Machine learning techniques to enable closed-loop control in anesthesia

Olivier Caelen [1,2], Gianluca Bontempi [2], Eddy Coussaert [3], Luc Barvais [4]
Université Libre de Bruxelles
Bruxelles, Belgium
{ocaelen,gbonte,coussae,lbarvais}@ulb.ac.be

François Clément
MEXYS- SA.
Mons, Belgium
fclement@mexys.com

Abstract

The growing availability of high throughput measurement devices in the operating room makes possible the collection of a huge amount of data about the state of the patient and the doctors' practice during a surgical operation. This paper explores the possibility of extracting from these data relevant information and pertinent decision rules in order to support the daily anesthesia procedures. In particular we focus on machine learning strategies to design a closed-loop controller that, in a near future, could play the role of a decision support tool and, in a further perspective, the one of automatic pilot *of the anesthesia procedure. Two strategies (direct and inverse) for learning a controller from observed data are assessed on the basis of a database of measurements collected in recent years by the ULB Erasme anaesthesiology group. The preliminary results of the learning approach applied to the regulation of hypnosis through the bispectral index (BIS) in a simulated framework appear to be promising and worthy of future investigation.*

1 Introduction

In recent years, a growing number of organizations in several domains have been allocating vast amounts of resources to construct and maintain large databases and create data warehouses. Health institutions are no exception and nowadays a lot of medical teams are using low cost computer technology enabling effective signal and data storage. In this new scenario, machine learning and data mining are key technologies in order to transform data into useful information for better diagnosis, event detection and decision aid. This paper deals with the anesthesia domain where several platforms have been made recently available for supporting the anesthesiologist in the operating room. An example is the TOOLBOX software [8] which has been used for several years by the group of anaesthesiology of the ULB Erasme hospital[5]. This tool, implemented in Smalltalk, monitors the patient'state and acts as a servo-controller on the multiple intravenous drug infusions, whose setting is regularly adjusted by the anesthesiologist, by simultaneously using pharmacokinetic and pharmacodynamic principles [2] (Figure 1). Before and during the operation TOOLBOX stores plenty of statistics and monitoring signals like: (i) generic information about the doctor, the patient and her general state, (ii) information about the type of surgery, (iii) the evolution of the hemodynamic and physiological parameters (e.g. the BIS) of the patient, (iv) the evolution of the drugs concentration levels chosen by the anesthesiologist.

This paper discusses and assesses the role of machine learning techniques in extracting information from the data collected by TOOLBOX during daily operations. In particular we will focus on the control of hypnosis through the bispectral index by acting on the Propofol (hypnotic) drug levels. The bispectral index (BIS) [20, 9] is a well-known measure adopted by anesthesiologists to rate the depth of the hypnosis. The BIS index represents the electroencephalographic signal in a normalized range from 100 to 0, where 100 stands for the "awake" status and 0 stands for electrical silence.

In our controller, the control action is expressed in terms of change of the current Propofol level and is computed on the basis of the current BIS, the desired BIS target, the Remifentanil (analgesic) level and some patient personal data (e.g. age and weight). The Remifentanil is taken into account because of its known impact on the interaction between Propofol and the BIS index [10].

A machine learning approaches to the design of a

[1]The work of the researcher was supported by the the F.I.R.S.T. project "Data mining prédictif en anesthésie intraveineuse informatisé en vue de certification" (project number EP1A320501R059F/415717) of the Région Wallonne, Belgium.

[2]Machine Learning Group; Département d'Informatique; ULB
[3]Département d'Informatique; Faculté de Médecine; ULB
[4]Service d'Anesthésiologie-Réanimation; Faculté de Médecine; ULB

[5]The Hôpital Erasme is the ULB (Université Libre de Bruxelles (Brussels, Belgium)) university hospital.

