



## Prediction of acute hypotensive episodes by means of neural network multi-models

Teresa Rocha<sup>a,\*</sup>, Simão Paredes<sup>a</sup>, Paulo de Carvalho<sup>b</sup>, Jorge Henriques<sup>b</sup>

<sup>a</sup> Instituto Superior de Engenharia de Coimbra, Departamento de Engenharia Informática e de Sistemas, Coimbra, Portugal

<sup>b</sup> CISUC, Centro de Informática e Sistemas da Universidade de Coimbra, Coimbra, Portugal

### ARTICLE INFO

#### Article history:

Received 16 June 2010

Accepted 26 July 2011

#### Keywords:

Hypotension episodes

Blood pressure prediction

Neural network multi-models

MIMIC-II database

### ABSTRACT

This work proposes the application of neural network multi-models to the prediction of adverse acute hypotensive episodes (AHE) occurring in intensive care units (ICU). A generic methodology consisting of two phases is considered. In the first phase, a correlation analysis between the current blood pressure time signal and a collection of historical blood pressure templates is carried out. From this procedure the most similar signals are determined and the respective prediction neural models, previously trained, selected. Then, in a second phase, the multi-model structure is employed to predict the future evolution of current blood pressure signal, enabling to detect the occurrence of an AHE.

The effectiveness of the methodology was validated in the context of the 10th PhysioNet/Computers in Cardiology Challenge—Predicting Acute Hypotensive Episodes, applied to a specific set of blood pressure signals, available in MIMIC-II database. A correct prediction of 10 out of 10 AHE for event 1 and of 37 out of 40 AHE for event 2 was achieved, corresponding to the best results of all entries in the two events of the challenge. The generalization capabilities of the strategy was confirmed by applying it to an extended dataset of blood pressure signals, also collected from the MIMIC-II database. A total of 2344 examples, selected from 311 blood pressure signals were tested, enabling to obtain a global sensitivity of 82.8% and a global specificity of 78.4%.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

Hypotension, a clinical condition characterized by abnormal low blood pressure values, is one of the recurrent situations occurring in intensive care units. Among the most frequent events, acute hypotensive episodes (AHE) are particularly critical, since they may result in irreversible organ damage and, eventually, death [1]. As a consequence, the characterization of such episodes is of fundamental importance in the management of critical ill patients. In fact, when promptly detected, it is possible to improve the clinical decision concerning which intervention is more appropriated for each specific condition (sepsis, myocardial infarction, cardiac arrhythmia, pulmonary embolism, hemorrhage, dehydration, or any of a wide variety of other causes of hypovolemia, insufficient cardiac output, or vasodilatory shock). Additionally, early detection of AHE will give professionals enough time to select a more effective treatment, without exposing the patient to additional risks of delaying therapy. Therefore, the development of methodologies able to detect not only the presence of this condition but also to predict its occurrence, is of

extreme importance concerning appropriated clinical interventions. Moreover, since clinical interventions to treat such events are usually invasive and aggressive, a prediction system that could identify an imminent episode would be a significant benefit to timely support non-invasive and preventive treatments.

It is clinically accepted that if there exists enough patient's clinical information, then a prediction system for hypotensive episodes, over a specific time period, can be developed. Typically, this information is based on the medical record, such as clinical history, laboratory tests and medications, as well as on information extracted from physiologic vital signals, such as electrocardiogram, blood pressure and respiration. In this context, Singla et al. [2] showed the correlation between some independent variables and the development of early hypotension episodes. These variables included age, sex, body mass index, history of hypertension, diabetes mellitus, anemia, heart rate, systolic and diastolic blood pressure. Similarly, Lin et al. [3] studied the association of specific variables with the increasing risk of hypotensive episodes, namely weight, height, American Society of Anesthesiologist physical status, surgical category (orthopedics, plastic surgery, general surgery, obstetrics, and urology) and systolic blood pressure. Based on these variables, Lin et al. proposed a logistic regression model to assess the risk of developing a hypotensive episode.

\* Corresponding author.

E-mail address: [teresa@isec.pt](mailto:teresa@isec.pt) (T. Rocha).

In practice, the development of automatic solutions for hypotensive episodes prediction has explored the correlation between patient's clinical information collected in real-time, such as arterial blood pressure (ABP), heart rate (HR) and oxygen saturation (SO<sub>2</sub>), and the onset of the hypotensive episode. In particular, Bassale [4] proposed the use of parametric and non-parametric methods to analyze and characterize ABP before hypotensive episodes. He concluded that ABP variability and shape features have the potential to predict such events. Crespo et al. [5] also suggested the use of changes in the ABP morphology occurring immediately before an episode of hypotension. They proposed the variance of the ABP signal and the variance of the wave slope as the most relevant features to consider when predicting AHE. Lehman et al. [6] presented a similarity-based searching and pattern matching algorithm, applicable to classification and forecasting tasks. Using real physiological measurements, they employed the methodology to forecast hypotensive episodes in intensive care units. Also, Saeed et al. [7,8] introduced a new temporal similarity metric, based on a transformation of time series data into an intuitive symbolic representation. They used wavelet decomposition to characterize time series dynamics at multiple time scales. Their algorithm was employed to identify similar physiologic patterns in hemodynamic time series from ICU patients, with potential to be used in the detection of imminent hemodynamic deterioration. Frolich et al. [9] suggested the use of baseline HR as a significant predictor of obstetric spinal hypotension. Basically, they showed that higher baseline HR could be a useful parameter to predict postspinal hypotension.

Using spectral analysis of HR and ABP variability, Pelosi et al. [10] identified precursors of hypotensive episodes during renal dialysis. Also using frequency analysis techniques, Reich et al. [11] investigated the correlation between HR variability analysis and hypotension events. Chamchad et al. [12] found a significant correlation between nonlinear HR variability dimension analysis and the presence of hypotension, occurring after spinal anesthesia for cesarean delivery. Hanss et al. [13] also concluded that HR variability analysis could be used to predict the occurrence of hypotension during spinal anesthesia. In particular, they investigated the ratio of low to high frequency peaks of the HR variability power spectrum (LF/HF) for the prediction of hypotension events after spinal anesthesia, in the specific cases of pregnant women [14] and elderly men [15]. Mancini et al. [16] showed the potential of SO<sub>2</sub> short-term variability in anticipating the hypotension onset. Recently, Lee and Mark [17] investigated the existence of discriminatory patterns in ICU data that could be indicative of impending hypotensive episodes. Based on neural, they proposed a binary classification scheme (normotensive vs. hypotensive) and an estimation strategy for predicting future mean blood pressure values.

This work proposes the forecast of acute hypotensive episodes through the development of predictive multi-models, applicable to the mean ABP (MAP) time-series signal. To achieve this goal, a generic methodology consisting of two main phases is considered. In the first phase, a correlation analysis procedure is carried out between the current MAP signal and a representative set of historical MAP evolution trends. The most similar ones are identified and the correspondent prediction neural multi-models, previously trained using those historical signals, selected. In the second phase, these models are employed to the current MAP signal to predict its future evolution and, therefore, the detection of an AHE occurrence. Basically, the prediction methodology consists of a multi-model scheme using neural network structures. Multi-models do not recursively use model outputs as inputs for step ahead predictions. Therefore, prediction errors are not propagated and long-term predictions can be accurately estimated. Among regression models, neural networks have shown considerable capabilities to learn and to generalize from

non-linear environments, enabling to capture the fundamental data dynamics. Moreover, multi-models can be trained by means of simple standard backpropagation algorithms. In fact, since an independent neural sub-model is used for each sampling instant and does not depend on previous predictions, a static training algorithm, as the referred backpropagation, can be employed.

