Proceedings of the Sixth IASTED International Conference
BIOMEDICAL ENGINEERING
February 13–15, 2008, Innsbruck, Austria
ISBN Hardcopy: 978-0-88986-721-5 / CD: 978-0-88986-722-2

TIME-SERIES ANALYSIS TECHNIQUES COMBINED WITH GAUSSIAN PROCESS CLASSIFIERS FOR PREDICTION OF CLINICAL STABILITY AFTER CORONARY BYPASS SURGERY

F. Güiza, H. Blockeel, M. Bruynooghe
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200a B-3001
Leuven, Belgium
Fabian.Guiza@cs.kuleuven.be

K. Van Loon, J.-M. Aerts, D. Berckmans
Division Measure, Model & Manage Bioresponses
Katholieke Universiteit Leuven
Kasteelpark Arenberg 30 B-3001
Leuven, Belgium
Kristien.Vanloon@biw.kuleuven.be

G. Meyfroidt, G. Van den Berghe
Department of Intensive Care Medicine
University Hospital Gasthuisberg
Herestraat 49 B-3000
Leuven, Belgium
Geert.Meyfroid@uz.kuleuven.be

ABSTRACT
In this study we present a combination of time-series analysis tools and a machine learning algorithm (Gaussian Process classifier) for the task of predicting the time frame in which the minimal clinical conditions of stability to start weaning of mechanical ventilation are reached. We perform a retrospective analysis of clinical data obtained from a Patient Data Management System of 103 elective coronary bypass surgery patients. Four hours of ICU data of 14 physiological variables, was used as input for five different time-series analysis models. A Gaussian Process Classifier, with the parameters of the calculated models as inputs, assigned to each patient a probability of belonging to the defined classes for clinical stability: within the first 8 hours, between 8 and 16 hours, between 16 and 24 hours, and after 24 hours. Including parameters of different types of time-series models as a representation of the time-varying signals, we incorporate knowledge of the dynamical behaviour of the patients. As a result we obtained aROCs above the medical requirements of 0.8 for some of the classes and above 0.7 for all classes. The use of the dynamics captured by the model representations led to increased performance in further ahead predictions.

KEY WORDS
Intensive Care, Time Series Analysis, Machine Learning, Gaussian Process Classifier

1. Introduction

In cardiac surgery, optimal use of intensive care unit (ICU) and operating room (OR) capacity requires the prediction of future availability of ICU beds. On the level of the management of the department, a number of beds are reserved for cardiac surgery patients. Each of these patients needs intensive care for at least a few hours immediately after surgery. Sometimes the clinical condition of these patients necessitates a longer stay in the ICU of multiple days, weeks or even months. It is clear that these patients occupy a large portion of the available ICU capacity. Before a patient can be discharged from the ICU, he has to be weaned off the mechanical ventilation. If they are still ventilated, they cannot be sent to a normal ward, the bed does not become available and the surgeon cannot operate on new patients. So in order to manage the planning of the intensive care unit and the operating theatre, it would be of use to have a system that provides an early alert if there is a high probability that a patient will or will not be disconnected from ventilation during the next day. The respiratory weaning protocol in the studied ICU defines the following minimal conditions, before the sedative drugs administered to the patients can be stopped: hemodynamic and respiratory stability, absence of bleeding and normothermia. Many cardiac intensivists experience that the trends of the vital parameters during the first hours of ICU stay may be just as important as age, comorbidities and type and duration of surgery to predict a short or prolonged length of stay from early on. The health status of every critically ill patient varies with time. Time-series models can capture the dynamics of different physiological variables describing the evolution of an individual patients state. Machine learning algorithms on the other hand can analyse data from a collection of patients and can be trained to make predictions on new previously unseen patients. Here we combine the advantages of both approaches for the selected prediction task, where the input data for the machine learning algorithm is obtained from the time-series model representations of the patients state. Including parameters of different

types of time-series models as a representation of the time-varying signals, we incorporate knowledge of the dynamical behaviour of the patients, which we show leads to better predictive performances. The objective of this preliminary study is to test whether time-series analysis tools, combined with a machine learning algorithm (Gaussian Processes Classifier), allow the prediction of the timeframe in which the minimal clinical conditions to start weaning of the mechanical ventilation are reached, based on standard measured clinical variables in the first hours after ICU admission.