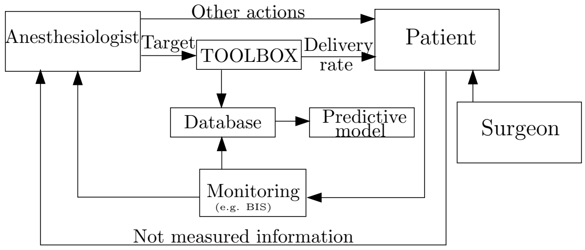

Figure 1. The TOOLBOX software and the anesthesia procedure. TOOLBOX accomplishes two main tasks: (i) servo-controlling of the drugs delivery rate on the basis of the targets fixed by the anesthesiologist and (ii) monitoring and storing in a database the patient signals and the anesthesiologist actions .

closed-loop discrete-time controller is discussed: the inverse/forward approach [1, 7]. The idea of *inverse/forward* control consists in computing the control action by using the information returned both by an inverse and a forward representation of the system's dynamics. The inverse controller relies on a machine learning model which returns a prediction of the change of the Propofol level as a function of the desired BIS value. In this case the predictive model itself acts as a controller since it outputs the control action once the desired BIS is entered as input. Although this technique has been shown to be effective in a number of domains [17, 15] there are some drawbacks (ill-conditioning, lack of inverse relationship) that limit its usefulness.

A possible solution is the coupling of the inverse technique with a forward approach. The idea of *forward control* is inspired by the fact that a forward model can predict the future outputs of the system on the basis of present inputs. Atkeson *et al.* [1] propose the adoption of a forward model initialized with the result returned by the inverse modeling procedure (inverse/forward approach). Our inverse/forward controller relies on a machine learning model which returns a prediction of the next value of the BIS index as a function of the modification of the Propofol level. The control strategy consists in (i) submitting to the predictive model a set of alternative actions (i.e. different Propofol level modifications) in the neighborhood of the action proposed by the inverse controller and (ii) selecting the best one (i.e. the action for which the predictive model returns the closest value to the desired target).

In particular we will focus here on assessing and comparing a version of the inverse/forward approach based on a identified linear model and a second version based on a local learning approach (lazy learning) [7].

The learning procedure is supported by a feature selection to reduce the input dimensionality of the prediction problem. It is well-known in the machine learning community that removing irrelevant and redundant features can dramatically improve the predictive accuracy of learning algorithms [4]. This is still more relevant in the anesthesia domain where there is a large number of potential variables (e.g. the patient age, the surgery type, the phase of the operation) which could influence the value of the BIS signal.

In order to assess our approach we use the historical data to simulate the control strategy in a selection of scenarios where the suggested control action is compared with the anesthesiologist recorded action. The obtained results are promising and worthy of future investigation.

To our knowledge, this is the first study in the literature which aims to learn a closed-loop controller for the BIS index exclusively on the basis of a large amount of measured data. Existing approaches relies on more conventional control schemes or on the combination of first principle compartmental models with black-box identification schemes. In [12] the closed-loop controls the anesthesia form induction to maintenance using evoked potential index as the controlled variable. A PID controller of the BIS index is shown to outperform a manual strategy in [16]. A compartmental model to study the effect of Isoflurane on BIS is studied in [9]. An approach combining neural network and compartmental systems has also been developed in [11]

This paper is structured as follows: Section 2 presents the inverse/forward strategy. Section 3 presents the learning methods adopted to learn the predictive models underlying the two controllers. Section 4 presents and discusses the assessment of the proposed approaches in a data driven simulation framework. The last section contains the conclusion and open issues for future work.

2 The inverse/forward approach

Figure 2 shows an example of the evolution of the BIS index during a short period of time. In this example the patient is a 51 year-old woman of 82 kilogrammes. Note that 536 seconds after the beginning of the operation, the anesthesiologist moves the target of Propofol from $0.5\mu g/ml$ to $2\mu g/ml$ when the target of Remifentanil is equal to $7ng/ml$. As a consequence the BIS moves from an average value 53.2 of in the interval $[506, 536]$ sec to an average value of 43.6 over the interval $[656, 776]$ sec.

The goal of our control architecture is to adjust the concentration of the Propofol drug in order to let the BIS of the patient (the controlled system) attaining the desired level. Suppose that the dynamics of the BIS index can be described by a single-input single-output (SISO) NARMAX (Nonlinear AutoRegressive Moving Average with eXternal input) discrete-time dynamic system

$$b(t+\Delta) = f_d(b(t), p(t), \delta p(t), r(t), a, w) + \epsilon(t) \qquad (1)$$

where, at time t, $b(t)$ is the BIS value, $p(t)$ is the current concentration of Propofol, $\delta p(t)$ is the Propofol concentra-