The effectiveness of the proposed approach was validated in the context of 2009 PhysioNet/Computers in Cardiology challenge—Predicting Acute Hypotensive Episodes. The data for training and validation purposes was obtained from MIMIC-II dataset [18] that includes data before and during the prediction horizon. The forecast was made using the trained neural multi-model structure, only considering the information available before the forecast period. The occurrence of an AHE within the forecast window (one hour) was assessed according to AHE definition [19]. A sensitivity of 94.74% and a specificity of 93.55% revealed the effectiveness of the strategy, which obtained the best results of the challenge.<sup>1</sup> Additionally, the generalization ability of the strategy was confirmed by applying it to a larger dataset of blood pressure signals, also belonging to the MIMIC-II database. For a total of 2344 examples, selected from 311 blood pressure signals, a global sensitivity of 82.8% and a global specificity of 78.4% were achieved.

The paper is organized as follows: in Section 2 the 2009 PhysioNet/Computers in Cardiology challenge is presented. In Section 3, it is described the general methodology for the prediction task as well as how it can be used to address the challenge. In Section 4, results using MIMIC-II dataset are presented and discussed, both in the context of the PhysioNet/Computers in Cardiology challenge, as well as considering an extended dataset. Finally, in Section 5, some conclusions are drawn.

## 2. The 2009 physionet/computers in cardiology challenge

### 2.1. Challenge goal

The 2009 challenge was the tenth in the annual series of open challenges hosted by PhysioNet in cooperation with Computers in Cardiology Conference. The goal of that year's challenge was to predict which patients in the available dataset (MIMIC-II) would experience an acute hypotensive episode beginning within the forecast window, motivated by the possibility of improving care and survival of these patients [19].

The challenge dataset included, among other data, a time series of mean arterial blood pressure (MAP) at one-minute intervals. Each sample of the series was an average of the blood pressure measured in the radial artery over the previous minute. Given such a time series, an AHE was defined, for the purposes of the challenge, as any period of 30 minutes or more during which at least 90% of the MAP measurements were at or below 60 mmHg. The forecast window was defined as the one-hour period immediately following a specified time instant  $T_0$  (Fig. 1). In the test sets, all data following  $T_0$  was withheld and the forecast should be made using the available information before  $T_0$ .

### 2.2. MIMIC-II project

Data used in this challenge was collected and contributed to PhysioNet by the MIMIC-II project (*Multi-parameter Intelligent Monitoring for Intensive Care*), a Bioengineering Research Partnership funded by the US National Institutes of Health and its National Institute of Biomedical Imaging and Bioengineering, with additional

<sup>1</sup> <http://www.physionet.org/challenge/2009/final-scores>.

support from Philips Medical Systems. The MIMIC-II project has collected data from about 30,000 ICU patients to date, including recorded physiologic signals and time series, as well as accompanying clinical data such as interventions performed in the ICU, laboratory tests, observations and medication [20,21]. Basically, this information is organized in two main databases: *numeric record* that contains time-series data and *clinical records* that deal with clinical information (observations, medication, etc.). The intent is that a MIMIC-II record should be sufficiently detailed to allow its use in studies that otherwise would require access to an ICU, e.g., for basic research in intensive care medicine, or for development and evaluation of diagnostic and predictive algorithms for medical decision support.

2.3. Training and test datasets

The 2009 challenge dataset consisted of selected patient records from the MIMIC-II database. In the training set, the records included all available data before and after instant  $T_0$ . In the test sets the records were truncated at  $T_0$ , being the data after this instant unknown during the period of the challenge and made available only after its conclusion. The records that were chosen for the challenge dataset included, at a minimum: (i) at least 12 h of data before  $T_0$ , and at least one hour of data after  $T_0$ ; (ii) ECG and arterial blood pressure (ABP) signals sampled at 125 Hz; (iii) time series of vital signs sampled once per minute (in the training set) and once per second (in the test sets). These comprised heart rate and mean systolic and diastolic ABP. The majority of the records included a variety of additional vital signals time series, most often containing respiration rate and oxygen saturation.

In particular, the training set consisted of 60 records (with data after  $T_0$ ) belonging to two groups: H and C. Records in group H contained an episode of acute hypotension beginning during the

forecast window (the one-hour period following  $T_0$ ), while records in group C contained no AHE within the forecast window. Within group H, 15 records belonged to subgroup H1—patients who received pressor medication, and 15 belonged to subgroup H2—patients who did not receive pressor medication. Within group C, 15 records belonged to subgroup C1—patients with no documented AHE at any time during their hospital stay, and 15 records belonged to subgroup C2—patients who had AHE before or after the forecast window.

The validation set consisted of two datasets, A and B. The test set A comprised 10 records, excluding data after  $T_0$ . From these, 5 records were from subgroup H1 (AHE in subjects receiving pressors) and 5 were from subgroup C1 (no AHE in subjects receiving pressors). The test set B consisted of 40 records, also excluding data after  $T_0$ . Between 10 and 16 from these belonged to group H and between 24 and 30 belonged to group C.

2.4. Challenge events

The challenge comprised two events. The event 1, using A dataset, focused on patients who were receiving pressor medication and aimed to distinguish between two groups of ICU patients: the ones who would experience an acute hypotensive episode and those who would not. The event 2, using B dataset, addressed the broad question of predicting an AHE in the general population.

3. Methodology

Fig. 2 gives an overview of the methodology proposed in this work. The detection of a future acute hypotensive episode (AHE) is carried out by means of multi-models trained using mean arterial blood pressure signals (MAP).

The input consists of a MAP signal available before  $T_0$ , the instant where the forecast period starts. From a correlation analysis procedure, between the current MAP signal and a set of MAP templates, representative of historical MAP evolution trends, the most similar templates are identified. After that, the correspondent neural multi-models, previously trained using those historical templates, are selected and employed to predict the future evolution of the current MAP input signal, from instant  $T_0$  until the end of the forecast window. Finally, the occurrence of an AHE is straightforward determined.

This section starts by introducing the general multi-model scheme approach for prediction purposes. Then, the neural network

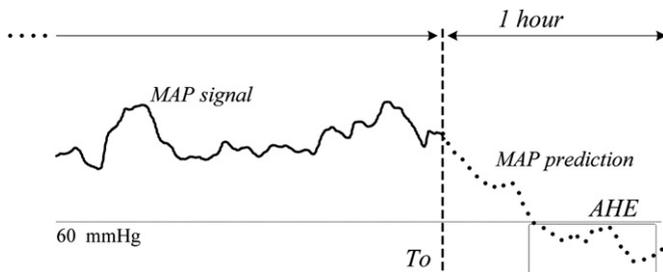


Fig. 1. 2009 PhysioNet—Computers in cardiology challenge goal.

