseline fitting, i^k and $\boldsymbol{v}^k$ are the current and the feature vector related to the potential corresponding to the current respectively. The R^2 score is used as the evaluation indicator to select the optimal baseline fitting result automatically from multiple polynomial fitting results of different orders for baseline subtraction and further redox peak current curve extraction, and the example of the peak current curve after baseline subtraction is shown in Fig. 4(b).

3.3. Improved peak Gaussian fitting model

Moving forward to the aforementioned estimation model [9], we propose here an improved peak Gaussian fitting model for the mathematical estimation of the extracted peak curve, where the redox peak extraction result can be constructed as a combination of multiple Gaussian components taking drug interactions into consideration, as follows:

$$i(V) = \sum_{\forall p} \Gamma_p C_p A_{p0} \cdot \{\prod_{\forall q \neq p}[1 \pm \frac{\varepsilon_q V_{max} C_q}{\upsilon_{max}(K_M + C_q)}]\} e^{-\frac{(V-V_p)^2}{2\sigma_p^2}} \tag{6}$$

where C_p and C_q are the concentrations of the drug and its interfering drug, V is the potential. The influence of its interfering drug on its own redox peak is expressed in Eq. (6) as a multiplicative term in braces which can affect the original peak current with the change of the concentration of interfering drug. Different operation symbols represent activation case and inhibition case of drug interactions respectively. A_{p0} is the original amplitude of each Gaussian component, ε_q is the interference strength, υ_{max} is the maximum reaction rate, V_{max} and K_M are the usual kinetic parameters for maximum velocity and substrate concentration at half maximum velocity [19], and V_p and σ_p are the position and width of each Gaussian component. Γ_p, which is related to the experimental setting and the diffusion coefficient D_p of the drug, is defined as:

$$\Gamma_p = 0.4463(\frac{n^3 F^3 \upsilon D_p}{RT})^{1/2} A \tag{7}$$

where n, F, υ, R, T, and A represent the number of electrons involved in the redox reaction, the Faraday constant, the scan rate, the gas constant, the temperature and the electrochemical active area of the working electrode respectively.

Combined with some adjustments, Gaussian Mixture Model (GMM) [20], a capable unsupervised ML algorithm used for fitting distributions, is transformed into a required regression-based parameter estimation method. Let $\boldsymbol{V} = \{v_1, \ldots, v_n\} \in \mathbb{R}$ denote the potential sequence of n samples whose probability distribution's shape is similar to the shape of the extracted peak current curve after baseline subtraction. The objective of the peak Gaussian fitting is to make the probability distribution of $\boldsymbol{V}$ approximate to weighted combination of Gaussian components equally with the number of peaks detected in Section 3.2, as follows:

$$P(\boldsymbol{X}|\boldsymbol{\Theta}) = \sum_{m=1}^{M} w_m \mathcal{N}(\boldsymbol{X}; \mu_m, \sigma_m^2) \tag{8}$$

where $\boldsymbol{X}$ is the sequence with the objective probability distribution, M is the number of the peaks, w_m is the mixing coefficient $\in \mathbb{R}$ with $\sum_{m=1}^{M} w_m = 1$, $\mu_m \in \mathbb{R}$ and $\sigma_m^2 \in \mathbb{R}$ are the mean and the variance of Gaussian component m respectively. Due to the hidden variables w_m, the process of finding the optimal approximation is completed by an Expectation–Maximization (EM) algorithm [20]. After performing the amplitude adjustment A_m, the approximation of the required parameters in Eq. (6) under the current experimental setup can be obtained according to the position correspondences and the calibration curve fitting, i.e., automatic parameter estimation is realized by peak Gaussian fitting. The correspondences of the redox peak influenced by drug interaction are as follows:

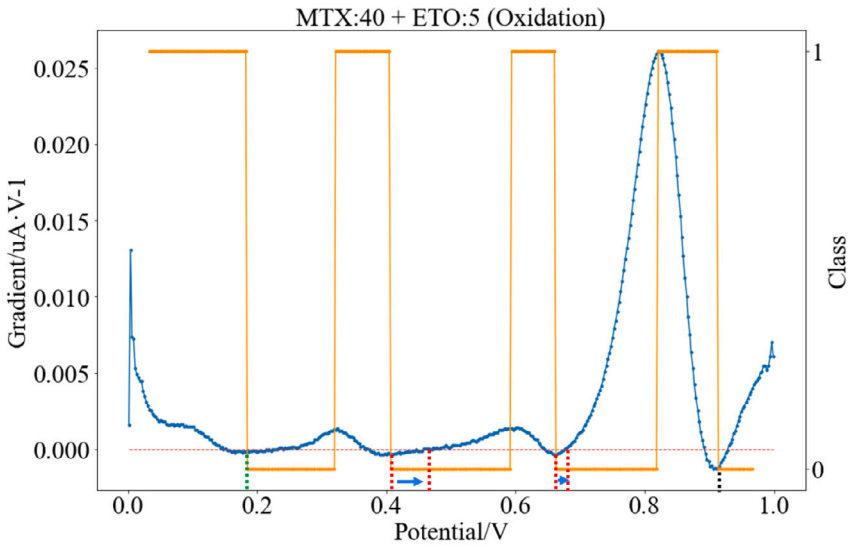
$$\begin{cases} \sigma_m \rightarrow \sigma_p \\ \mu_m \rightarrow V_p \\ A_m w_m \rightarrow \Gamma_p C_p A_{p0}\{\prod_{\forall q \neq p}(1 \pm \frac{\varepsilon_q V_{max} C_q}{\upsilon_{max}(K_M + C_q)})\} \end{cases} \tag{9}$$

3.4. Dataset generation

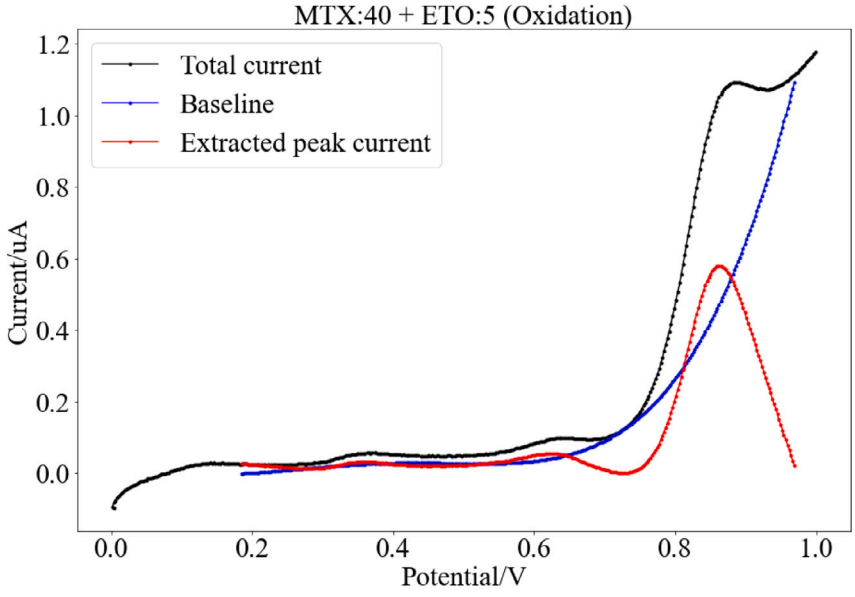
Based on the automatic parameter estimation result, the calculation of the current from Eq. (6) becomes a function upon the drug concentrations of interfering drugs. Such concentrations are considered here as independent variables, by opening the possibility to achieve large simulated dataset generation. Simultaneously, to deduce the impact of noise from multiple sources, the margins between linear regression-based fitting results and multiple sets of parameter estimation results with different concentration groups are used as the reasonable range of each required parameter, thus enabling the construction of large simulated dataset based on a small quantity of measured CV samples.

4. Results

Section 4.1 presents the performance of the main steps of the method proposed to build large simulated datasets based on automatic parameter estimation. Several experiments and analyses are implemented in Section 4.2 to evaluate the effectiveness of our proposed method for generalized drug concentration quantification in the simulated dataset.



(a)



(b)

Fig. 4. Automatic peak detection. (a) Gradient variations of each sample of the CV result are classified with Class = 0 and Class = 1 representing increasing and decreasing respectively. (b) The extracted peak current is obtained after baseline subtraction.

