Effects of inter-annual climate variability on grape harvest timing in rainfed hilly vineyards of piedmont (NW Italy)

Continuous monitoring of propofol in human serum with fouling compensation by support vector classifier

Simone Aiassa[a,b,*], Ivan Ny Hanitra[b], Gabriele Sandri[a,b], Tiberiu Totu[b], Francesco Grassi[a,b], Francesca Criscuolo[b], Giovanni De Micheli[b], Sandro Carrara[b] and Danilo Demarchi[a]

[a]*Department of Electronics and Telecommunications, Politecnico di Torino, Turin 10129, Italy*
[b]*Integrated Systems Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland*

ARTICLE INFO

Keywords:
Anaesthesia
Continuous drug monitoring
Electrode fouling
Machine learning
Propofol

ABSTRACT

We present a new method for electrochemical sensing, which compensates the fouling effect of propofol through machine learning (ML) model. Direct and continuous monitoring of propofol is crucial in the development of automatic systems for control of drug infusion in anaesthesiology. The fouling effect on electrodes discourages the possibility of continuous online monitoring of propofol since polymerization of the surface produces sensor drift. Several approaches have been proposed to limit the phenomenon at the biochemical interface; instead, here, we present a novel ML-based calibration procedure. In this paper, we analyze a dataset of 600 samples acquired through staircase cyclic voltammetry (SCV), resembling the scenario of continuous monitoring of propofol, both in PBS and in undiluted human serum, to demonstrate that ML-based model solves electrode fouling of anaesthetics. The proposed calibration approach is based on Gaussian radial basis function support vector classifier (RBF-SVC) that achieves classification accuracy of 98.9 % in PBS, and 100 % in undiluted human serum. The results prove the ability of the ML-based model to correctly classify propofol concentration in the therapeutic range between 1 µM and 60 µM with levels of 10 µM, continuously up to ten minutes, with one sample every 30 s.

1. Introduction

Continuous Therapeutic Drug Monitoring (TDM) is the clinical practice of constantly measuring the drug concentration in the effort of optimizing the drug dosage to each patient. Real-time monitoring of sedative drug concentrations (such as midazolam and sufentanil) was proved to be beneficial to avoid inadequate sedation and its complications in intensive care units for mechanically ventilated patients (Nies et al., 2018). In general intravenous anaesthesia, the optimal therapy requires the continuous monitoring of the concentration of the anaesthetic to maintain a certain level of sedation during the surgery, which is essential to avoid useless deep states and prolonged chemical coma. Propofol (2,6-diisopropyl phenol) is the most commonly used intravenous anaesthetic (Sahinovic et al., 2018). However, high concentrations of propofol are dangerous: common side effects include irregular heart rate, low blood pressure, addiction and stopping of breathing. In particular, propofol infusion syndrome is a rare complication of propofol administration presenting a high death rate, requiring an accurate prevention (Mirrakhimov et al., 2015). A real-time feedback-controlled algorithm that uses as input the continuously monitored concentration of propofol ensures both more precise and safer sedation than what is currently achieved without closed-loop system (Simalatsar et al., 2018).

Biosensors are playing a relevant role in numerous fields of application including healthcare, environment, and food industries (Kumar et al., 2019) and may be used for point-of-care diagnostics (Mahato et al., 2018). Electrochemical sensors have taken high relevance for propofol monitoring (Kivlehan et al., 2015; Stradolini et al., 2018a), especially, cyclic voltammetry (CV) techniques (Aiassa et al., 2019a). With this approach, some electronic systems were proposed for the monitoring of anaesthetics, e.g., propofol and paracetamol (Stradolini et al., 2018c; Aiassa et al., 2019b). The sensor is a three electrodes electrochemical cell consisting of a reference electrode (RE), a counter electrode (CE) and a working electrode (WE). A potentiostat applies a voltage sweep to the electrochemical cell, and then it samples the current produced by the induced redox. By analyzing the shape of the so-called voltammogram (current upon the voltage), it is possible to extract the concentration of the drug in the sample.

Although numerous sensors have been proposed for detecting and measuring propofol, achieving Limit Of Detection (LOD, the minimum concentration detectable by the sensor) of 0.5 µM (0.1 µg/ml) in only 25 s (Hong et al., 2016), most of those sensors leverage on blood-spot sampling with disposable sensors, most of the time single-use (Dincer et al., 2019), which cannot be used in an automatic TDM closed-loop system.

In continuous detection of propofol, when a difference of potential is applied to trigger the direct oxidation, a propofol free radical is generated (Langmaier et al., 2011). Unfortunately, free radicals react with O_2 or they undergo polymerization, leading to the formation of a polymeric film (Heyne et al., 2003; Stradolini et al., 2018b). The polymeric thin film covering the electrodes may degrade the sensor signal leading to the fouling phenomenon (Yang et al., 2013). The fouling film is robust, with low permeability, thermally stable and chemically inert itself. Being composed of high molecular weight species, it adheres tightly to the electrode. The fouling is highly limiting the development of closed-loop systems for monitoring of anaesthetics because continuous monitoring of propofol requires high stability in time.