2. Materials and Methods

2.1 Data Generation

An ICU is a very data rich environment. Patients are connected to patient monitors, gathering and displaying physiological measurements such as heart rate, blood pressures, temperature, cardiac output, etc. Therapeutic devices such as the patient's mechanical ventilator, or infusion pumps used to administer drugs and intravenous fluids, are also a source of data. Laboratory analysis of the patient's blood and urine is performed in a central laboratory or at bedside by point-of-care testing. Doctors, nurses or other health care professionals chart other data manually. A Patient Data Management system (PDMS), is a software package where all these patient related data can be collected and stored in one patient file, regardless of the data source or sampling rate [1]. The data used in this study was obtained from the Patient Data Management System (Metavision®, iMD-Soft®) of the ICU of the university Hospital Leuven, Belgium. A copy of the original PDMS database was created for the purpose of this study. For reasons of privacy, all data referring to the identity of the patients was removed from this database copy. Ethical Committee approval for this retrospective non-interventional study was obtained, and the need for informed consent was waived. A total of 103 non-urgent (elective) coronary bypass surgery patients were selected from this database. For each of these patients, 14 physiological variables were used as inputs for developing the models in this study (Table 1). Most variables were stored every minute from the patient monitor or respirator while blood-loss was manually recorded approximately once every hour and blood gas analysis was performed at least once every four hours (providing the partial oxygen tension in arterial blood ($\mathrm{PaO_2}$) and Serum Lactate measurements).

2.2 Modeling Analysis

2.2.1 Time series analysis techniques

MULTIVARIATE AUTOREGRESSIVE MODELS: A time series is a sequence of observations taken sequentially in time. Most time series consist of elements that are serially dependent. This means that a coefficient or a set of coefficients can be estimated that describes consecutive elements of the series from previous elements [2]. The general equation of a multivariate autoregressive model (MAR) can be written as

$$Y(t) = \sum_{m=1}^{M} A(m)Y(t-m) + E(t) \tag{1}$$

Every observation is made up of a linear combination of M prior observations (the order of the model) and a white noise term. $Y(t) = [y_1(t), ..., y_K(t)]'$ is the vector of simultaneous measured values at time t for K variables and $E(t) = [e_1(t), ..., e_K(t)]'$ is a prediction error vector. The matrices $A(m)$ are the MAR coefficients and are estimated using a stepwise least squares algorithm. In this study, the coefficients of matrix A are used for further analysis.

MULTIPLE INPUT/MULTIPLE OUTPUT AR MODELS: The relationship between inputs and outputs of a given system can be modeled as follows [3]:

$$Y(t) = \sum_{m=1}^{M} A(m)Y(t-m) + \sum_{n=1}^{N} B(n)U(t-n) + E(t) \tag{2}$$

where $Y(t) = [y_1(t), ..., y_K(t)]'$ is the vector of K simultaneously measured outputs at time t, similarly $U(t) = [u_1(t), ..., u_L(t)]'$ is the vector of simultaneously measured inputs, and $E(t) = [e_1(t), ..., e_K(t)]'$ is a prediction error vector. This model is called an ARX model, where AR refers to the autoregressive component and X to the extra exogenous input U. The coefficients of $A(m)$ and $B(n)$ are also computed using a stepwise least squares estimation method. The $K \times K$ coefficients of the estimated A matrix and the $K \times L$ coefficients of the estimated B matrix are used in further analysis.