4.1. Automatic Gaussian fitting for simulation

Taking the first concentration group (MTX: 40 μM, ETO: 5 μM) as an example, Fig. 4 shows that peak detection and extraction are accomplished by the proposed method based on the k-Means algorithm and polynomial regression. It first achieves the classification of the gradient changing trends of each sample on the CV result into two classes (Fig. 4(a)), thus automatically locating four positions of the switch from descending status to ascending status, namely, approximate positions of articulation between redox peaks. According to automatic peak types analysis, the articulation positions of Type 1 are kept (Green dashed line), and the articulation positions of Type 2 are adjusted to the more reasonable positions based on the gradient (Red dashed line). The last possible articulation at around 0.9 V (Black dashed line) has been deleted, because its subsequent curves no longer meet the requirements to form a new redox peak. Consequently, three anodic peaks are detected automatically and the peak current curve is extracted after baseline subtraction (Fig. 4(b)).

Fig. 5 illustrates the result of the peak Gaussian fitting based on the GMM algorithm. The same number of Gaussian components as the redox peak number obtained from peak detection are fitted to approximate the extracted peak current of oxidation (Fig. 5(a)) and reduction (Fig. 5(b)) respectively. Automatic parameter estimation is then accomplished by extracting the parameters associated with each Gaussian component.

In order to generate the simulated data for model training and validation, the parameter estimation results of three sets (C_{MTX}, C_{ETO} = [(40 μM, 5 μM), (40 μM, 10 μM), (40 μM, 30 μM)]) and drug interaction described in [16] are used to solve equation sets and perform curve fitting in order to calculate reasonable fluctuation ranges for amplitude, mean and variance with uniform distribution of each Gaussian component that constitutes the peak current and unknown parameters value. Then, a large simulated dataset (100,000 sets totally) is established within the concentration range (5 μM, 45 μM). Fig. 6 presents the comparison between the measured result and examples in simulated dataset when the concentrations of MTX and ETO are 40 μM and