The literature presents a few attempts to solve the propofol fouling problem. Good results were achieved by coating the electrode with a PVC membrane (Kivlehan et al., 2015). Still, its low mechanical resistance constrains to carry out electrode cleaning in order to ensure the long term performances (Stradolini et al., 2018b), which, on the other hand, limits the final application since it requires constant human intervention. Commercial pencils are suit-

*Corresponding author
simone.aiassa@polito.it (S. Aiassa)
ORCID(s): 0000-0001-7985-7186 (S. Aiassa)

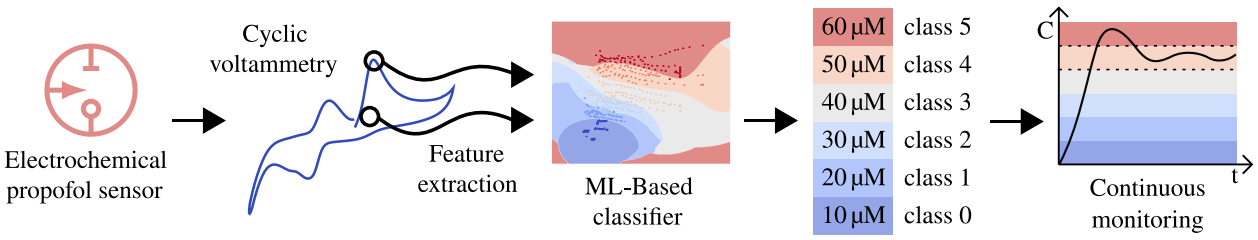

Figure 1: Proposedmachine-learning based approach for continuous monitoring of propofol: from left to right, the sample is analyzed through electrochemical sensors and CV to extract relevant features to be fed a ML-based that determines the range of concentration of the propofol, compensating the fouling effect.

able for propofol detection, and their composition balances the fouling effect (Stradolini et al., 2018a). Nevertheless, sensor response highly depends on the conformation of the sensor, where any small change in geometry, size, and composition results in a drastic drop in sensor performance.

In this work, we propose a novel machine learning (ML)-based calibration procedure to compensate the fouling effect. Our ML classifier is implemented to identify the correct concentration of propofol in a given sample to contribute to the development of a system for closed-loop controlled infusion of anaesthetic. Our classifier is designed, and its parameters are optimized using a large dataset of 480 samples acquired in PBS background. Later, the model is validated with a smaller, still representative, dataset of 120 samples for direct monitoring of propofol in undiluted human serum.

2. Machine learning for continuous monitoring

Soft modelling techniques based on ML models are commonly coupled to biosensors as tools for solving complex mathematical models related to biochemical processes (Esteban et al., 2006; De Vito et al., 2018). Applications such as peak deconvolution, pH, temperature and fouling compensation have been successfully applied to e-tongues and e-noses cyclic voltammograms by the aid of ML models, such as support vector machines and artificial neural networks (ANNs) (del Valle, 2017; De Vito et al., 2018; Wang et al., 2019). Those systems extract relevant features from biochemical analysis, and they provide them to ML-based algorithms to build complex mathematical models. For ANN-based ML-based techniques, key components to be extracted from a cyclic voltammogram are the peak position, peak height, peak width half-height, peak sum of derivatives, and charge (Asir et al., 2019).

The fouling phenomena in phenolic compounds are characterized by a strong non-linear response, limiting the application of pre-established mathematical models. Chemometric techniques based on ML have been successfully applied to compensate for the lack of theoretical models (Apetrei and Apetrei, 2013; Maleki et al., 2017; Li et al., 2019; Sheng et al., 2019). To the best of our knowledge, an ML algorithm specifically designed to balance the fouling of propofol has never been proposed.

Recent works suggested that monitoring controlled-delivery of anaesthetics may be achieved with a 10 % accuracy around the target concentration, and with one measurement every 30 s, continuously in time, while the therapeutic concentration of propofol ranges between 0.25 mg/l and 10 mg/l (1 : 60 µM) (Simalatsar et al., 2018). The goal of this work is to develop a technique suitable for the continuous measurement of propofol concentration every half a minute and able to discriminate the concentration level lower than 12 µM. Fig. 1 presents the proposed ML-based method. Through an electrochemical sensor and CV technique, the redox of propofol is analyzed. From the voltammogram, several peculiar features are extracted and fed to the ML-based classifier. The classifier determines the concentration level of propofol in the sample in the form of classes, compensating fouling of propofol. According to the requirements, we subdivided the therapeutic range into classes representing concentration levels of 10 µM. The developed ML classifier provides the anesthesiologist with a tool for direct determination of the concentration of propofol in human serum. The classifier, compared to a regressor, directly provides the information required to maintain the constant-dose in the range of interest.

3. Materials and methods

In this work, chemicals, setup, and sensor (Section 3.1) are defined leveraging of previous works that had been proven to be suitable for the monitoring of anaesthetics (Stradolini et al., 2018a; Tuoheti et al., 2020). Section 3.2 describes the procedure for acquiring the samples, which resembles the real scenario of the target application. The dataset elaboration is detailed in Section 3.3, and used to train an ML-based algorithm, which is fully described in Section 3.4.