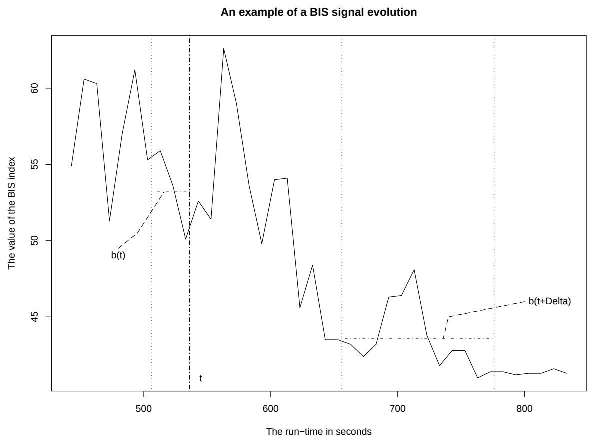

Figure 2. The evolution of the BIS index during a short period of time. At time $t = \mathbf{536}$ sec the anesthesiologist changes the setting of the target of Propofol from $\mathbf{0.5}$ to $\mathbf{2}$.

tion modification (control action) and $r(t)$ is the concentration of Remifentanil. Also, a and w are the age and the weight of the patient, respectively. In our study the time step Δ is set to 120 seconds. In fact, in order to smooth the signal fluctuations, $b(t)$ is the time average of the BIS signal over the interval $[t-30, t]$ and $b(t+\Delta)$ is the time average of the BIS index over the interval $(t+120, t+240)$.

Once we assume that an inverse of the dynamics (1) exists, the control action $\delta p(t)$ can be expressed as the explicit function

$$\delta p(t) = f_i(b(t), b(t+\Delta), p(t), r(t), a, w) + \epsilon(t) \qquad (2)$$

This equation states that there is an unique input value at time t that can drive the output from $b(t)$ to $b(t+\Delta)$.

In a generic nonlinear configuration, the inverse mapping f_i, if it exists, may not have an analytically closed form. Therefore, learning methods can be useful in order to approximate the inverse mapping on the basis of a training set. Assume that the inverse mapping (2) exists and that a sufficiently accurate model $\hat{f}_i(\cdot, \alpha_N)$, with set of parameters α_N, can be learned on the basis of an historical dataset of N samples. The *inverse control* technique consists into computing at each step

$$u_I(t) = \hat{f}_i\left(b(t), b^*, p(t), r(t), a, w, \alpha_N\right) \qquad (3)$$

where b^* is the reference target.

Although this technique has been shown to be effective in a number of domains [17] there are some drawbacks that limit its usefulness. It is well-known in linear control theory that the stability of the inverse control system is guaranteed only in case of minimum-phase systems. Moreover in a generic nonlinear case, if the space of control actions has a different dimensionality than that of the output, the inverse model can be ill defined. Finally, if the input/output relation is characterized by a many-to-one mapping, then the inverse modeling technique is unable to find an inverse [3].

A possible solution is the coupling of the inverse technique with a forward approach. The idea of *forward control* is inspired by the fact that a forward model $\hat{f}_d(\cdot, \alpha_N)$ of the dependency (1) can predict the future outputs of the system on the basis of present inputs. It therefore seems reasonable to turn this around and ask what control action at the present instant of time would bring the future output nearest to the desired value.

Atkeson *et al.* [1] propose the adoption of a forward model as a way to perform a numerical inversion of the plant. This requires searching in all the possible control actions, the one which produces the output minimizing the distance to the reference. The search routine can be initialized with the result (3) returned by the inverse modeling procedure.

In our context the forward model returns, for a given modification of the target of Propofol, a prediction

$$\hat{b}(t+\Delta) = \hat{f}_d(b(t), p(t), \delta p(t), r(t), a, w, \alpha_N) \qquad (4)$$

The principle of the direct controller is simple. Let b^* the target BIS value and δp the control variable. Once the model $\hat{f}_d$ is available and a set U of alternative control actions is fixed, the proposed control action is given by the following minimization procedure

$$\begin{aligned}\delta p(t) &= u_D(t) = \\ &= \arg\min_{u \in \mathcal{U}} \left| \hat{f}_d(b(t), p(t), u, r(t), a, w, \alpha_N) - b^* \right| \qquad (5)\end{aligned}$$

In other terms the control action is the one which minimizes the deviation of the predicted future value of b from the target value b^*.