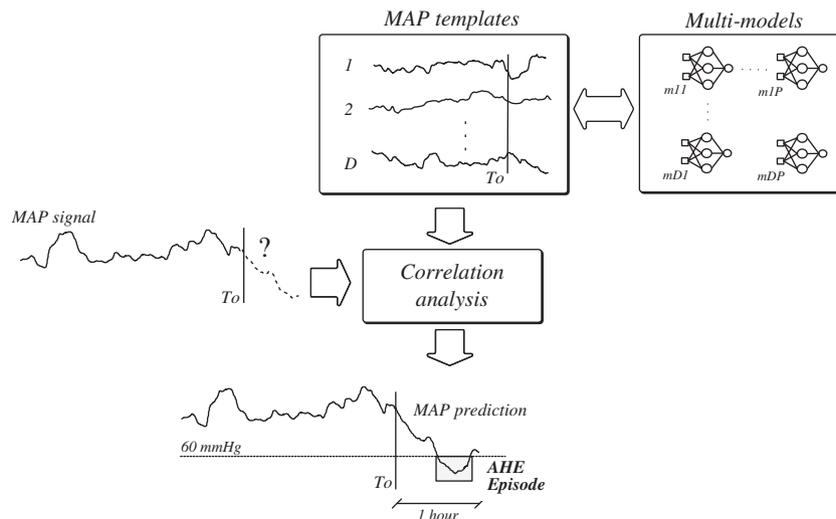


Fig. 2. Proposed neural network multi-model scheme for AHE prediction.

models structure and their incorporation into the multi-model scheme are presented. Finally, the prediction of AHE based on this strategy is addressed.

### 3.1. Multi-models

Regression representations are common techniques for modeling and prediction tasks. By means of autoregressive representations, information from past instants can be used to estimate future values, usually one-step ahead. Consider a single-input single-output (SISO) system described by the following discrete-time nonlinear autoregressive (NAR) model:

$$y(k) = f_1(y(k-1), y(k-2), \dots, y(k-N)) \tag{1}$$

where  $y(k)$  is the measured value (scalar) at instant  $k$ ,  $N$  is the order of the model and  $f_1$  is the model/mapping such that  $f_1: \mathfrak{R}^N \rightarrow \mathfrak{R}$ . At each instant  $k$ , it is assumed the availability of the information  $\varphi(k) \in \mathfrak{R}^{N+1}$  is composed of current and past data ( $N$  instants).

$$\varphi(k) = \{y(k), y(k-1), y(k-2), \dots, y(k-N)\} \tag{2}$$

Using this autoregressive representation, the output at instant  $k+1$ ,  $y(k+1)$ , can be estimated using the available information  $\varphi(k)$  and the mapping  $f_1$ , as follows:

$$\hat{y}(k+1|k) = f_1(y(k), y(k-1), \dots, y(k-N+1)) \tag{3}$$

Following the same approach, the output at instant  $k+2$  can also be estimated using the same formulation, as given by:

$$\hat{y}(k+2|k) = f_1(\hat{y}(k+1|k), y(k), \dots, y(k-N+2)) \tag{4}$$

Using this representation, the value  $\hat{y}(k+1|k)$  is one input of the  $f_1$  mapping, which is used for the prediction of value  $\hat{y}(k+2|k)$ . Since the value  $\hat{y}(k+1|k)$  is an estimation of the actual output  $y(k+1)$ , prediction errors are propagated and long-term predictions cannot be accurately performed by means of autoregressive models.

Using a multi-model strategy, one independent sub-model is employed for each sampling instant within the prediction horizon. Consequently, future predictions do not depend on previous predictions, allowing to obtain more accurate estimations. According to multi-model approach [22], Eq. (4) can be reformulated using Eqs. (3) and (1), originating the recursive Eq. (5), equivalent to the non-recursive Eq. (6)

$$\hat{y}(k+2|k) = f_1(f_1(y(k), \dots, y(k-N+1)), y(k), \dots, y(k-N+2)) \tag{5}$$

$$\hat{y}(k+2|k) = f_2(y(k), y(k-1), \dots, y(k-N+1)) \tag{6}$$

The mapping  $f_2$  is a new mapping such that  $f_2: \mathfrak{R}^N \rightarrow \mathfrak{R}$ . In general, a specific future time instant  $P$  can be expressed by Eq. (7), being  $f_P$  a mapping such that  $f_P: \mathfrak{R}^N \rightarrow \mathfrak{R}$

$$\hat{y}(k+P|k) = f_P(y(k), y(k-1), \dots, y(k-N+1)) \tag{7}$$

Thanks to this structure, predictions over a forecast horizon do not depend on previous predictions, but only on information available at current instant  $k$ ,  $\varphi(k)$ . On the other hand, one independent model ( $f_i$ ) has to be used for each sampling instant within the prediction horizon. As a result, if a future instant  $P$  has to be predicted,  $P$  distinct regression models have to be derived. Fig. 3 illustrates this prediction process using multi-models.

### 3.2. Neural-network regression models

Artificial neural networks (ANN) have been widely used in several areas due to their powerful capacity to capture nonlinear mappings, high accuracy for learning and good robustness properties [23]. They also have the capacity to learn the behavior of poorly understood phenomena and systems where the dependency between inputs and outputs are too complex to be mathematically

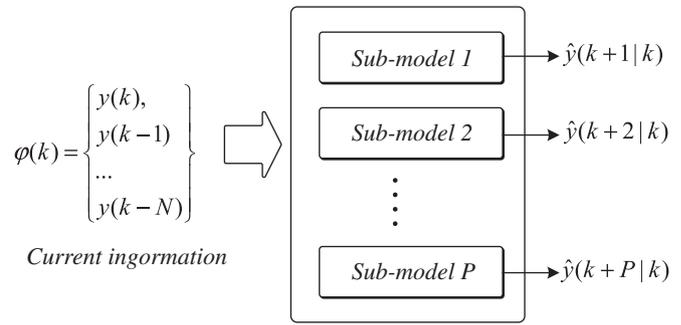


Fig. 3. The multi-model scheme for prediction future P instants.

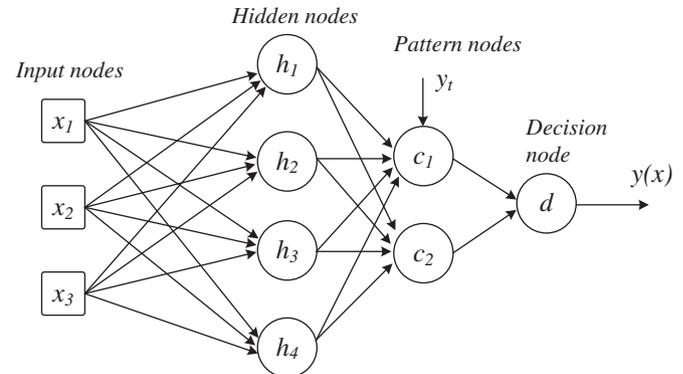


Fig. 4. Example of a GRNN architecture.

described. Additionally, the prediction results of a well-trained neural network are usually accurate.

Therefore, each regression sub-model ( $f_i$ ) is here implemented by means of a neural network model. In particular, generalized regression neural networks (GRNN), a type of radial basis function networks, are considered. In effect, GRNN models can be seen as a normalized radial basis function (RBF) network, in which there is a hidden unit centered at every training case. These RBF units are called “kernels” and are usually probability density functions, such as Gaussian functions. The weights from hidden to output layer are just the target values, so the output is simply a weighted average of the target values of training cases close to the given input case. As a consequence, the only parameters to be learned are the widths of the RBF units [24].