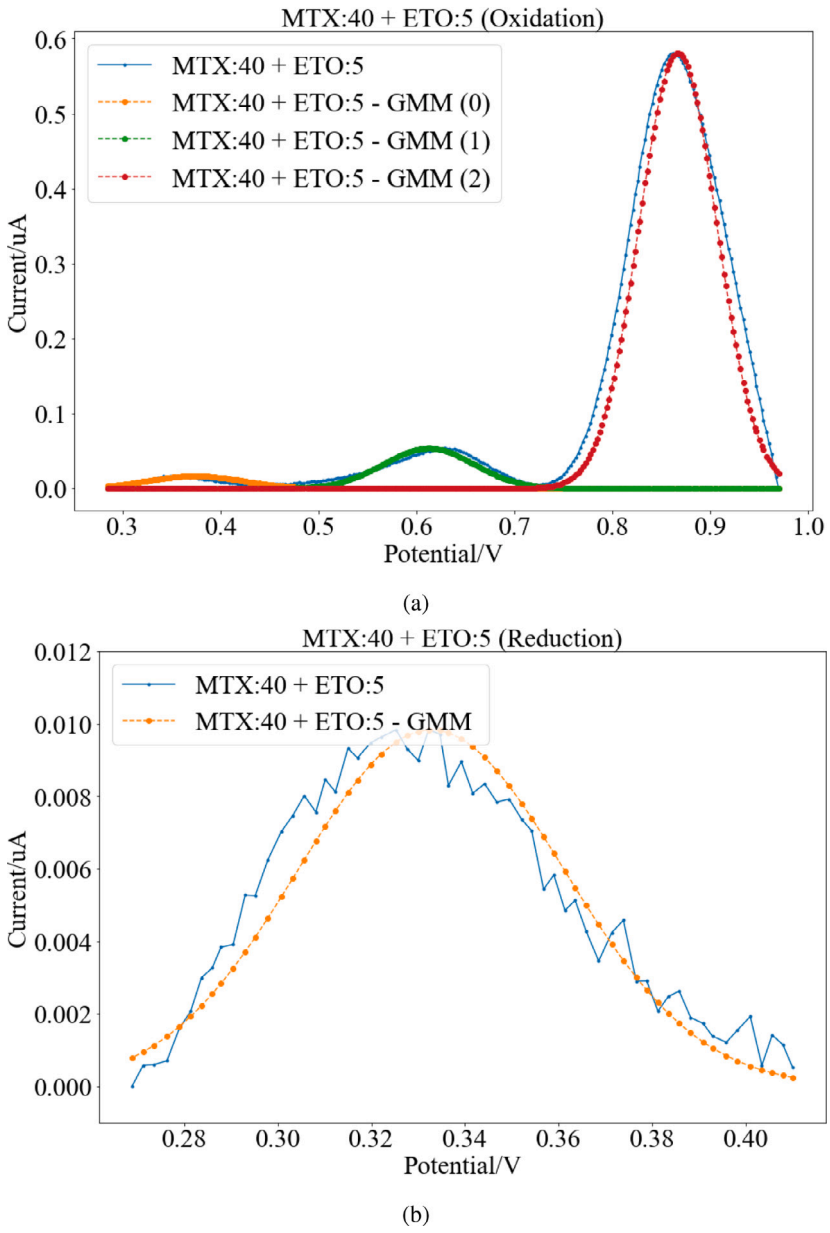


Fig. 5. Peak Gaussian fitting results based on the GMM algorithm. When the sample contains 40 μM MTX and 5 μM ETO, the peak Gaussian fitting results of oxidation side (a) and reduction side (b) are obtained.

5 μM respectively. In order to consider the existence of instability, noise, and other influencing factors (such as the fouling phenomena mentioned in the literature [10]) in the measured data, the values of each parameter will be randomly selected within the corresponding range. Therefore, for a fixed concentration group, the generated data will have some deviation from the sample data used to estimate the aforementioned parameters, which can simulate the CV results obtained in small experimental environment changes and different scanning times and be used to train more stable models.

4.2. Drug concentration quantification

To verify the reliability of automatically generated datasets in the "Simulation" (Fig. 2(b)), a simple and powerful Multi-Layer Perceptron (MLP) [20] with one hidden layer of 16 neurons is used to achieve generalized drug concentration quantification. For each generated CV result in the training set and test set, amplitude and variance of each redox peak are extracted as the features to build the feature vector

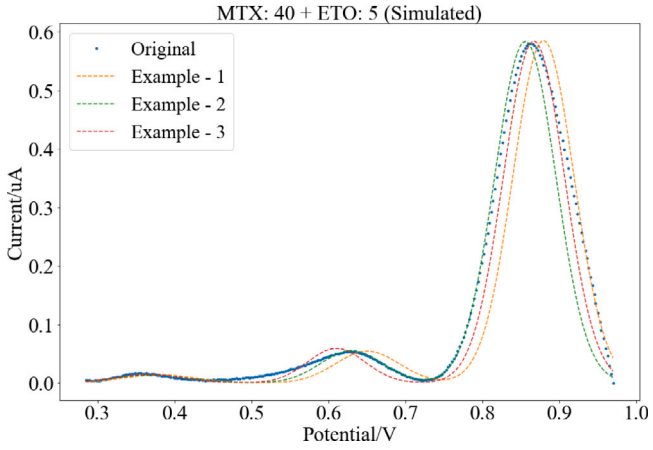


Fig. 6. The examples in simulated dataset.