3 Learning the predictive models

This section describes the learning procedure adopted to estimate the predictive models (3) and (4). Consider a discrete-time dynamic system where the evolution of the variable b can be described by a relation of the form (1) where the mapping f_d is unknown . Let v be a vector containing all the variables apart form b. Suppose we have collected a set of measures $[b(t_i), v(t_i)]$, $i = 1, \ldots, N$, called the *training set*. The prediction problem consists in estimating the future value $b(t+\Delta)$ when the vector $q(t) = [b(t), v(t)]$ (in the following also query point) is taken as input. Learning methods can be used to estimate the function f_d on the basis of a finite training set. The simplest learning approach boils down to a conventional linear identification : $\hat{b}(t+\Delta) = \alpha_N^T q(t)$ where the vector of parameters is estimated on the basis of the observed dataset by using linear regression techniques (e.g. least-squares) [18]. However, when linear identification does not return a sufficiently accurate prediction, the control designer may want to use alternative methods for learning non-linear relationships. This paper adopts a method of local modeling, called lazy learning, which proved to be successful in many problems of non-linear modeling [5] and in two international competitions on data analysis and time series prediction [6]. In local modeling the value of an unknown mapping is estimated focusing on the region surrounding the point where the estimation itself is required. The procedure essentially consists of these steps: i) for each query point $q(t)$, select a set of neighbors and weight their relevance according to some relevance criterion (e.g., the distance) ii) choose a local regression function h in a restricted family of parametric functions iii) compute the regression value $h(q)$. So doing the approach requires keeping in memory the set of observations for each prediction, instead of discarding it as in a global modeling approach (e.g., linear regression). At the same time, local modeling requires only simple approximators (e.g., constant and/or linear) to model the dataset in a neighborhood of the query point. Moreover, the method is intrinsically adaptive, since the availability of new measurements requires simply the updating of the observation set.

Lazy learning is a particular instance of local modeling which provides an automatic way of selecting the optimal number of neighbors for each query point. The idea consists in starting from a minimum number of neighbors and recursively adding neighbors until the predicted performance of the corresponding local approximation decays significantly or until a maximum number of examples is reached. This procedure allows the detection of a linearity region around the query point. For more details on local modeling methods and the distinctive features of lazy learning, we refer the reader to [5].

3.1 Feature selection

A common issue to all the techniques which estimate nonlinear dependencies from observed data is how to deal with the dimensionality of the problem. Large dimensionality is known to make the estimation more badly conditioned and prone to overfitting. A typical solution is to have recourse to feature selection techniques. Examples of feature selection approaches are filter [14] and wrapper techniques [13]. In this paper, we use a wrapper technique where a leave-one-out cross-validation procedure is used to assess the accuracy of the input sets.

The goal was to define which variables are the most relevant for the control task. The selection procedure considered a set composed of the features mentioned in (1) and two more: the type of surgery and the phase of the operation. As far as the forward model of the relation (1) is concerned, the feature selection returned the same set of variables for the two types of models (linear and lazy): $\{b, p, \delta p, r\}$. As far as the model of the inverse relation (2) is concerned, the feature selection returned this set of variables: $\{b, b(t+\Delta), p, r, a\}$.

Note that dimensionality reduction is useful not only in statistical terms but also as a way of returning to the anesthesiologist high-level information about which variables play an important role on the evolution of the patient physiological parameters. For instance, in our case the procedure confirmed the importance of taking into account the titration of Remifetanil in the definition of a control strategy for the BIS.

4 The simulation study

The simulation study is based on a set of data obtained from 329 surgical operations carried out at the "Hôpital Erasme". During each operation, the patient receives Propofol as hypnotic and Remifentanil as analgesic. The flow rates of the drugs are driven by TOOLBOX according to the concentration targets chosen by the anesthesiologist and the well-known pharmacokinetic sets of Schnider (for Propofol) and Minto (for Remifentanil) [19].

The monitored data collected by TOOLBOX are stored in a MySQL [6] dataset. We used the statistical language *R* [7] for the analyses and the package *RMySQL* [8] for the connection to the database. From the database we extracted a learning dataset containing the values of all the variables of interest at $N = 1491$ different time instants. A time instant t_i is selected if the target of Propofol is adjusted by the anesthesiologist at time t_i and when no further modification of the Propofol target takes place in the interval $[t_i - 60, t_i + 60]$ sec.