3.1. Chemicals

A stock solution of 5.4 mM propofol is prepared on the day of use with 2,6-Diisopropylphenol (propofol) purchased from Tokyo Chemical Industry Co., Ltd. and dissolved in 0.1 M NaOH. The samples are prepared with seven concentrations of propofol, equally spaced in its therapeutic range: 10, 20, 30, 40, 50, and 60 µM. The samples for primary analysis, and the training and testing sets for the ML classifiers are prepared in Phosphate Buffer Saline (PBS, 10 mM, pH 7.4) from Sigma Aldrich®. The final testing is done with real and undiluted human serum with the addition of propofol in the therapeutic range at body temperature, which artificially mimics the infusion of anaesthetic done in the clinical environment. The samples for classification in human serum are prepared in undiluted human serum, heat-inactivated from human male AB plasma, from Sigma-Aldrich®, and they are continuously kept at 37 °C by a hot plate stirrer from VWR®.

The sensor is a three-electrode electrochemical cell consisting of two 0.5 mm diameter HB mechanical pencil lead from Papeteria Migros, and one 0.3 mm diameter platinum wire as shown in Fig. 2. Similarly to what presented in (Tuoheti et al., 2020), the WE consists of one pencil lead, exposed to the sample for 10 mm of its length, resulting in an active area of around 32 mm^2. The CE is the other pencil lead exposed to the sample for 15 mm to maintain a ratio between the area of WE and CE smaller than one. The RE is the platinum wire exposed to the sample for 10 mm. A silicon disk keeps the cell stable, and the electrodes are directly wired

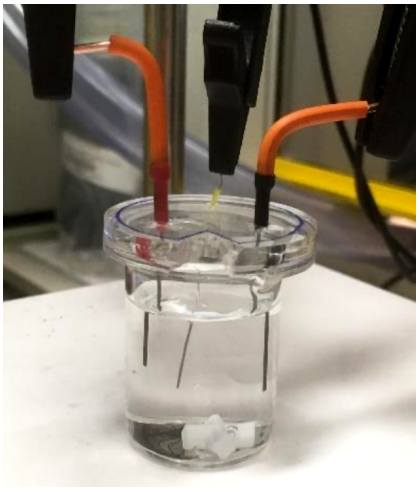

Figure 2: Proposed experimental setup: the three-electrode electrochemical cell consisting of two pencil lead (WE and CE) and a platinum wire (RE) that are immersed in the sample containing propofol dissolved in ether PBS or Human Serum. The wires are connected to the potentiostat to perform CV.

on their top to the instrumentation. The design of the electrochemical cell is the result of a long study performed through potassium ferrocyanide to define the best geometry for a disposable and miniaturized sensor. It is worth to be noticed that the platinum pseudo-reference electrode utilized in this work did not show any visible difference of stability when compared to silver reference electrodes present on commercial screen-printed electrodes. The selectivity of this type of pencil graphite sensors has been proven to be optimal for the detection of propofol in human serum for the target application through interference studies in previous works (Stradolini et al., 2018b).

3.2. Data collection

According to the description in Section 2, the proposed sensor is capable of compensating fouling in real-time, at a fixed concentration of analyte in its therapeutic range. For this reason, all the samples prepared according to Section 3.1 are measured twenty consecutive times, with an interval time of 30 s, to be consistent with the continuous monitoring target of this work.

The proposed sensor is connected to a commercial potentiostat, the Metrohm Autolab PGSTAT 302N, driven by the software Nova 1.11. The analysis of the sample is performed via a staircase cyclic voltammetry (SCV) procedure. The SCV is executed with a scan rate of 0.1 V/s, with a driving voltage ranging in $-0.8 : 0.7$ V, with a step voltage height of 5 mV, and a step time length of 30 ms, starting from 0 V, reaching up to 0.7 V, lowering down to -0.8 V, and finally settling again at 0 V. It is worth noticing that the procedure is not formally cyclic because the proposed method entails that the fouling is limited to its minimum. Cycling the voltage scan more than one time per measure would catastrophically reduce the lifespan of the sensor.

3.3. Dataset

The dataset consists of $m = 480$ samples acquired from four different sensors in PBS and $m = 120$ samples obtained in human serum. The dataset in PBS has been extended with respect to the human serum one in order to improve the optimization process of the ML-based algorithm and its parameters. Each recorded sample is a voltammogram relating the current measured in function of the potential applied to the electrochemical cell. The voltammograms acquired from propofol measurements at known concentrations are analyzed in order to extract three relevant features the peak current, the potential at peak current, and the total charge. The peak Faradaic current i_p is the current resulting from primary propofol oxidation, with the removal of the baseline charging current. E_p is the cell potential at which the peak current is achieved. i_p and E_p are the most relevant features for characterizing the electrochemical measurements (Carrara, 2012). Moreover, the integral of the voltammogram in the window $0 : 0.7$ V is computed. The latter is the total charge exchanged during primary propofol oxidation, denoted as Q, that has been proved to be relevant in the determination of drugs (Aiassa et al., 2020). Lastly, the ordinal number of measurements performed with the same sensor, n_{meas}, is added to the feature list. The input features matrix $\mathbf{X}$ is first standardized by removing column mean, and scaled to unit variance. This is of paramount importance since the features have different unit scales. The dataset is then split into a training and test set of ratio 80 % / 20 %.