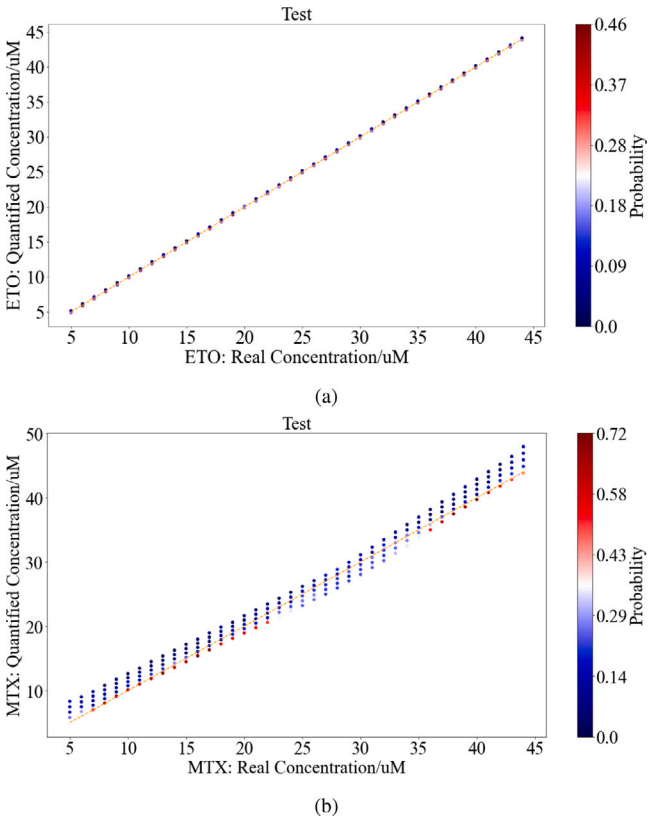


Fig. 7. Validation in the simulated test set: (a) result of ETO concentration quantification; (b) result of MTX concentration quantification.

for model training. Additionally, the Mean Average Percentage Error (MAPE) is introduced to evaluate the accuracy of generalized quantification. The aforementioned dataset is then split into a training and test set of ratio 90%/10%.

Fig. 7 demonstrates the validation result in test set split from the simulated dataset. For ETO and MTX respectively, the range of quantification results corresponding to each label within predefined limits is divided into five subsets and the probability of each subset is calculated.

Due to the little influence on the ETO's redox peaks with the concentration fluctuation of MTX, the labels are approximately positively correlated with ETO's concentrations in a linear relationship, regardless of the values of MTX's concentration in the preset range. Therefore, the MLP-based quantification model is able to simulate the relationship between the input feature vectors and the output more accurately and quantification results of ETO concentration in simulated test set are nearly in the vicinity of the labels (MAPE: 0.32%) as shown in Fig. 7(a). The inhibition interaction of ETO to MTX makes it difficult to predict the concentration of MTX within a wide preset range of two drugs, namely, the concentration quantification results of MTX are relatively scattered. However, after probability density analysis of the quantification results, most of the predicted results are still close to their labels with MAPE reaching at 4.78% (Fig. 7(b)).

Subsequently, the large simulated dataset constructed by the method proposed in this paper through automatic parameter estimation based on measured CV samples can effectively train the quantification model, thereby achieving comparatively ideal drug cocktail concentration quantification results in the simulated dataset. Namely, it can be a beneficial complement to the training set when a sufficiently large amount of measured data is not available.

5. Conclusion

An effective ML-based method is demonstrated to solve the problem of insufficient dataset availability for quantification model training in the development of personalized therapy for cancer treatments. The proposed model is based on the integrated application of multiple ML algorithms capable to automatically estimate the parameters needed to construct large simulation datasets from several CV samples. Through extensive analysis, the basic ML model trained by the dataset established by the method in this paper can achieve satisfactory quantification of multiple drug concentrations with mutual interactions in simulated test set, demonstrating the rationality of our simulated dataset automatically generated by our methodology and the capability used in generalized quantification model training. Future work will construct more precise pseudo data with different characteristics for model training and establish more powerful concentration quantification models with the capable structure to leverage the interactions between multiple drugs for the application of more effective and reliable TDM and personalized therapy.

CRediT authorship contribution statement

Lin Du: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yann Thoma:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Francesca Rodino:** Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. **Sandro Carrara:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sandro Carrara reports financial support was provided by Swiss National Science Foundation. If there are other authors they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Swiss National Science Foundation, Switzerland, with the project entitled Intelligent Platform for Drug Response in Precision Oncology, grant number 200021_207900/1.

References

[1] D. Levêque, G. Becker, The role of therapeutic drug monitoring in the management of safety of anticancer agents: a focus on 3 cytotoxics, Expert Opin. Drug Saf. 18 (11) (2019).
[2] S. Ohdo, S. Koyanagi, N. Matsunaga, Chronopharmacological strategies: Intra- and inter-individual variability of molecular clock, Adv. Drug Deliv. Rev. 62 (9) (2010) 885–897.
[3] A.M. Jarrett, D. Faghihi, D.A. Hormuth, E.A.B.F. Lima, J. Virostko, G. Biros, D. Patt, T.E. Yankeelov, Optimal control theory for personalized therapeutic regimens in oncology: Background, history, challenges, and opportunities, J. Clin. Med. 9 (5) (2020).
[4] T. Buclin, Y. Thoma, N. Widmer, P. André, M. Guidi, C. Csajka, L.A. Decosterd, The steps to therapeutic drug monitoring: A structured approach illustrated with imatinib, Front. Pharmacol. 11 (2020).