CEPSTRAL COEFFICIENTS: Cepstrum analysis is a nonlinear signal processing technique. The cepstrum is defined as the inverse Fourier transform of the short-time logarithmic amplitude spectrum [4] and can be used to analyze time series data. The distance between the cepstral coefficients of different time series can be used as a similarity measure between these time series. On the basis of autoregression coefficients from linear models the linear predictive coding (LPC) cepstrum can be calculated. The LPC cepstral coefficients for an univariate time series of length N can be obtained from the autoregressive coefficients from equation 1 as follows [4]:

$$c_n = \begin{cases} a_1 & \text{if } n = 1 \\ a_n + \sum_{m=1}^{n-1} \left(1 - \frac{m}{n}\right) a_m c_{n-m} & \text{if } 1 < n \leq \mathbf{M} \\ \sum_{m=1}^{\mathbf{M}} \left(1 - \frac{m}{n}\right) a_m c_{n-m} & \text{if } \mathbf{M} < n \end{cases} \tag{3}$$

The parameters a_i are the coefficients of the calculated MAR models. The $c_i\ (i = 1, ..., N)$, are the cepstral coefficients.

Table 1. Physiological Variables

VAR	PHYSIOLOGICAL VARIABLE	UNIT	SAMPLING FREQUENCY
1	Heart rate	bpm	1 / min
2	Arterial Blood Pressure, systolic	mmHg	1 / min
3	Arterial Blood Pressure, diastolic	mmHg	1 / min
4	Arterial Blood Pressure, mean	mmHg	1 / min
5	Pulmonary Artery Pressure, systolic	mmHg	1 / min
6	Pulmonary Artery Pressure, diastolic	mmHg	1 / min
7	Pulmonary Artery Pressure, mean	mmHg	1 / min
8	Central Blood Temperature	°C	1 / min
9	Peripheral Skin Temperature	°C	1 / min
10	Positive End Expiratory Pressure (PEEP)	mbar	1 / min
11	Fraction of Inspired Oxygen (FiO_2)		1 / min
12	Partial Oxygen Tension in Arterial Blood (PaO_2)		1 / 4 h
13	Serum Lactate	mmol/L	1 / 4 h
14	Blood loss	mL	1 / h

2.2.2 Gaussian Processes for Classification

Given a training set $\{\mathbf{X}, \mathbf{t}\}$ comprised of N training input vectors $(\mathbf{X} = \mathbf{x}_1, ..., \mathbf{x}_N)$ and their corresponding N binary class labels $(\mathbf{t} = t_1, ..., t_N)$, such that $t_i = +1$ if $\mathbf{x}_i$ belongs to a given class C and $t_i = -1$ if $\mathbf{x}_i$ does not belong to the class. In a probabilistic binary classification task the objective is to determine for an unlabeled test input vector $(\mathbf{x}_*)$ the probability of belonging to the class C: $\pi_C(\mathbf{x}_*) = p(t_* = +1|\mathbf{x}_*)$. From this, the probability of it not belonging to the class can also be computed: $p(t_* = -1|\mathbf{x}_*) = 1 - \pi_C(\mathbf{x}_*)$. In the remainder of this text the input vectors to the classifiers $(\mathbf{X})$ will be referred to as *examples*.

Gaussian Processes, a type of kernel method, are machine learning techniques that have been successfully used to model and forecast real dynamic systems because of their flexible modeling abilities and their high predictive performances. They allow multi-dimensional inputs and they assign a confidence value to their predictions. The main advantage of using a Gaussian Process Classifier over other kernel method classifiers is that it produces an output with a clear probabilistic interpretation [5]. In Gaussian Process binary classification, a Gaussian Process over a function $f(\mathbf{x})$ is defined and then transformed through a logistic function $\sigma()$ so that its outputs lie in the $[0, 1]$ interval. This way they can be interpreted as probabilities: $\pi_C(\mathbf{x}_*) = p(t_* = +1|\mathbf{x}_*) = \sigma(f(\mathbf{x}))$. Conditioning the predictive distribution $\pi_C()$ on the training data allows for a probabilistic prediction on a test input example [6]. The Gaussian Process prior over the function $f(\mathbf{x})$ takes the form $f \sim N(\mathbf{0}, \mathbf{K}(\mathbf{X}, \mathbf{X}))$, with a zero mean function and a covariance function given by a positive semi-definite kernel function $k(\mathbf{x}_i, \mathbf{x}_j)$. The kernel function thus determines the similarity between the input examples $\mathbf{x}_i$ and $\mathbf{x}_j$. Inference of the predictive distribution requires the solution to integrals which are analytically intractable, a problem that is solved either by resorting to Monte Carlo sampling or analytical approximations to the integrals. In this study we follow the latter approach through the use of expectation propagation [5].

