

Support Vector Machine with External Recurrences for Modeling Dynamic Cerebral Autoregulation*

Max Chacón¹, Darwin Diaz¹, Luis Ríos¹, David Evans², and Ronney Panerai²

¹ Universidad de Santiago de Chile; Departamento de Ingeniería Informática,
Av. Ecuador No 3659 - Casilla 10233; Santiago-Chile
mchacon@diinf.usach.cl, darwindiaz@gmail.com,
lrios@diinf.usach.cl

² Medical Physics Group, Department of Cardiovascular Sciences, University of Leicester,
Leicester Royal Infirmary, Leicester LE1 5WW, UK
dhe@le.ac.uk, rp9@le.ac.uk

Abstract. Support Vector Machines (SVM) have been applied extensively to classification and regression problems, but there are few solutions proposed for problems involving time-series. To evaluate their potential, a problem of difficult solution in the field of biological signal modeling has been chosen, namely the characterization of the cerebral blood flow autoregulation system, by means of dynamic models of the pressure-flow relationship. The results show a superiority of the SVMs, with 5% better correlation than the neural network models and 18% better than linear systems. In addition, SVMs produce an index for measuring the quality of the autoregulation system which is more stable than indices obtained with other methods. This has a clear clinical advantage.

Keywords: Support Vector Machine, biological signals, cerebral autoregulation.

1 Introduction

Support Vector Machines (SVMs) have shown their usefulness by improving over the performance of different supervised learning methods, either as classification models or as regression models. But the small number of papers involving the prediction of temporal series or signal modeling [1-2] shows a lack of assessment in this respect. When these applications are restricted to the field of biomedical signals, SVMs are used as classical classifiers following a process of extraction of signal characteristics [3-4]. At this time we are not aware of any applications that use SVMs as recurrent structures for modeling biomedical signals.

To evaluate SVMs in the field of biomedical signal modeling and prediction we have chosen to model the Autoregulation Blood Flow System (ABFS). Describing this system adequately is a complex problem for which currently there is no model that represents the phenomenon properly. This means that we can not have reliable methods that allow detecting, diagnosing and monitoring different cerebrovascular

* This works was supported by FONDECYT, Chile, under project 1050082.

conditions such as stroke, carotid artery disease, severe head injury, Alzheimer's disease, vascular dementia, arterial hypertension and others.

The ABFS performs the function of maintaining Cerebral Blood Flow (CBF) approximately constant within a given range, even if there are changes in Arterial Blood Pressure (ABP). The study of this system has been possible due to the development of Transcranial Doppler as a tool for measuring Cerebral Blood Flow Velocity (CBFV), which can be assumed to be equivalent to CBF from the studies of Newell et al [5].

The classical way of representing this system is to attempt the dynamic characterization of the ABP-CBFV relationship. Different techniques have been used to provoke changes in ABP and then observe the corresponding CBFV. Chiefly amongst them is the sudden release of thigh cuffs inflated above systolic pressure for 3 min. [5-9]. Different system identification methods have been used to evaluate the dynamic ABP-CBFV relationship. Those carried out in the time domain should be noted; they model the ABP-CBFV relation by means of a differential equation proposed by Aaslid-Tiecks (A-T) [9] which has the advantage of measuring the autorregulation of each subject with a unidimensional index. These methods assume that there is a linear relation between ABP and CBFV, but it is well known that there are a number of nonlinearities in the system [6,10-12].

As far as we know, the only papers that modelled the thigh cuff maneuver responses in a nonlinear way are those of Panerai [10-11]. The first of them uses Winer-Laguerre filters (with second order polynomials) and the second uses artificial neural networks (Time Lagged FeedForward, with Laguerre memory), but neither of them improved the predictive capacity of linear models such as those of A-T [9].

To carry out this work we have used 84 thigh cuff maneuvers performed to 15 healthy subjects. We propose to use the SVMs as static regression machines, adding to them external recurrences, allowing the temporal ABP-CBFV relation to be treated. To compare the results we used the same data to build models with neural networks and we also applied the equations of Aaslid-Tiecks [9].

2 Methods

2.1 Data Collection and Pre-processing

The study considers 16 voluntary subjects who did not have a history of cardiovascular problems, hypertension or nervous system disease. The average age \pm SD of the subjects was 30 ± 7 years, with a range of 23 to 47 years. The measurements were made in a room at a temperature of approximately 30 °C at the University of Leicester, in England. The study was approved by the Leicestershire Ethics Committee.