Fig. 4 depicts an example of a particular GRNN, consisting of three inputs  $x_i$ ,  $i=1, 2, 3$  and one output,  $y$ . Furthermore, four hidden neurons,  $h_i$ ,  $i=1, \dots, 4$ , are considered, as a result of four training cases  $\{x_{ti}, y_t\}$   $t=1, \dots, 4$   $i=1, \dots, 3$ .

Receiving the vector of input values ( $x = [x_1 \ x_2 \ x_3]^T$ ) from the input layer, the hidden layer  $h_i(\cdot)$  computes the Euclidean distance between the input vector and the neuron’s center of the kernel (predetermined by the training inputs  $x_t$ ), and then computes the RBF kernel function,  $\varphi_t(x - x_t)$ . The resulting values are passed to the numerator and denominator neurons in the pattern layer, respectively,  $c_1(\cdot)$  and  $c_2(\cdot)$ . This second unit (denominator) adds the values coming from each of the hidden neurons  $\varphi_t(\cdot)$ , while the first unit (numerator) adds those values multiplied by the actual target value ( $y_t$ ) for each hidden neuron. Finally, the decision layer,  $d$ , divides the value in the numerator unit by the value in the denominator unit to derive the predicted value,  $y(x)$ . Mathematically, a GRNN can be described by (8):

$$y(x) = \frac{c_1(\cdot)}{c_2(\cdot)} = \frac{\sum_{t=1}^{NP} y_t \varphi_t(x - x_t)}{\sum_{t=1}^{NP} \varphi_t(x - x_t)} \tag{8}$$

The vector  $x$  represents the current input,  $\varphi_t(\cdot)$  identifies the particular kernel radial basis function,  $y_t$  is the target output correspondent to the input  $x_t$  and  $NP$  is the total number of training pairs  $\{x_t, \varphi y_t\} t=1, \dots, NP$ . In case a Gaussian radial basis function is used, follows Eq. (9), where the parameter  $\sigma$  represents the kernels width

$$\varphi_t(\cdot) = \exp^{-(x-x_t)^T(x-x_t)/2\sigma^2} \tag{9}$$

The principal advantages of GRNN are their aptness for smooth function-approximation, their ability to predict behavior of systems based on few training samples and their interpolation properties between training samples [24]. They enable a fast learning and are often more accurate than multilayer perceptron networks. Moreover, they are especially advantageous due to its ability to converge to the underlying function of the data with only few training samples available. These properties make GRNN a very useful tool to perform predictions [25]. On the other hand, like kernel methods, they suffer from the curse of dimensionality dilemma, requiring more memory space to store the trained model.

As mentioned, although multi-models are used for long-range prediction, each neural network can be trained by means of a *standard backpropagation* algorithm. Actually, as referred, the dimension of the training dataset  $\{x_t, y_t\} t=1, \dots, NP$ , predetermines the number of hidden neurons ( $NP$ ). Thus, the training of a GRNN only involves the estimation of the kernels width,  $\sigma$ , Eq. (9). In the application of the GRNN structure to the particular problem of AHE prediction, the number of previous instants considered by each model (designated here as the *order*) together with the time period before the starting of the forecast window (designated here as *size*) determine the dimension of the dataset, therefore the number of the hidden layers. These *size* and *order* parameters will be detailed in section 4.1.1. The width of each hidden neuron is automatically determined using Matlab, based on an algorithm proposed by [26].

### 3.3. Application to the prediction of AHE

As illustrated in Fig. 2, the prediction of an AHE is based on a set of multi-models trained using mean arterial blood pressure

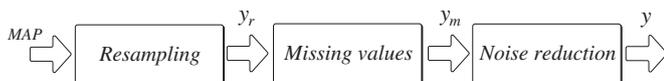


Fig. 5. Pre-processing stages.

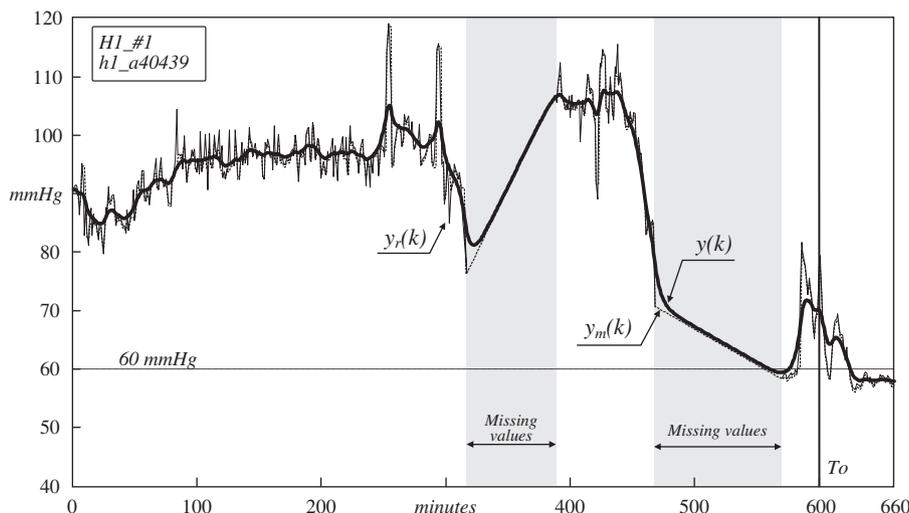


Fig. 6. Pre-processing phases: yr—resampling of MAP signal (1 sample/min); Ym - replacing of missing values; y—filtered signal.

signals (MAP), obtained from MIMIC-II *numerics record* dataset (H and C datasets). No information from *clinical records* was used.

The input consists of a discrete MAP signal (sampled once per minute) considering the information available before  $T_0$ . This signal passes through a set of pre-processing techniques, namely to deal with missing information, noise reduction and normalization. Then, a correlation analysis procedure is carried out using the processed MAP signal and a set of MAP templates, representative of historical MAP evolution trends. From this correlation analysis the most similar templates are identified and the correspondent multi-models, previously trained, are selected. These specific neural multi-models are then employed to predict the future evolution of the particular MAP input signal, from instant  $T_0$  until the end of the forecast window (one hour). Finally, an AHE is identified if, according to the challenge definition, at least 90% of the MAP predicted signal is at or below 60 mmHg during a period of 30 min or more.

#### 3.3.1. Pre-processing

Firstly, a pre-processing stage is applied to the original MAP signal. This processing involves resampling the raw signals, deal with missing values and perform a noise reduction, as illustrated in Fig. 5.

In the first phase all MAP signals are resampled to 1 sample per minute. Thus, for signals presenting a sampling rate of 1 Hz (A and B testing datasets) the average of the blood pressure measured in each 60 samples is considered.

To deal with missing values a simple procedure is carried out. In case the lack of MAP a first order linear interpolation is performed. This process uses the values in the limits of the missing interval, e.g. the last available value (on the left) and the first available value (on the right).

For noise reduction a simple first order filter is used, considering a pole at 0.8.

$$y(k) = \frac{0.2 q^{-1}}{1-0.8 q^{-1}} y_m(k) \tag{10}$$

Finally, the same duration is assumed, equal to  $T=11$  h, for all signals. Specifically, periods of 10 h and 1 h are considered, respectively, before and after the instant  $T_0$ .

Fig. 6 illustrates the pre-processing procedure for the particular signal H1\_#1 (training set group H1 signal #1,  $h1\_a40439$ ). As can be seen, missing values, approximately between instants 320 and 380 min and between 480 and 580 min, are replaced using the described technique.