2.3 Protocol

The task considered is the prediction of the time frame in which the patients fulfill the clinical criteria for stability that will lead to weaning from mechanical ventilation. Patients are considered stable if they satisfy the following criteria:

- Hemodynamical stability:
  - Dobutamine rate $\leq$5 μg/kg/min
  - Noradrenaline rate $\leq$ 0,2 μg/kg/min
  - Serum lactate $<$ 2 mmol/L
- Respiratory stability:
  - $PaO_2 \geq$ 75 mmHg
  - $FiO_2 \leq$ 0,5
  - PEEP $\leq$ 8 mbar
- Temperature stability:
  - Blood temperature $>$ 36 °C
  - Peripheral skin temperature $>$ 30 °C
- Blood loss stability:
  - Sum of blood loss of all drains $<$ 100 ml/h

Dobutamine is a drug used to support the function of the heart by increasing heart rate and forcing contractions. Noradrenaline is a drug used to support blood pressure by increasing vascular tone. Serum Lactate is a waste product

of metabolism in the absence of oxygen and is used as a marker for insufficient tissue perfusion.

These criteria should be met for at least 30 consecutive minutes in order for patients to be considered sufficiently recovered from their cardiac surgery so as to start the weaning from mechanical ventilation. To enable future comparisons with predictions performed by intensivists, the considered task is restated as follows: Predict in which of the following time frames (classes) the patient will begin to satisfy the stability criteria.

Class 1: earlier than 8 hours

Class 2: between 8 and 16 hours

Class 3: between 16 and 24 hours

Class 4: later than 24 hours

A binary probabilistic classifier is learned for each of the four classes. In the test cohort, a patient is assigned to the class for which it has the largest probability. Training examples for each classifier are labeled positive $(t = +1)$ if the moment when the patient becomes stable starts within the corresponding time interval and are labeled negative $(t = -1)$ otherwise.

Prior to the analysis some preprocessing was performed on the data obtained from the PDMS. This included the removal of outliers and the filling-in of missing values (a rare occurrence) through linear interpolation.

Data from each patient collected at different moments during ICU stay is used to generate the different time-series models. The parameters of these models are used in a next step as training examples for the classifiers. The data from the first 4 hours of ICU stay is used to generate a first example, to which the appropriate class label is assigned as is shown in Figure 1. A second example is generated from the same patient by sliding the 4-hour window of used data (the gray area of Figure 1) to 30 minutes after admission, the class-membership thresholds (vertical dashed lines) are displaced by 30 minutes accordingly, but the moment when the stability criteria are met (solid vertical line) remains the same. Further training examples are generated by sliding the 4-hour window (and thresholds) in 30-minute steps, and assigning to each the appropriate class label. The last training example that can be generated from one patient uses the 4-hour window of data that ends just before the moment of stability takes place.

The different time-series analysis techniques described above, were applied to each of the 4-hour intervals of data in order to generate the examples used as inputs for the Gaussian Process Classifier. The interval duration of four hours was chosen since time intervals of shorter duration led to non-stable time-series models. The types of examples (input vectors) used to train the Gaussian Process classifiers in our experiments are explained below.

1. Signal Average: Each example is a 56 dimensional vector containing four values for each of the 14 physiological variables of Table 1. Each of the four values is an average over a one-hour interval of the measured time series signal.