3.4. Support vector classifier

Given the nonlinearity induced by the fouling on the sensor response, non-linear classifiers are considered in this work. Kernelized-SVCs are suitable for non-linear classification problems by constructing the optimal hyperplane in the features space induced by a kernel function (Cortes and Vapnik, 1995). The SVC predicts propofol concentration class of unknown samples according to distance or similarity of the samples to the training instances. All processing stages and algorithms are implemented within a Python 3.7.4 environment, using NumPy and scikit-learn libraries (Pedregosa et al., 2011).

Let $\mathbf{X} = \{\mathbf{x}_i\}_{i=1,\cdots,m}$, with $\mathbf{x}_i \in \mathbb{R}^n$, denote the input tensor of m samples and n features. Let $\mathbf{y} \in \mathbb{R}^m$ denote the vector labelling the propofol concentration of each measurement, where each target concentration (10, 20, 30, 40, 50, and 60 µM) is encoded into a categorical class with values from zero to five. The objective of the six-class classifier is to construct a predictor based on the training set $(\mathbf{X}_{train}, \mathbf{y}_{train})$ that is able to divide the input features space into a collection of regions belonging to each class. The decision boundaries are refined in the training process, for which the metric used is the classification accuracy:

$$accuracy = \frac{\#\ correctly\ predicted\ test\ samples}{\#\ test\ samples} \tag{1}$$

During inference, the predicted class indicates the range of concentration of propofol in the sample. In kernelized-SVCs the optimization objective is

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \frac{1}{2}\boldsymbol{\alpha}^T \cdot \mathbf{M} \cdot \boldsymbol{\alpha} - \mathbf{e}^T \cdot \boldsymbol{\alpha} \\ \text{subject to} \quad & \mathbf{y}^T \cdot \boldsymbol{\alpha} = 0, \\ & 0 \le \alpha_i \le C,\ i = 1, \cdots, m, \end{aligned} \tag{2}$$

where $\mathbf{M}$ is an m by m positive semi-definite matrix $M_{ij} = y_i \cdot y_j \cdot K(\mathbf{x}_i, \mathbf{x}_j)$, and $K(\mathbf{x}_i, \mathbf{x}_j) \equiv \phi(\mathbf{x}_i)^T \cdot \phi(\mathbf{x}_j)$ is the equation defining the SVC kernel. $\mathbf{e}$ is a vector of ones of length m. C is a regularization hyper-parameter tuning the tolerance of margin violations. Linear, polynomial, Gaussian radial basis function (RBF), and sigmoid kernels are investigated. The training instances are shuffled, and ten-splits cross-validation is performed, where the metric used

is the prediction accuracy based on decision boundaries. The ten splits are stratified to preserve the number of samples per class in each training and validation split (80 % / 20 %). An analysis is carried out to assess how the type of kernel, and the features to be considered in the dataset, influence SVC performance. Then, a cross-validation grid-search on the kernelized-SVC hyper-parameters is carried out to optimize the classifier. Finally, the classification performance of the optimal SVC is evaluated on the test set. LIBSVM library is used, where a one-versus-one scheme is adopted for the multi-class classification (Chang and Lin, 2011).

4. Results

Section 4.1 presents the performance of the sensor for propofol detection through standard metrics and shows the limitation of a linear model in the given application. Several experiments and analysis are carried out to implement the here proposed ML-based algorithm. Namely, the optimal kernel is selected (Section 4.2), the best set of features is defined (Section 4.3), and the parameters of the classifier are optimized (Section 4.4). The previous experiments are performed with the dataset obtained from the 480 samples in PBS background. Such a stable background may facilitate the identification of the best ML model for the compensation of propofol fouling. Finally, in Section 4.5, the proposed ML-based SVM is validated on human serum.

4.1. Limits of standard linear model

Fig. 3 presents the resulting voltammogram obtained by repetitively acquiring the Faradaic current in a sample of 60 µM of propofol, considering PBS (Fig. 3a), and human serum (Fig. 3b) as background. Even though both the sensor and the concentration of propofol are not varying, the curves are radically varying. Fig. 3a illustrates fouling on the carbon surface of the electrode in PBS. Namely, every new measurement the fouling layer increases on the interface, reducing the height of the primary oxidation peak (A), and shifting the peak itself to the right. Peak B and C are substantially changing in time, similarly to A, presenting fouling as well. In this work, the analysis focuses on the primary peak A, which is the best candidate to determine the concentration of propofol, due to higher magnitude which improves sensitivity, and higher distance from other peaks, which enhances selectivity.

Fig. 3b shows a detail of peak A in the undiluted human serum, obtained by baseline subtraction and filtering. The human serum contains proteins completely absent in PBS, which embed and adsorb the propofol. For this reason, the free propofol detectable in serum is lower than in PBS, resulting in reduced Faradaic current, signal strength, and reduced sensitivity (Stradolini et al., 2018c). As expected, the current peak is lower (one fifth), and the passivation due to fouling is again visible.