### 3.3.2. MAP templates and GRNN multi-models

To define the MAP templates a representative historical dataset composed of past and future tendencies has to be considered. In the proposed strategy, the template dataset consists of 59 training records (H and C) is available in Physionet/CinC challenge [27]. Actually, one signal was excluded (C2\_#4, c2\_40234), since it presented a significant discontinuity in the neighborhood of the instant  $T_0$ , which caused significant difficulties during the training phase. Therefore, a matrix of  $D$  templates,  $Y \in \mathfrak{R}^{D \times T}$ , with  $D=59$  and  $T=660$  min is defined.

$$Y = \begin{bmatrix} Y^1 \\ Y^2 \\ \dots \\ Y^D \end{bmatrix} \quad \text{with } Y^i \in \mathfrak{R}^T, i = 1 \dots D \quad (11)$$

The pre-processing phase, described in the last section, is applied to each  $Y^i$  template.

To address future predictions, each of these time series templates (H and C) is modeled using the GRNN multi-model approach. These models are trained using past and future information (before and after  $T_0$ ). Moreover, to reduce the number of sub-models, each GRNN structure is trained to deal with 15 step ahead predictions. As result, given the forecast period (1 h), 4 neural sub-models are considered for each MAP template ( $f_1, f_2, f_3, f_4$ ). Consequently, the total number of models is  $59 \times 4=236$ .

### 3.3.3. Prediction of MAP signals

Given a new MAP testing signal, truncated at time instant  $T_0$ , the MAP forecast is done based on previous trained GRNN multi-models. To select the specific multi-models, a correlation analysis procedure takes place. Basically, correlation coefficients  $CC_i$  between new MAP signal and stored MAP templates are, in a first stage, calculated.

$$CC_i = \frac{\sum_{k=To-size}^{To} (X_k - \bar{X})(Y^i_k - \bar{Y}^i_k)}{\sqrt{\sum_{k=To-size}^{To} (X_k - \bar{X})^2} \sqrt{\sum_{k=To-size}^{To} (Y^i_k - \bar{Y}^i_k)^2}} \quad i = 1, \dots, D \quad (12)$$

The correlation vector,  $CC \in \mathfrak{R}^D$ , is computed for a specific period of time ( $size$ ), starting before the forecast window until instant  $T_0$  ( $size$  parameter will be introduced in the next section). The signal  $X$  represents the new MAP signal, with dimension  $X \in \mathfrak{R}^T$ . As referred, the signal  $Y^i$  represents the  $i$ th template ( $i=1, \dots, D$ ) of the matrix  $Y$ . The scalars  $\bar{X}$  and  $\bar{Y}^i$  define the means of vectors  $X$  and  $Y^i$ , respectively.

In a second stage, the MAP templates that present a correlation coefficient verifying a given threshold value are selected. In particular, being  $CC_L \in \mathfrak{R}^L$  ( $L \leq D$ ) the vector composed of all positive correlation coefficients (sorted in descending order), the first  $M$  templates are selected, such that Eq. (13) is verified.

$$\frac{\sum_{i=1}^M CC_L(i)}{\sum_{i=1}^L CC_L(i)} > tolerance \quad (13)$$

The parameter  $tolerance$  is a pre-defined scalar, aiming to only select the most relevant multi-models (with the highest correlation). Therefore, using this approach a variable number of multi-models is selected for each particular MAP signal.

Finally, the occurrence of an AHE is assessed according to the AHE definition [27], considering the predicted MAP signal  $y_p(k)$  over the forecast window. For the forecast of the MAP signal, a weighted average of the predictions performed by the  $M$  multi-models is computed, as described by Eq. (14).

$$y_p(k) = \frac{\sum_{i=1}^M CC_L(i) \times Y^i}{\sum_{i=1}^M CC_L(i)} \quad k = T_0, \dots, T_0+60 \quad (14)$$

## 4. Results

Following, the main topics of validation results are presented. It should be referred that all the implementations done in this work (regarding database access, signal processing, classification, training and validation) were carried out using Matlab [28].

### 4.1. GRNN multi-models

#### 4.1.1. GRNN size and order

When modeling each MAP signal template, the selection of the order ( $N$ ) and the size ( $S$ ) is of particular importance. The parameter size is defined as the period before the starting of the forecast window, from where information is used for training purposes. The parameter order defines the previous instants considered by each model, Eq. (1). In order to estimate these parameters an optimization procedure was carried out, by means of the minimization of the least square prediction error over the forecast window

$$\min_{size, order} \sum_{k=To}^{k=To+60} (y(k) - y_p(k))^2 \quad (15)$$

Variables  $y(k)$  and  $y_p(k)$  define, respectively, the actual and the predicted MAP signal. The referred minimization procedure was performed considering different values for the order and size parameters, specifically in the ranges  $order \in [60 \dots 90]$  and  $size \in [120 \dots 180]$ , with increments of 10 min.

Fig. 7 depicts the histogram regarding the order and size parameters obtained for all test datasets (H and C datasets, 60 signals). As can be seen, the predominance of the models can be described by  $(order, size)=(60, 170)$  and  $(order, size)=(80, 140)$ . The statistic analysis of mean and standard deviation results in the following values:  $mean(size)=152$  min,  $std(size)=15$  min;  $mean(order)=70$  min,  $std(order)=9$  min. Based on these results, it can be concluded that the evolution of the MAP signals can be characterized, on average, using the past 70 minutes (order), and the models should be trained using the past 2.5 h (size).

#### 4.1.2. GRNN training

The GRNN structures were defined and trained using the *newgrnn* function [25], available in Matlab toolbox. Basically, the training of a GRNN is performed in a single step (no back-propagation of error is involved). In particular, the training

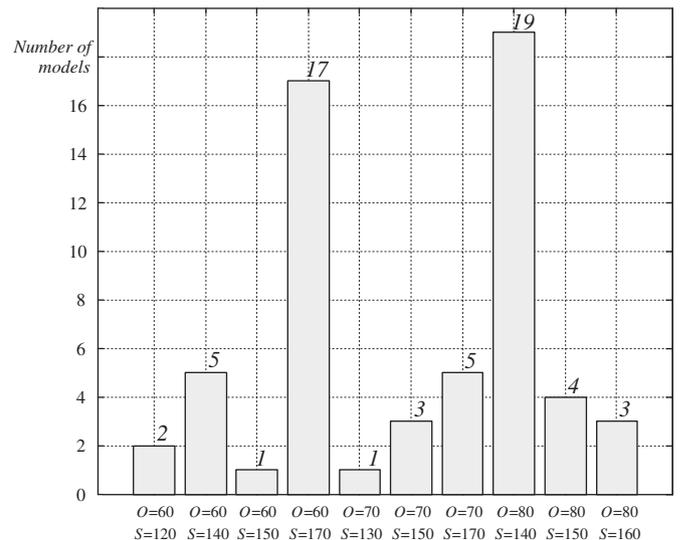


Fig. 7. Size and order parameters, estimated for training signals (H and C dataset).