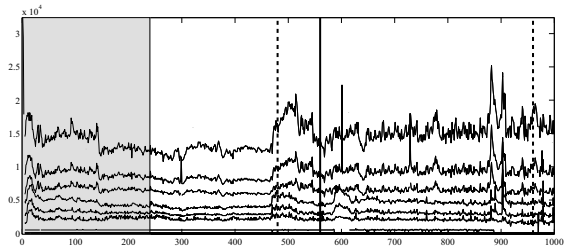

Figure 1. Gray area corresponds to 4-hour interval of data used to generate the example. Dashed vertical lines depict the 8 and 16 hour class-membership thresholds and the solid vertical line indicates the moment when the patient satisfies the stability criteria. The example generated from this data is labeled as belonging to Class 2.

2. MAR coefficients: Each example is a 64 dimensional vector containing the a_{ij} coefficients of the first-order MAR model for eight variables of one sample per minute, namely signals 1 to 8 from Table 1. Signals 9 to 10 are excluded from this and the remaining time-series models because their low sampling rate or slow varying dynamics did not yield stable models.

3. ARX coefficients: Each example is a 48 dimensional vector containing the coefficients of the A and B matrices of an input/output model with two inputs (variables 1 and 8 from Table 1) and six outputs (variables 2 to 7 from Table 1). Taking human physiology into account, the input-output selection was deemed relevant for this study.

4. Cepstral coefficients (CEP): Each example is an 80 dimensional vector containing the 10 first cepstral coefficients of variables 1 to 8 from Table 1 computed directly from the time-series.

5. Cepstral coefficients from ARX models (CEPARX): Each example is an 80 dimensional vector containing the 10 first cepstral coefficients of variables 1 to 8 from Table 1 using equation 3.

What follows is a description of the procedure to compute the predicted probabilities of belonging to each of the 4 classes for each patient. This procedure is repeated for each type of time-series model previously described.

All examples generated for all N patients from one type of time-series model and their corresponding class labels are collected in one dataset. From this dataset those examples generated from patient P_i are removed and the one generated from the first 4 hours of ICU stay is used as test example $(\mathbf{x}_{*i})$. The data from the remaining patients is sampled such that there are an equal number of positive and negative labeled examples, and will be used as training set $\{\mathbf{X}_i, \mathbf{t}_i\}$ for the Gaussian Process classifier. Once the classifier has been trained, the predicted probability of belonging to a class C, $\pi_C\left(\mathbf{x}_{*i}\right)$ is recalibrated to compensate for the effects of learning on the artificially balanced training set so that it applies to the original (non-sampled) distribu-

Table 2. Classifier Performance for Different Models

INPUTS	CLASS 1	CLASS 2	CLASS 3	CLASS 4
SIGAVG	**0.81**-0.58	0.57-0.56	0.62-0.55	0.67-0.61
MAR	0.64-0.51	0.59-0.50	0.55-0.51	0.66-0.52
ARX	0.60-0.50	0.57-0.51	0.68-0.60	0.62-0.50
CEP	0.66-0.50	**0.72**-0.56	0.57-0.58	0.66-0.58
CEPARX	0.73-0.55	0.64-0.52	**0.85**-0.66	**0.70**-0.63

tion. The described process is repeated for each patient P_i where $i = 1, ..., N$, and for each class $C \in 1, 2, 3, 4$ so that a probability of belonging to each class is assigned to each of the N patients. If a hard-classification is required, each patient is assigned to the class for which it has the highest probability. The obtained probabilities allow for the computation of an aROC for the classifier of each class.

Sampled examples are used to generate the training set since it has been observed that learning from a balanced dataset typically leads to more accurate models avoiding over-fitted solutions that occur when learning from heavily skewed distributions [6].