The full dataset acquired is elaborated to extract the sensor calibration according to the linear model commonly used in electrochemical sensors (Aiassa et al., 2019a) to evaluate the limitation introduced by the fouling phenomenon. The calibration relates linearly the primary oxidation current peak and the propofol concentration recalling the Randles-Ševćik equation (Carrara, 2012). The calibration is performed using 80 % of the samples, to be consistent with the ML-based method. The linear calibration procedures resulted in a sensor with a sensitivity of 162.9 ± 10.3 nA/µM and a LOD of 2.4 ± 0.1 µM with PBS as background. In human serum, the sensor presented a sensitivity of 28.8 ± 7.7 nA/µM and LOD of 4.9 ± 1.3 µM. The sensitivity has been calculated as the coefficient of regression through linear regression fit, while LOD has been computed as three times the standard deviation of the blank signal around the peak, over the sensitivity (Stradolini et al., 2018a). Both sensitivity and LOD variations are computed with the residual sum of squares of three times standard deviation of each observation.

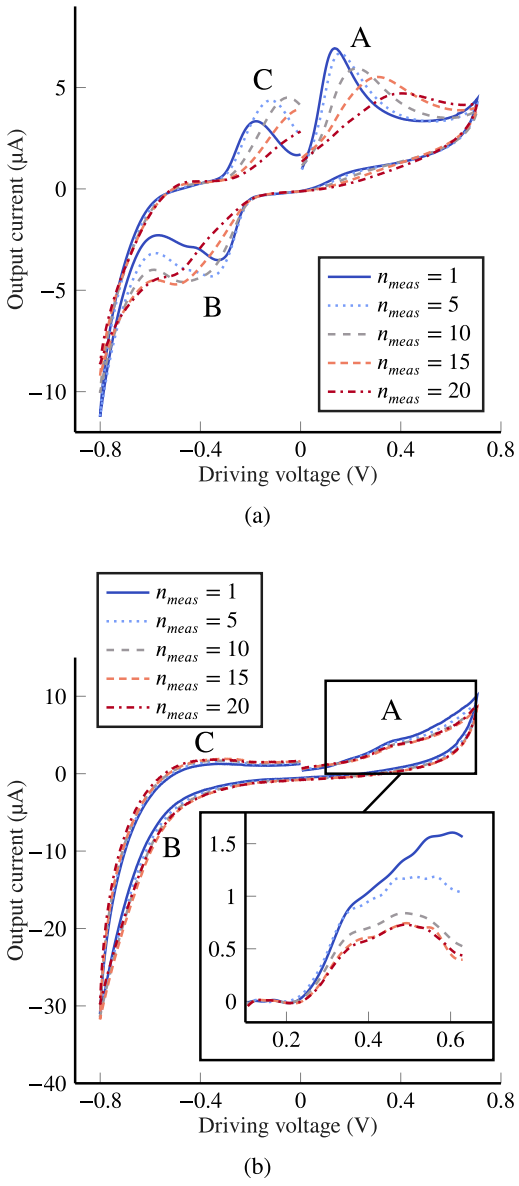

Figure 3: Voltammogram from propofol detection in time at fixed 60 µM in PBS (a), and in human serum (b). In human serum, the peak is graphically highlighted by baseline subtraction, and filtering since the lower free concentration of propofol reduces its visibility. In both cases, the peak A lowers in current after each new measurement due to the fouling phenomenon.

Despite being promising, the two extracted linear calibrations present their limits only when considering the measurements in time. Reporting the linear model to a six-classes classifier, as the ML-based model will do in this work, the classification accuracy tested on the remaining 20 % of the samples are 69.8 % and 33.3 % in PBS and human serum, respectively. This analysis proves that it is not possible to develop a system for continuous monitoring of propofol concentration without compensating the non-linear fouling effect with a non-linear model.

Table 1
Selection of the appropriate kernel for propofol classification: comparison among different kernelized-SVCs in $i_p - E_p$ features space with their default hyper-parameters. The best results are achieved with RBF-SVC.

Kernel type	Linear	Polynomial	RBF	Sigmoid
Kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$	$\mathbf{x}_i^T \cdot \mathbf{x}_j$	$(\gamma \cdot \mathbf{x}_i^T \cdot \mathbf{x}_j + r)^d$	$\exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$	$tanh(\gamma \cdot \mathbf{x}_i^T \cdot \mathbf{x}_j + r)$
Kernel hyper-parameters	-	$\gamma = 1/m$, $r = 0$, $d = 3$	$\gamma = 1/m$	$\gamma = 1/m$, $r = 0$
Soft-margin penalty parameter	$C = 10$	$C = 10$	$C = 10$	$C = 10$
Classification accuracy	43.8 %	86.5 %	90.6 %	33.3 %

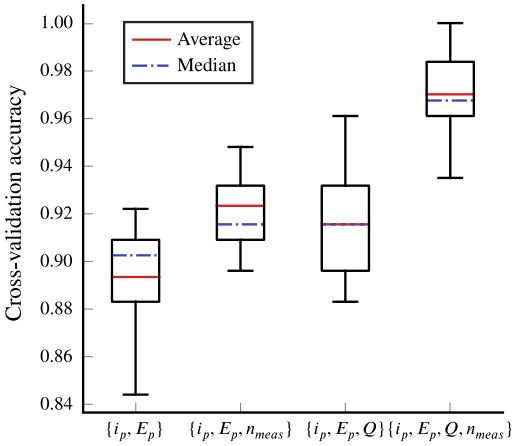

Figure 4: Cross-validation classification accuracy of RBF-SVC trained with four combinations of input features. The box plots extend from lower to upper quartile values of cross-validation accuracy, with median (dotted blue line) and average (red line), and whiskers show the range of classification accuracy. The complete set of feature ($\{i_p, E_p, n_{meas}, Q\}$) shows visibly an higher accuracy, that cannot be reached with less features.