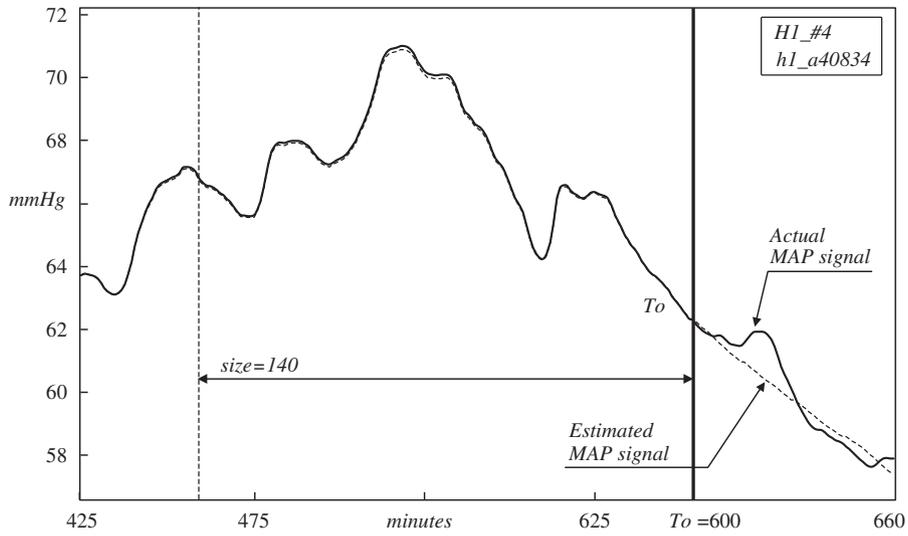


Fig. 8. GRNN training—testing signal H1\_#4 (h1\_a40834).

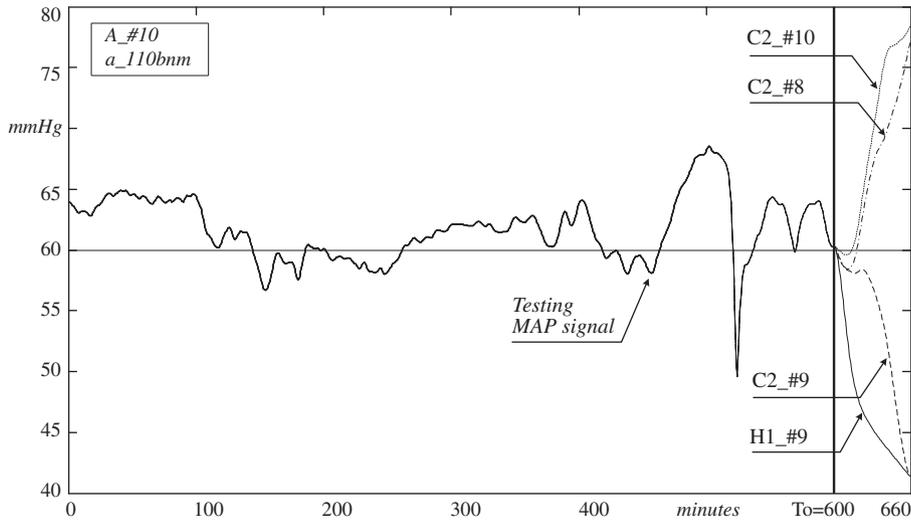


Fig. 9. Prediction of MAP signal A\_#10 (a\_110 bnm) using neural multi-models.

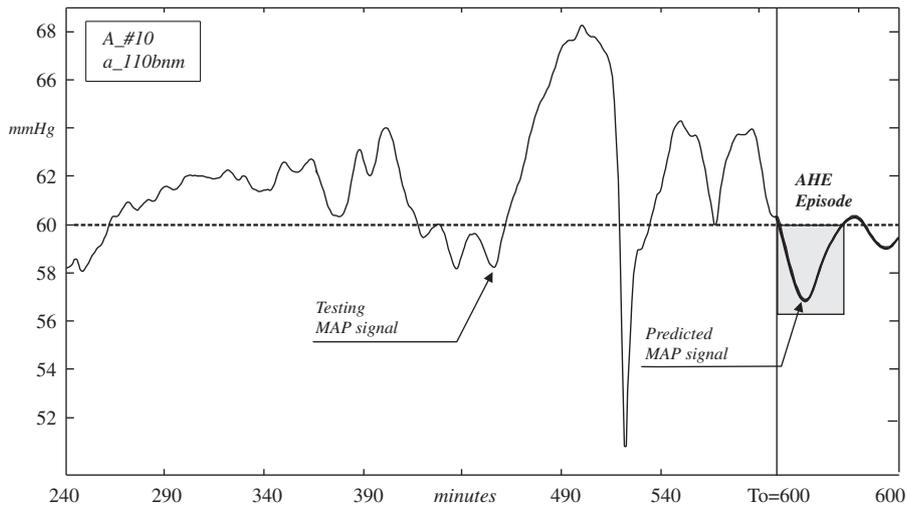


Fig. 10. MAP prediction and AHE identification; testing signal A\_#10 (a\_110 bnm). (a) MAP signal B\_#24, FN event. (b) MAP signal B\_#5, FP event. (c) MAP signal B\_#26, FP event.

comprises the determination of kernels widths. Fig. 8 presents the training results for the particular testing record H1\_#4 (h1\_a40834). For this specific signal, the size and order values are, respectively, 140 and 80 min.

It is important to stress that the neural network multi-models predict future behavior of signals over the whole prediction horizon, only using information available before the starting of the forecast window (instant  $T_0$ ).

4.2. AHE prediction in the challenge context

Using the present strategy, test datasets available in the Physionet/CinC challenge (10 records of A dataset, and 40 records of B dataset) were used for validation purposes. Firstly, each of these 50 signals was correlated with the MAP templates, considering a specific period of size minutes before instant  $T_0$  (which depends on the specific signal). The correspondent GRNN models, determined from the correlation analysis procedure, were used to predict future MAP values.

Fig. 9 shows the prediction over the forecast window for the particular signal A\_#10 (a\_110bnm). For the determination of models to be employed, a tolerance of 0.15 was considered in Eq. (13). From this computation, 4 templates ( $M=4$ ) were obtained. These templates correspond to the training records C2\_#9 (c2\_a40329), H1\_#9 (h1\_a41835), C2\_#8 (c2\_a40306) and C2\_#10 (c2\_a40355), with correlation coefficients of 0.5370, 0.5126, 0.5031 and 0.4997, respectively.

The final MAP predicted signal is computed as the weighted average of all four estimated predictions, Eq. (14), being the identification of an episode straightforward performed using the definition of AHE. In this work, an AHE is considered to occur if in a period of 20 min or more, at least 90% of the MAP measurements are at or below 60 mmHg (instead of the 30 min originally defined by the challenge). This reduction is mainly due to the processing phase, since it introduces some delay in the signal evolution as well as it produces a smoothness of the original signal. Fig. 10 shows the prediction of the specific MAP signal A\_#10 (a\_110bnm) over the forecast horizon, resulting from the weighted average of the mentioned templates. For this particular case, using the previous definition an AHE is identified.

Table 1 presents the global results for the 2009 PhysioNet/Computers in Cardiology challenge. As can be observed, for the event 1 (using dataset A) 5 AHE were identified, corresponding to the signals {1, 2, 4, 9, 10}. For the event 2 (using dataset B) 15 episodes were identified.

The proposed methodology achieved a correct prediction of 10 out of 10 AHE for event 1 and of 37 out of 40 AHE for event 2, which were the best results of all entries in the two events of the challenge.

4.2.1. Discussion of results

The obtained results can be described in terms of sensitivity, specificity and accuracy. Sensitivity ( $SE$ ), Eq. (16), gives the percentage of actual AHE that were correctly identified; specificity ( $SP$ ), Eq. (17), gives the percentage of AHE that did not occur and were correctly identified, and accuracy ( $AC$ ), Eq. (18), gives

Table 1  
AHE detection.