The kernel function used in the study is the squared exponential with ARD (automatic relevance determination) defined as follows:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{M}(\mathbf{x}_i - \mathbf{x}_j)\right) \quad (4)$$

where $\mathbf{M} = \mathrm{diag}(l)^{-2}$ is a diagonal matrix and its diagonal elements $l_1, \ldots, l_D$ are characteristic length-scales for each dimension of the input examples. Recall that each example $\mathbf{x}$ corresponds to a vector obtained from the different time-series models. The values of the parameters of the diagonal matrix $\mathbf{M}$ determine the relevance of the corresponding input dimension. The σ_f^2 parameter is the standard deviation of the process, which controls its magnitude. During training, the parameters $\boldsymbol{\theta} = \left\{\sigma_f^2, l_1, \ldots, l_D\right\}$ are iteratively updated according to the expectation propagation algorithm so as to maximize the likelihood of the class labels given the training data [5].

3. Results

The leftmost number in each entry of Table 2 corresponds to the aROC (area under the receiver operating characteristic curve) of the corresponding Gaussian Process probabilistic binary classifier for each of the 4 classes. The rightmost number is the aROC obtained when using a logistic regression model [7], included here as a baseline for performance.

The Automatic Relevance Discrimination feature of the selected kernel function for the Gaussian Process classifier, allowed us to select the 10 most relevant input dimensions for each classifier. Using only these most relevant input dimensions, new classifiers were learned. This decrease in the number of dimensions used reduces the computational time required for the models to be learned. Their aROCs are shown in Table 3.

Table 3. Classifier Performance for Most Relevant Dimensions for Different Models

INPUTS	CLASS 1	CLASS 2	CLASS 3	CLASS 4
SIGAVG	**0.77**-0.66	0.65-0.63	0.54-0.54	0.67-0.50
MAR	0.61-0.56	0.55-0.53	0.73-0.50	0.66-0.55
ARX	0.56-0.54	0.60-0.53	0.80-0.67	0.62-0.52
CEP	0.66-0.53	0.65-0.55	0.65-0.56	0.66-0.62
CEPARX	0.55-0.55	**0.69**-0.68	**0.83**-0.67	**0.70**-0.62

Table 4. Classifier Performance for Combinations of Most Relevant Dimensions for Different Models

INPUTS	CLASS 1	CLASS 2	CLASS 3	CLASS 4
SIGAVG MAR	0.76-0.62	0.64-0.65	0.74-0.54	0.60-0.58
SIGAVG ARX	**0.77**-0.61	**0.70**-0.67	0.70-0.68	**0.74**-0.53
SIGAVG CEP	0.75-0.60	0.68-0.60	0.82-0.62	0.63-0.57
SIGAVG CEPARX	0.72-0.58	0.68-0.68	**0.84**-0.58	0.64-0.60

Combining the 10 most relevant dimensions of 2 types of time-series models led to 20-dimensional examples with performances as shown in the entries of Table 4.

4. Discussion

These initial results, with aROCs below the desired 0.8 for medical standards of discrimination, for classes 2 and 4, corroborate the difficulty of the selected prediction task. Depending on the interval when the patient becomes stable, different inputs are more predictive for the task. For example, the further ahead the stability criteria are met, the higher the relevance of the dynamics of the physiological variables for prediction. This can be seen in the high performance of the CEPARX-input classifier for the third class (Table 2). For the prediction within the first 8-hour interval, the Signal Average model results in a high performance with an aROC above 0.8 (Table 2), revealing that for early on predictions the dynamics are not as predictive as the actual signal values. The results in Table 3 closely follow those of Table 2 but the overall performance of the classifiers decreases because of the information-loss from the discarded dimensions. The information reduction does not have such a strong impact on the CEPARX for