4.2. Kernel selection

The selection of the most appropriate kernel for the classification of propofol is exploited implementing linear, polynomial, RBF, and sigmoid SVCs with their default hyper-parameters. The classifiers decision boundaries are constructed from the training set in the space built by the combination of standardized peak current and standardized potential at peak current ($i_p - E_p$ space). The latter are the main features to characterize propofol electrochemical measurements. The different classifiers are trained with ten-splits cross-validation on the training set and evaluated on the test set. The four different kernel function, the kernel hyper-parameters, and their classification performance are reported in Table 1. The decision boundaries of each kernelized-SVC are visualized in Supplementary Material, Fig. S3. Linear and sigmoid kernels are not suitable for detection of propofol (see Table 1). Meanwhile, non-linear kernels enable computing the decision hyperplanes in the space of higher dimension. Polynomial and RBF are the most accurate kernels, and the decision boundaries are smoothly separating the classes. RBF kernel is chosen for the subsequent experiments since it yields the higher classification accuracy of 90.6 % on the test set.

4.3. Features selection

Different combinations of input features are evaluated on the RBF-SVC. This optimization will help to understand the effect of the different features extracted from the voltammograms on the classification accuracy of the ML model. The possible combinations of features considered are $\{i_p, E_p\}$, $\{i_p, E_p, n_{meas}\}$, $\{i_p, E_p, Q\}$, and $\{i_p, E_p, n_{meas}, Q\}$. All the different input sets are fed to the classifier, and a ten-splits cross-validation is performed on the training set. The cross-validation accuracies are presented in Fig. 4 with a whisker plot. It highlights that classification accuracy scales with the amount of features included in the dataset. Indeed, it could be noticed from Supplementary Material, Fig.S3, that samples belonging to classes two and three are mis-classified in features space $i_p - E_p$, since the samples are intermingled. Besides, for the set of features $\{i_p, E_p, Q\}$, classification accuracy is more dispersed, and it is lower than using n_{meas} instead of Q. There is a high correlation between the charge exchanged during propofol oxidation and the peak oxidation current. When the training set contains all four features, significant improvement in cross-validation accuracy is observed, and the latter reaches 0.970 ± 0.020.

4.4. Hyper-parameters optimization

RBF-SVC hyper-parameters are tuned through cross-validation grid-search in order to optimize the hyper-parameters of our proposed ML model. The non-linear coefficient γ is swept from 10^{-9} to 10^3, while the soft margin penalty parameter C is swept from 10^{-2} to 10^{10}. Both sweeps are performed in logarithmic scale, training 169 SVC models. The training set $(\mathbf{X}_{train}, \mathbf{y}_{train})$ comprising the four features is shuffled, and ten-splits cross-validations are implemented. The training and validation accuracies are computed for each split. Their average value is retained for classifier comparison. Supplementary Material, Fig. S2, displays the heatmap of the cross-validation accuracy on hyper-parameters, highlighting the hyper-parameter space yielding the most accurate classifier. It is observed that for large values of γ, the support vectors are not able to separate the samples. They influence very few training instances. Conversely, a low value of γ over-constrains the classifier model, that ends up behaving like a linear classifier. The classifier does not capture the complexity of the non-linear dataset. As for the soft-margin penalty parameter C, a lower value is preferred to reduce over-fitting trading-off accuracy. Larger values for C tend to generalize better, but maximum accuracy is reached for $C = 10^4$, and it does not improve beyond. RBF-SVC models with C lower than 10^4, and yielding cross-validation accuracy above 97.5 % are evaluated on the test set. RBF-SVC with parameters $\{C = 100, \gamma = 1\}$ yields classification accuracy of 98.9 %.

4.5. Validation in human serum

After the selection of the best ML-based model for propofol fouling compensation in PBS buffer, the proposed RBF-SVC classifier is validated in undiluted human serum at the body temperature (37 °C). The CV dataset from propofol measurement in human serum is pre-processed as the dataset from PBS. The column