4.1 The data-driven validation

Two are the most common ways of validating a closed-loop control strategy: either adopting the controller in the real setting, or testing its performance in a simulated environment. In our case the early stage of this work and the evident ethical issues make the first option too premature. At the same time, given the complexity of the controlled system (i.e. the patient) no simulated model would have been convincing enough.

To solve this dilemma we adopted the most commonly used strategy in machine learning when the assessment of the generalization accuracy of a predictive model is at stake:

[6] http://www.mysql.com/
[7] http://www.r-project.org/
[8] http://cran.r-project.org/src/contrib/Descriptions/RMySQL.html

the use of training and test procedure (also known as cross-validation). This consists on using a portion of the data for training our controller and leaving the remaining part for testing. This should be in principle equivalent to activating the controller in a specific configuration and gives us the opportunity to compare the performance of the automatic controller to the performance of the controller reputed at the moments as the most reliable one, that is the human anesthesiologist.

Although this approach does not account for the complexity and the dynamic effects of a real testbed we consider it as a preliminary and safe way to assess the property of an automatic controller in such a sensible environment.

Three different cross-validation criteria are used to validate the different controllers and all of them relies on the notion of leave-one-out error. The leave-one-out error made by a controller is the quantity : $E^{loo}(t_i) = \delta p(t_i) - u^{(-i)}(t_i)$ where $u^{(-i)}$ is the action returned by a controller trained on all the data apart from the ones concerning the instant t_i and $\delta p(t_i)$ is the action (increment) taken by the anesthesiologist at time t_i.

The first criterion is the $NMSE$ (*Normalized Mean Squared Error*):

$$NMSE = \frac{\sum_{i=1}^{N} \left(E^{loo}(t_i)\right)^2}{\sum_{i=1}^{N} \left(\delta p(t_i) - \widehat{\mu}_p\right)^2} \tag{6}$$

where $\widehat{\mu}_p = 1/N \sum_{i=1}^{N} \delta p(t_i)$. Note that if $NMSE > 1$ than we may interpret it by the fact that the control prediction error is worser than the control error that we would obtain if the average of the target modification $\widehat{\mu}_p$ would be used as controller.

The second criterion is the mean MAE of the absolute errors

$$MAE\,(U) = \frac{1}{N} \sum_{i=1}^{N} \left|E^{loo}(t_i)\right| \tag{7}$$

This measure has the same dimension of the δp variable and gives an indication of the average size of the errors made by the controller.

The last criterion is the percentage P of cases that the target modification returned by the controller has the same sign as the modification proposed by the anesthesiologist

$$P = \frac{100}{N} \sum_{i=1}^{N} I\left[\delta p(t_i) \cdot u^{(-i)}\right] \tag{8}$$

where $I\,[A] = \begin{cases} 1 & \text{if } A \geq 0, \\ 0 & \text{if } A < 0. \end{cases}$

This criterion returns a measure of the frequency with which the controller suggests a strategy similar to the one adopted by the anesthesiologist.

Inverse controllers			
model	$NMSE$	MAE	P
linear	0.67	0.28	75.7
lazy	0.65	0.26	75.7

Table 1. Inverse controller.

Forward controller			
model	$NMSE$	MAE	P
linear	0.70	0.30	72.43
lazy	0.67	0.28	72.1

Table 2. Inverse/forward controller.

4.2 Results

The experimental session compares the accuracy of the inverse and the inverse/forward approach (see section 2) where two types of prediction models (a linear model and a lazy model) are used. In the inverse/forward formulation (Equation 5) we limit to consider a small set U composed of the values $[\delta p - 0.1, \delta p, \delta p + 0.1]$ where δp is the outcome of the inverse approach.

For each control approach and for each model type, we report the values of the criteria defined in section 4.1 for the set of variables returned by the feature selection procedure (Section 3.1).

The figures in Table 1 show that the inverse approach returns an NMSE significantly lower than one and that 75% of the time the sign of the modification of the Propofol titration returned by the controller coincides with the one chosen by the anesthesiologist. However, the values of table 2 show that the forward approach does not bring any improvement to the inverse strategy. Also, we note that in both the approaches (inverse and inverse/forward), the accuracy of the control action is improved by the use of the lazy-learning as a learning algorithm.

Finally, it is worth noticing to remark that the amount of dosage proposed by the learned controller is slightly lower and sensibly less standard deviation than the one of the anesthesiologist (see Table 3).