	AHE detected
Dataset A	1, 2, 4, 9, 10
Dataset B	2, 3, 5, 7, 9, 14, 17, 18, 22, 23, 25, 26, 34, 38, 39

the total percentage of correct predictions.

$$SE = \frac{TP}{TP + FN} \tag{16}$$

$$SP = \frac{TN}{TN + FP} \tag{17}$$

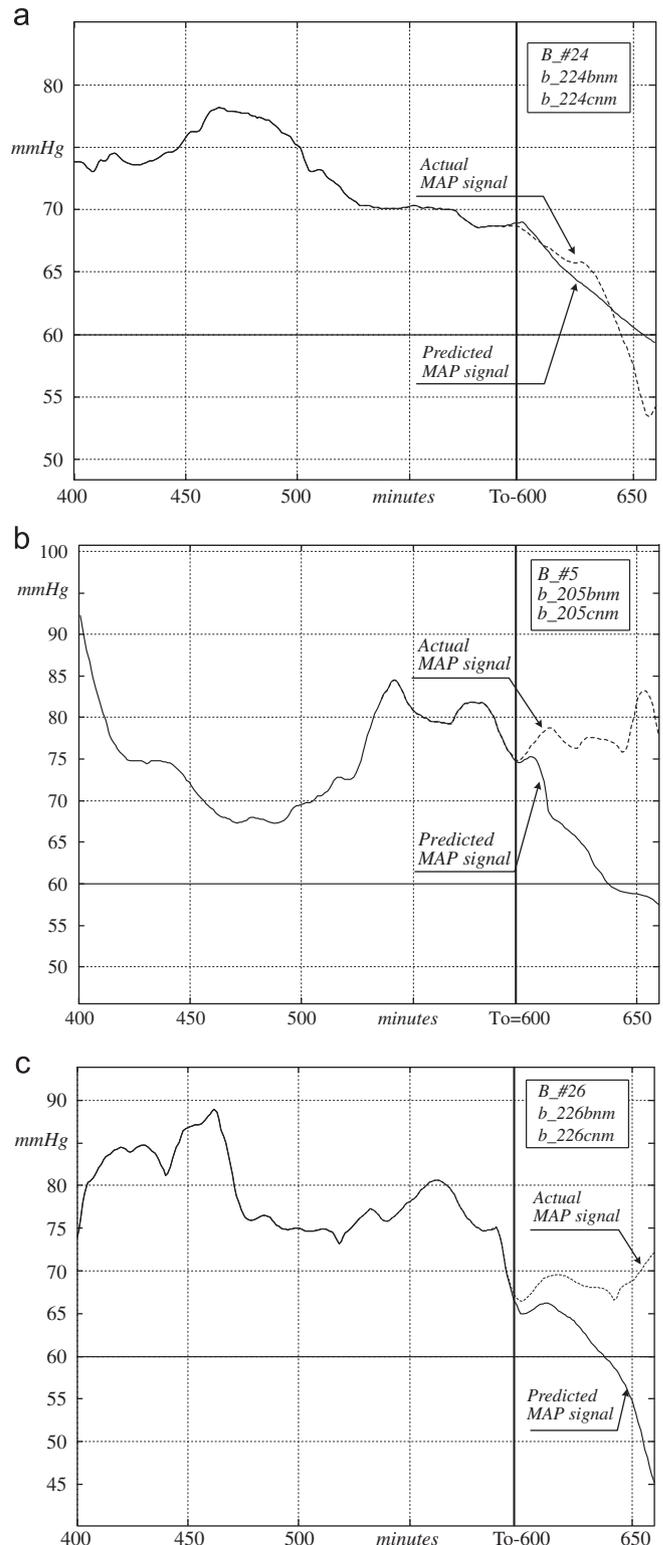


Fig. 11. MAP signals incorrectly classified: predicted and actual values.

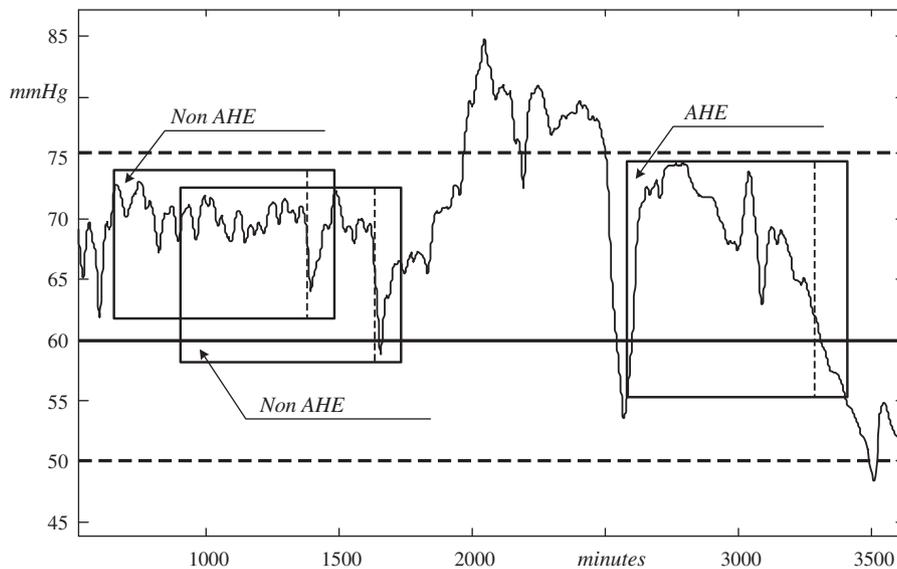


Fig. 12. An example of a test case selection from the MIMIC-II database.

$$AC = \frac{TP+TN}{TP+TN+FP+FN} \quad (18)$$

The variables TP, TN, FP and FN define, respectively, true positive, true negative, false positive and false negative events detected.

In the validation phase, three incorrect predictions were made: one FN event {B\_#26} and two FP events, {B\_#5, B\_#24}. As result, the global sensitivity, specificity and accuracy values were, respectively, SE=94.74%, SP=93.55% and AC=94.00%.

Fig. 11 shows the predicted results as well as the actual MAP signals for the AHE incorrectly classified (the last were made available only after the conclusion of the challenge).

Concerning the FN episode (record B\_#24), as can be seen in Fig. 11(a), actual MAP signal presents a sudden drop approximately at instant  $T_0+35$  minutes. Although the multi-model strategy was able to capture the evolution trend of the signal it was not fast enough to identify the AHE event.

Regarding the FP episodes (records B\_#5 and B\_#26), there is not an obvious justification for the verified situations. Since the proposed strategy only uses MAP history and no clinical information (observations, medication), possibly for these particular cases other sources of information (such as clinical record data) should be considered, in order to achieve a correct prediction. On the other hand, given that the prediction scheme is based on a set of representative templates, these particular signal evolutions can be interpreted as uncommon behaviors that are not characterized by these templates.

In conclusion, although the obtained results are relevant, the experiments performed have suggested that additional clinical information, such as medication, should be considered in a future implementation. Additionally, one of the drawbacks of the proposed strategy relies on the number of models involved (236 models). A methodology able to reduce this number should be also considered, in order to improve the robustness of the approach.