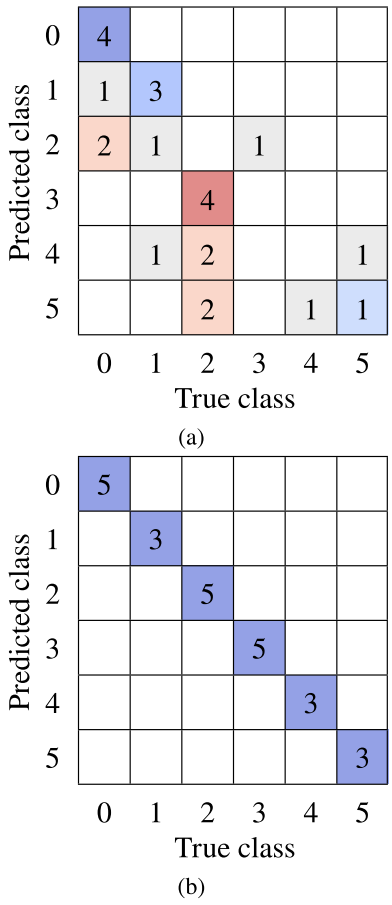

Figure 5: Results of validation in undiluted human serum: the confusion matrices show the difference between real concentration (true class) and the concentration estimated (predicted class). While the standard linear model (a) leads to wrong estimation 33 % of the time, the proposed ML-based model (b) always outputs a correct result.

mean-centering and standardization to unit-variance are applied, and the dataset is split into training/test set of ratio (80 % / 20 %). All four features $\{i_p, E_p, n_{meas}, Q\}$ are sent to the classifier and a cross-validation grid-search is carried out for the optimization of the hyper-parameters C and γ. The RBF-SVC models yielding cross-validation accuracy superior to 94.0 % are selected, and evaluated on the test set. The same RBF-SVC which obtained the higher performance in PBS (with $\{C = 100, \gamma = 1\}$), yields the best results also in human serum. The maximum classification accuracy achieved by our model, in human serum, is 100 %, with a null generalization error. The lower error and higher gain achieved by the classifier in serum with respect to the PBS is justified by the smaller dataset, 120 samples for human serum and 480 samples for PBS.

Fig. 5 displays the results of the validation carried directly in undiluted human serum, in the form of confusion matrices, which presents the prediction accuracy graphically. As visible in Fig. 5a the standard linear model features a classification accuracy of 33.3 %, leading to a wrong estimation of the concentration in 66 % of the cases. Meanwhile, the proposed ML-based model with the RBF-SVC classifier (Fig. 5b) compensates the fouling resulting in an accuracy of 100 %. With this result, we prove that our sensor is suitable for continuous monitoring of propofol for up to ten minutes, with one sample every 30 s, discriminating concentration levels of 10 µM.

5. Discussion

Previous works had proven that it is possible to detect propofol with small LOD with inexpensive and disposable sensors (Hong et al., 2016). At the same time, Simalatsar et al. (2018) opened the need for sensors for continuous monitoring of anaesthetics for improved TDM-assisted anaesthesiology practice. With this goal in mind, Kivlehan et al. (2015) and Stradolini et al. (2018a) faced the problems of fouling and electrode-passivation in continuous measurement of propofol using new materials and new mechanical procedures. In this work, we demonstrated, on one hand, the difficulties of using standard linear models in continuous measuring of propofol, on the other hand, we proposed a novel soft-modelling based solution to compensate via ML-based method the problem of fouling.

The proposed sensor itself features LOD of 4.9 ± 1.3 µM in real undiluted human serum, which is ten times more than the LOD reached by (Hong et al., 2016), but still below the minimum concentration of interest (10 µM). More comparisons on the LOD with respect to the state-of-the-art are available on Supplementary Material, Table S1. The kernelized-SVM has been proven to be optimal for compensating the problem of fouling since it reaches 100 % of accuracy in real undiluted human serum at steady 37 °C in discriminating 10 µM of propofol. Moreover, extensive experiment and optimization have demonstrated that the best kernel for SVM in this application is RBF, with the best parameter set to $\{C = 100, \gamma = 1\}$, considering as input feature the Faradaic primary current peak (i_p), the potential at the current peak (E_p), the total charge exchanged during Faradaic process (Q), and the ordinal number of measurements performed with a given sensor since it was new (n_{meas}).

6. Conclusion

We developed a novel ML-assisted method to compensate the fouling effect of propofol on electrochemical sensors to improve the anaesthesiology practices. Through extensive analysis, we demonstrate that the proposed model based on Gaussian RBF-SVC helps to obtain high classification accuracy (higher than 98.9 %) both in PBS and in human serum. Our ML-based model discriminates 10 µM concentration with 100 % classification accuracy, directly in undiluted human serum at body temperature, and continuously up to ten minutes to meet the requirement for the development of a system for closed-loop controlled-infusion of anaesthetics. Future work will include the implementation of the proposed model in a portable electronic device for continuous monitoring of anaesthetics and its test with clinical samples.

Acknowledgements

This work is supported by Politecnico di Torino and Compagnia di San Paolo under the initiative "Joint research projects with top universities".

CRediT authorship contribution statement

Simone Aiassa: Conceptualization, Methodology, Validation, Investigation, Writing - Original Draft. **Ivan Ny Hanitra:** Software, Validation, Writing - Original Draft. **Gabriele Sandri:** Methodology, Validation, Investigation. **Tiberiu Totu:** Investigation, Writing - Original Draft. **Francesco Grassi:** Conceptualization, Investigation. **Francesca Criscuolo:** Methodology, Writing

- Review & Editing. **Giovanni De Micheli:** Supervision, Writing - Review & Editing. **Sandro Carrara:** Supervision, Writing - Review & Editing. **Danilo Demarchi:** Project administration, Writing - Review & Editing.