	$anesth.$	$best\ controller$
mean of δp	2.053	2.0434
sd of δp	0.361	0.188

Table 3. Mean and standard deviation of the Propofol modification

5 Conclusion and future works

This paper proposes the use of machine learning techniques to learn closed-loop controllers for supporting the activity of anesthesiologists during surgical operations. The results of the inverse approach appear to be promising and ask for more convincing validation in realistic clinical settings. We expect however that, on a shorter horizon, the utility of the approach could be appreciated also in out of the loop configurations. In fact, anesthesiologists still lack of decision support tools able to suggest, validate and confirm course of actions during daily practice. We advocate that adaptive expert systems able to learn from historical data, once made available in the surgical block, could play a major role in scenarios where anesthesia procedure are performed either in emergency situations or by inexpert personnel.

References

[1] C. G. Atkeson, A. W. Moore, and S. Schaal. Locally weighted learning for control. *Artificial Intelligence Review*, 11(1–5):75–113, 1997.

[2] J. M. Bailey, W. M. Haddad, and T. Hayakawa. Closed-loop control in clinical pharmacology: Paradigms, benefits and challenges. In *Proceedings of the 2004 American Control Conference*, pages 2268–2277, 2004.

[3] C. M. Bishop. *Neural Networks for Statistical Pattern Recognition*. Oxford University Press, Oxford, UK, 1994.

[4] A. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97:245–271, 1997.

[5] G. Bontempi. *Local Learning Techniques for Modeling, Prediction and Control*. PhD thesis, IRIDIA-Université Libre de Bruxelles, 1999.

[6] G. Bontempi, M. Birattari, and H. Bersini. Lazy learners at work: the lazy learning toolbox. In *Proceeding of the 7th European Congress on Intelligent Techniques and Soft Computing EUFIT '99*, 1999.

[7] G. Bontempi, M. Birattari, and H. Bersini. Lazy learning for modeling and control design. *International Journal of Control*, 72(7/8):643–658, 1999.

[8] F. Cantraine and E. Coussaert. The first object oriented monitor for intravenous anesthesia. *PubMed - indexed for MEDLINE*, 16:3–10, 2000.

[9] A. Gentilini, C. W. Frei, A. H. Glattfedler, M. Morari, T. J. Sieber, R. Wymann, T. W. Schnider, and A. M. Zbinden. Multitasked closed-loop control in anesthesia. *IEEE Engineering in Medicine and Biology*, pages 39–53, 2001.

[10] B. Guignard, C. Menigaux, X. Dupont, D. Fletcher, and M. Chauvin. The effect of remifentanil on the bispectral index change and hemodynamic responses after orotracheal intubation. *Anesth Analg*, 90:161–168, 2000.

[11] T. Hayakawa, W.M. Haddad, N. Hovakimyan, and J. M. Bailey. Neural network adaptive dynamic output feedback control for nonlinear nonnegative systems using tapped delay memory units. In *Proceeding of the American Control Conference*, 2004.

[12] G.N. Kenny and H. Mantzaridis. Closed-loop control of propofol anaesthesia. *Anaesth*, 83:223–8, 1999.

[13] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324, 1997.

[14] D. Koller and M. Sahami. Toward optimal feature selection. *International Conference on Machine Learning*, pages 284–292, 1996.

[15] M. Kuperstein. Neural model of adaptive hand-eye coordination for single postures. *Science*, 239:1308–1311, 1988.

[16] N. Liu, T. Chazot, A. Genty, A. Landais, A. Restoux, K. McGee, P.A. Laloe, B. Trillat, L. Barvais, and M. Fischler. Titration of propofol for anesthetic induction and maintenance guided by the bispectral index : Closed-loop versus manual control. a prospective randomized multicenter study. *Anesthesiology*. submitted.

[17] W. T. Miller. Sensor-based control of robotic manipulators using a general learning algorithm. 3:157–165, 1987.

[18] R. H. Myers. *Classical and Modern Regression with Applications*. PWS-KENT, Boston, MA, 1990.

[19] C.S. Nunes and T.F. Mendona. Modelling propofol and remifentanil pharmacodynamic interaction using two different pharmacokinetic models: a comparative study. *Cardernos de Matematica*, 1:46, 2005.

[20] J.C. Sigl and N.G. Chamoun. An introduction to bispectral analysis for the electroencephalogram. *Clin Monitor*, 10:392–404, 1994.