#### 4.3. AHE prediction using an extended dataset

Finally, the prediction strategy was tested considering a larger number of cases, obtained from MIMIC II waveform database (mimic2db) that contains 3915 records of adult ICU patients. From these, a representative sub-set of 311 records of blood

pressure signals was selected. To increase the probability to obtain an AHE episode, segments where blood pressure values were higher or lower than a pre-specified threshold (respectively 75 and 50 mmHg) were discarded. As a result, a total of 2344 segments were obtained, each one with duration of 11 h. From these, 285 contained an AHE episode.

Fig. 12 illustrates the selection procedure of the cases to be tested.

#### 4.3.1. Discussion of the results

Applying the proposed prediction methodology to the 2344 cases previously referred, a global sensitivity of SE=82.8% and a global specificity of SP=78.4% were obtained. Although these results are not so relevant as the ones achieved in the challenge, they still are satisfactory. Moreover, it is important to note that the neural network models have not been re-trained. Actually, they were the same that were trained using the original dataset consisting of 60 records, which can justify the deterioration of the AHE detection performance. Possibly, the use of a larger training dataset could contribute to improve these results.

## 5. Conclusions

In this work, generalized regression neural network models, integrated into a multi-model structure, were proposed to address time-series prediction over a forecast horizon. Although this is a generic technique, with potential to be employed in different areas, it was validated in the prediction of acute hypotensive episodes. The methodology consists of two steps: in the first, a correlation analysis procedure is carried out between the current signal and a representative set of historical evolution trends signals. The most similar ones (templates) are identified and the correspondent prediction neural models, previously trained using those historical signals, selected. In the second step this multi-models structure is employed to the current signal to predict its future evolution.

Applied to the mean arterial blood pressure (MAP) time-series, considered in the 2009 PhysioNet/Computers in Cardiology challenge, the referred strategy allowed to adequately capture MAP evolution and, consequently, to detect the occurrence of hypotensive episodes. In this context, a correct prediction of 10 out of 10 AHE for event 1 and of 37 out of 40 AHE for event 2 was

achieved (SE=94.74% and SP=93.55%), enabling to obtain the best results of all entries in the two events of the challenge. One the other hand, applied to an extended dataset of blood pressure signals also collected from MIMIC-II, the strategy attained values of SE=82.8% and SP=78.4%.

Despite the success achieved by the proposed methodology, future work should consider other sources of information, as well as the reduction of the number of templates and, consequently, the number of multi-models involved.

### Conflicts of interest

None declared.

### Acknowledgments

This work was supported by HeartCycle, a project partly funded by the European Community's Seventh Framework Program, FP7-216695, and by CISUC, Center for Informatics and Systems of University of Coimbra, Portugal.

### References

- [1] J. Piccini, K. Nilsson, The Osler Medical Handbook, in: R. Scott Stephens, S. Haldar, C. Wiener (Eds.), Hypotension and Shock, 2nd ed., Elsevier, Saunders, 2006 (Chapter 20).
- [2] D. Singla, et al., Risk factors for development of early hypotension during spinal anesthesia, *J. Anaesth. Clin. Pharmacol.* 22 (4) (2006) 387–393 387.
- [3] C.-S. Lin, et al., Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks, *Comput. Methods Programs Biomed.* 92 (2008) 193–197.
- [4] J. Bassale, Hypotension Prediction—Arterial Blood Pressure Variability, Technical Report, 2001.
- [5] C. Crespo, et al., Precursors in the arterial blood pressure signal to episodes of acute hypotension in sepsis, in: Proceedings of the 16th International EURASIP Conference BIOSIGNAL, vol. 16, 2002, pp. 206–208.
- [6] L. Lehman, M. Saeed, G. Moody, R. Mark, Similarity-based searching in multiparameter time series databases, *Computers in Cardiology* 35 (2008) 653–656.
- [7] M. Saeed, Temporal Pattern Recognition in Multiparameter ICU Data.; Doctoral dissertation, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 2007.
- [8] Saeed, M., Mark, R.; A Novel Method for the efficient retrieval of similar multiparameter physiologic time series using wavelet-based symbolic representations, in: Proceedings AMIA Annual Symposium Proceedings 2006, 2006, pp. 679–683.
- [9] A. Frolich, D. Caton, Baseline heart rate may predict hypotension after spinal anesthesia in prehydrated obstetrical patients, *Can. J. Anesth.* 49 (2002) 185–189.
- [10] G. Pelosi, et al., Impaired sympathetic response before intradialytic hypotension: a study based on spectral analysis of heart rate and pressure variability, *Clin. Sci.* 96 (1999) 23–31.
- [11] D. Reich, S. Hossain, M. Krol, et al., Predictors of hypotension after induction of general anesthesia, *Anesth. Analg.* 101 (2005) 622–628.
- [12] D. Chamchad, V. Arkoosh, J. Horrow, et al., Using heart rate variability to stratify risk of obstetric patients undergoing spinal anesthesia, *Anesth. Analg.* 99 (2004) 1818–1821.
- [13] R. Hanss, et al., Heart rate variability predicts severe hypotension after spinal anesthesia, *Anesthesiology* 104 (2006) 537–545.
- [14] R. Hanss, B. Bein, T. Ledowski, et al., Heart rate variability predicts severe hypotension after spinal anesthesia for elective cesarean delivery, *Anesthesiology* 102 (2005) 1086–1093.
- [15] R. Hanss, B. Bein, H. Francksen, et al., Heart rate variability-guided prophylactic treatment of severe hypotension after subarachnoid block for elective cesarean delivery, *Anesthesiology* 104 (2006) 635–643.
- [16] L. Mancini, et al., Short term variability of oxygen saturation during hemodialysis is a warning parameter for hypotension appearance, *Comput. Cardiol.* 35 (2008) 881–883.
- [17] J. Lee, R. Mark, An investigation of patterns in hemodynamic data indicative of impending hypotension in intensive care, *BioMedical Engineering Online* 9.1, 62, 2010.
- [18] <<http://www.physionet.org/challenge/2009/training-set.shtml>> (training data). <<http://www.physionet.org/pn3/challenge/2009/>> (test data).
- [19] PhysioNet/Computers in Cardiology Challenge 2009, Predicting Acute Hypotensive Episodes Home Page. <<http://www.physionet.org/challenge/2009/>>.
- [20] M. Saeed, C. Lieu, G. Raber, R.G. Mark, I.I. MIMIC, A massive temporal ICU patient database to support research in intelligent patient monitoring, *Comput. Cardiol.* 29 (2002) 641–644.
- [21] <<http://mimic.physionet.org/>>.
- [22] Lawrynczuk, M., Computationally efficient nonlinear predictive control based on RBF neural multi-models, In: Proceedings of the 18th International Conference on Artificial Neural Networks, Prague, Czech Republic, 3–6 September, 2008.
- [23] S. Haykin, *Neural Networks and Learning Machines*, third ed., Prentice Hall, 2008.
- [24] Bauer, M.r1995.
- [25] P. Sherrrod, DTREG Predictive Modeling Software, Copyright© 2003–2010. <[www.dtreg.com](http://www.dtreg.com)>.
- [26] P. Wasserman, *Advanced Methods in Neural Computing*, Van Nostrand Reinhold, New York, 1993, pp. 155–161.
- [27] G.B. Moody, L.H. Lehman, Predicting acute hypotensive episodes: the 10th annual PhysioNet/computers in cardiology challenge, *Comput. Cardiol.* (2009).
- [28] MatLab; Mathworks, Inc., 2007.