References

Aiassa, S., Carrara, S., Demarchi, D., 2019a. IEEE Sens. Lett. 3, 1–4.

Aiassa, S., González Martínez, J.D., Demarchi, D., Carrara, S., 2020. New measurement method in drug sensing by direct total-charge detection in voltammetry, in: 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1–6.

Aiassa, S., Stradolini, F., Tuoheti, A., Carrara, S., Demarchi, D., 2019b. Quasi-digital biosensor-interface for a portable pen to monitor anaesthetics delivery, in: 2019 15th Conference on Ph. D Research in Microelectronics and Electronics (PRIME), pp. 265–268.

Apetrei, I.M., Apetrei, C., 2013. Food Res. Int. 54, 2075–2082.

Asir, S., Dimililer, K., Kirsal-Ever, Y., Özsöz, M., Shama, N.A., 2019. Electrochemical determination of potassium ferricyanide using artificial intelligence, in: 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1–4.

Carrara, S., 2012. Bio/CMOS interfaces and co-design. Springer Science.

Chang, C.C., Lin, C.J., 2011. ACM Trans. Intell. Syst. Technol. 2, 1–39.

Cortes, C., Vapnik, V., 1995. Mach. Learn. 20, 273–297.

De Vito, S., Esposito, E., Salvato, M., Popoola, O., Formisano, F., Jones, R., Di Francia, G., 2018. Sens. Actuators B Chem. 255, 1191–1210.

Dincer, C., Bruch, R., Costa-Rama, E., Fernández-Abedul, M.T., Merkoçi, A., Manz, A., Urban, G.A., Güder, F., 2019. Adv. Mater. 31, 1806739.

Esteban, M., Ariño, C., Díaz-Cruz, J., 2006. Trends Analyt. Chem. 25, 86–92.

Heyne, B., Kohnen, S., Brault, D., Mouithys-Mickalad, A., Tfibel, F., Hans, P., Fontaine-Aupart, M.P., Hoebeke, M., 2003. Photochem. Photobiol. Sci. 2, 939–945.

Hong, C.C., Lin, C.C., Hong, C.L., Lin, Z.X., Chung, M.H., Hsieh, P.W., 2016. Biosens. Bioelectron. 86, 623–629.

Kivlehan, F., Chaum, E., Lindner, E., 2015. Analyst 140, 98–106.

Kumar, A., Purohit, B., Maurya, P.K., Pandey, L.M., Chandra, P., 2019. Electroanalysis 31, 1615–1629.

Langmaier, J., Garay, F., Kivlehan, F., Chaum, E., Lindner, E., 2011. Anal. Chim. Acta 704, 63–67.

Li, Z., Huang, S., Chen, J., 2019. RSC Adv. 9, 34196–34206.

Mahato, K., Maurya, P.K., Chandra, P., 2018. 3 Biotech 8, 149.

Maleki, N., Kashanian, S., Maleki, E., Nazari, M., 2017. Biochem. Eng. J. 128, 1–11.

Mirrakhimov, A.E., Voore, P., Halytskyy, O., Khan, M., Ali, A.M., 2015. Crit. Care Res. Pract. 2015, 1–10.

Nies, R.J., Müller, C., Pfister, R., Binder, P.S., Nosseir, N., Nettersheim, F.S., Kuhr, K., Wiesen, M.H., Kochanek, M., Michels, G., 2018. J. Intensive Care 6, 1–8.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. J. Mach. Learn. Res. 12, 2825–2830.

Sahinovic, M.M., Struys, M.M., Absalom, A.R., 2018. Clin. Pharmacokinet. 57, 1539–1558.

Sheng, Y., Qian, W., Huang, J., Wu, B., Yang, J., Xue, T., Ge, Y., Wen, Y., 2019. Microchim. Acta 186, 543.

Simalatsar, A., Guidi, M., Roduit, P., Buclin, T., 2018. Smart Health 9, 101–114.

Stradolini, F., Kilic, T., Di Consiglio, A., Ozsoz, M., De Micheli, G., Carrara, S., 2018a. Electroanalysis 30, 1363–1369.

Stradolini, F., Kilic, T., Taurino, I., De Micheli, G., Carrara, S., 2018b. Sens. Actuators B Chem. 269, 304–313.

Stradolini, F., Tuoheti, A., Kilic, T., Ntella, S.L., Tamburrano, N., Huang, Z., De Micheli, G., Demarchi, D., Carrara, S., 2018c. IEEE Trans. Biomed. Circuits Syst. 12, 1056–1064.

Tuoheti, A., Aiassa, S., Criscuolo, F., Stradolini, F., Tzouvadaki, I., Carrara, S., Demarchi, D., 2020. IEEE Trans. Nanobioscience 19, 339–346.

del Valle, M., 2017. Materials for electronic tongues: Smart sensor combining different materials and chemometric tools, in: Materials for Chemical Sensing. Springer, pp. 227–265.

Wang, T., Lu, Y., Cao, Z., Shu, L., Zheng, X., Liu, A., Xie, M., 2019. Sensors 19, 5324.

Yang, X., Kirsch, J., Fergus, J., Simonian, A., 2013. Electrochim. Acta 94, 259–268.