class 3, revealing that the information contained in the selected significant cepstrum coefficients is still predictive. Different models can be combined to yield learning examples that contain information of both the absolute signal measurements and their dynamics. The results of a first attempt are shown in Table 4. Using the information of the complete models would result in examples of prohibitively large dimensions in terms of computation time, therefore only the 10 most relevant dimensions of each model were used. When compared to their single-model counterparts (Table 3) the combined models result in an overall increase in performance, although only for Class 4 do we obtain a combined-model that results in a higher aROC. As a general remark, it can be seen that the Gaussian Process classifier outperforms the Logistic Regression classifier. For some cases however, such as for Class 2 in table 4, the results of both classifiers are nearly identical indicating that there is not sufficient information in the representations for the Gaussian Classifier to exploit its non-linear flexibility to improve on the performance of a linear model. Analysis of the dimensions that are found to be more relevant for the different models could potentially lead to discovery of domain knowledge in the prediction task. The cepstral coefficients that are most predictive for classes 1, 2 and 3, both in the CEP and in the CEPARX case, are derived from the heart rate signal. Other relevant coefficients in these first three classes differ, with blood temperature being the least relevant. For the prediction of the 4th class, the systolic arterial blood pressure is more relevant than the heart rate for the CEP case; while for the CEPARX case, arterial and pulmonary pressures are the most relevant and none of the heart rate coefficients appears as one of the 10 most relevant. Amongst the physiological variables found to be most relevant for prediction of class 1 with the Signal-Average input model were $\mathrm{PaO_2}$, $\mathrm{FiO_2}$ and blood loss. These signals were excluded from the remaining time-series analysis because they led to unstable models, and their absence could explain their poor performance in predicting the first class.

5. Conclusion

In this study we have shown a first step at automatically determining the future course of elective coronary bypass surgery patients, by predicting the moment when the clinical stability criteria is met that results in weaning from mechanical ventilation. Prediction was performed using data only from the first 4 hours of a patients ICU stay. The results show the complexity of the prediction task with different physiological variables and different representations of their dynamics becoming more relevant depending on the moment when the stability criteria are met. Slow-varying signals with low sampling rates appear more predictive for patients that meet the stability criteria early on, while the dynamics of various signals appear more predictive when stability occurs at later stages of the patients ICU stay. We obtained aROCs above 0.7 for all four predicted classes, and for two of them, above the medical desired value of 0.8.

6. Future work

The development of a definite model for the prediction of respiratory weaning, would benefit from including other non-dynamic variables that are known to be predictive for ICU length of stay after cardiac surgery. To improve on the generalization capabilities of the classifiers it would also be of use to increase the number of patients used during training. This increase both in the number of physiological variables and patients will however require more complex implementations of the algorithms presented such that they are able to cope with the data increase while still remaining computationally tractable. Possible variants of the Gaussian Process classifier include the use of sparse methods, aggregation, dimensionality reduction techniques and the inclusion of more specialized kernels that better incorporate the available prior knowledge. The information obtained from the calculations done in this research should be combined in one algorithm that predicts in which class a given patient belongs. In order to do so, also other time series analysis that describe relationships between different variables could be included if they are found to increase predictive performance.

References

[1] R.H. Ireland, H.V. James, M. Howes, Design of a summary screen for an ICU patient data management system, *Med Biol Eng Comput*, 35, 1997, 397-401.

[2] G.E. Box, G.M. Jenkins and G.C. Reinsel, *Time Series Analysis: Forecasting and Control* (New Jersey: Prentice-Hall International, 1994).

[3] L. Ljung, *System Identification: Theory for the User* (New Jersey: PTR Prentice-Hall, 1987).

[4] K. Kapalkis, D. Gada, V. Puttagunta, Distance measures for effective clustering of ARIMA time-series, *Proc. IEEE International Conf. on Data Mining*, 2001, 273-280.

[5] C.E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (Cambridge: MIT Press, 2006).

[6] C.M. Bishop, *Pattern Recognition and Machine Learning*(Springer, 2006).

[7] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning, Data Mining, Inference and Prediction*(Springer, 2001).

[8] I. Korhonen, R. Takalo, V. Turjanmaa, Multivariate autoregressive model with immediate transfer paths for assessment of interactions between cardiopulmonary variability signals, *Med Biol Eng Comput*, 34, 1996, 199-206.

[9] R.M. Rangayyan, *Biomedical Signal Analysis: A Case Study Approach* (New York: Wiley Interscience, 2002).