[5] P. Estrela, J.L. Hammond, N. Formisano, P. Estrela, S. Carrara, J. Tkac, Electrochemical biosensors and nanobiosensors, Essays Biochem. 60 (1) (2016) 69–80.
[6] H.R.S. Lima, J.S. da Silva, E.A. de Oliveira Farias, P.R.S. Teixeira, C. Eiras, L.C.C. Nunes, Electrochemical sensors and biosensors for the analysis of antineoplastic drugs, Biosens. Bioelectron. 108 (2018) 27–37.
[7] S. Aiassa, S. Carrara, D. Demarchi, Optimized sampling rate for voltammetry-based electrochemical sensing in wearable and IoT applications, IEEE Sens. Lett. 3 (6) (2019) 1–4, http://dx.doi.org/10.1109/LSENS.2019.2918575.
[8] Y. Yoon, M.J. Kim, J.J. Kim, Machine learning to electrochemistry: Analysis of polymers and halide ions in a copper electrolyte, Electrochim. Acta 399 (2021) 139424.
[9] S. Carrara, A. Cavallini, V. Erokhin, G. De Micheli, Multi-panel drugs detection in human serum for personalized therapy, Biosens. Bioelectron. 26 (9) (2011) 3914–3919.
[10] S. Aiassa, I. Ny Hanitra, G. Sandri, T. Totu, F. Grassi, F. Criscuolo, G. De Micheli, S. Carrara, D. Demarchi, Continuous monitoring of propofol in human serum with fouling compensation by support vector classifier, Biosens. Bioelectron. 171 (2021) 112666.
[11] B.-C. Gu, K.-J. Chung, B.-W. Chen, Y.-H. Dai, C.-C. Wu, Electrochemical detection combined with artificial neural networks for the simultaneous intelligent sensing of caffeine and chlorogenic acid, Electrochim. Acta 463 (2023) 142820.
[12] S. Asir, K. Dimililer, Y. Kirsal-Ever, M. Özsöz, N.A. Shama, Electrochemical determination of potassium ferricyanide using artificial intelligence, in: 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT, 2019, pp. 1–4, http://dx.doi.org/10.1109/ISMSIT.2019.8932910.
[13] A. Zimmer, A. Tendler, I. Katzir, A. Mayo, U. Alon, Prediction of drug cocktail effects when the number of measurements is limited, PLOS Biol. 15 (10) (2017) 1–16.
[14] S. Kuroda, S. Kagawa, T. Fujiwara, Chapter 12 - selectively replicating oncolytic adenoviruses combined with chemotherapy, radiotherapy, or molecular targeted therapy for treatment of human cancers, in: E.C. Lattime, S.L. Gerson (Eds.), Gene Therapy of Cancer (Third Edition), third ed., Academic Press, San Diego, 2014, pp. 171–183.
[15] A.A. Khand, S.A. Lakho, A. Tahira, M. Ubaidullah, A.A. Alothman, K. Aljadoa, A. Nafady, Z.H. Ibupoto, Facile electrochemical determination of methotrexate (MTX) using glassy carbon electrode-modified with electronically disordered NiO nanostructures, Nanomaterials 11 (5) (2021).
[16] F. Rodino, M. Bartoli, S. Carrara, Simultaneous and selective detection of etoposide and methotrexate with single electrochemical sensors for therapeutic drug monitoring, IEEE Sens. Lett. 7 (8) (2023) 1–4, http://dx.doi.org/10.1109/LSENS.2023.3300817.
[17] L. Jiang, K. Zheng, Towards the intelligent antioxidant activity evaluation of green tea products during storage: A joint cyclic voltammetry and machine learning study, Food Control 148 (2023) 109660.
[18] J.A. Hartigan, M.A. Wong, Algorithm AS 136: A K-means clustering algorithm, J. R. Stat. Soc. C (Appl. Stat.) 28 (1) (1979) 100–108, URL http://www.jstor.org/stable/2346830.
[19] T. Naka, N. Sakamoto, Kinetics of membrane-bound enzymes: Validity of quasi-steady-state approximation for a Michaelis-Menten-type reaction, J. Membr. Sci. 74 (1) (1992) 159–170.
[20] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.