Each of the 16 subjects underwent six thigh cuff maneuvers (a total of 96). Seven maneuvers (in four subjects) had excessive noise at their origin, and three of them were concentrated in one subject. In order to have maneuvers that would allow a symmetric performance of training and tests (2 or 3 maneuvers for training and the same number for testing), the subject who had the fewest maneuvers was discarded

and three subjects remained with only four maneuvers. As a result, 15 subjects underwent a total of 84 maneuvers.

CBFV was monitored in the middle cerebral artery using a Scimed QVL-120 Transcranial Doppler system with a 2-MHz transducer. ABP was measured with a noninvasive Finapres 2300 Ohmeda pressure monitor.

Pressure and flow data were collected and stored on a digital audio tape on an eight channel recorder (Sony PC108M) for later processing. The data from the tape were transferred to a microcomputer in real time. Fast Fourier transform was used to extract the maximum frequency (i.e. velocity) envelope with a 5-ms time window. The ABP signal was sampled at 200 [samples/sec]. Both signals were filtered with an eighth order lowpass Butterworth filter with a 20-Hz frequency cut-off. The beginning of each cardiac cycle was detected from the diastolic value of the ABP wave, and mean values for ABP (MABP) and CBFV were calculated for each cardiac cycle, interpolated with a spline and resampled at 0.2 s interval to produce signals with a uniform sampling rate (5 samples/s).

2.2 Modeling Techniques

Support Vector Machine

The SVM algorithm that we used corresponds to the one called v-SVM, which was introduced by Schölkopf et al. [13] in the late 1990s. This algorithm is based on the results of the statistical theory of learning given by Vapnik [14], which introduces regression as the fitting of a tube of radius ϵ to the data. The decision boundary for determining the radius of the tube is given by a small subset of training examples called Support Vectors (SV).

Vapnik’s SVM Regression estimates the values of \vec{w} to obtain the function

$$f(\vec{x}) = (\vec{w} \cdot \vec{x}) + b, \quad \vec{w}, \vec{x} \in \mathbf{R}^N, b \in \mathbf{R}, \quad (1)$$

by introducing the so called ϵ -insensitive loss function shown in equation 2:

$$|y - f(\vec{x})|_\epsilon = \max\{0, |y - f(\vec{x})| - \epsilon\}, \quad (2)$$

which does not penalize errors smaller than $\epsilon > 0$ (where ϵ corresponds to a value chosen a priori).

The algorithm is implemented by minimizing the functional risk $\|\vec{w}\|^2$ to which is added a penalty for leaving points outside the tube (identified by slack variables ξ). In this way the risk function to be minimized is given by equation 3, where C is a constant that determines the trade-off between the complexity of the model and the points that remain outside the tube. Figure 1 shows a geometric interpretation for the case of a linear regression.

$$\text{minimize } \theta(\vec{w}, \xi) = \frac{1}{2} \|\vec{w}\|^2 + \frac{C}{l} \sum_{i=1}^l \xi_i, \quad (3)$$

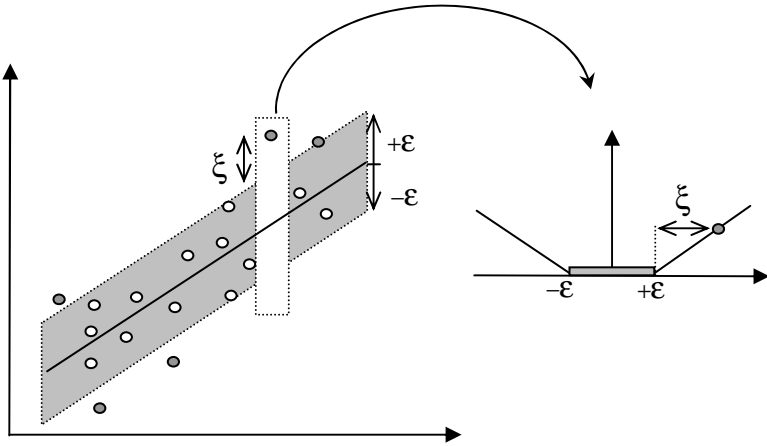


Fig. 1. Geometric interpretation of the SV regression in which the regression equation is estimated by means of a tube of radius ϵ . The trade-off between the complexity of the model and the points left outside the regression tube is controlled by the slack variables ξ . The dark points correspond to the SVs.

The variation of the ν -SVM introduced by Schölkopf et al. [13], consists in adding a variable $\nu \geq 0$ that weights the size of ϵ making a trade-off between the complexity of the model and the slack variables (equation 4).

$$\text{minimize } \theta(\bar{w}, \xi) = \frac{1}{2} \|\bar{w}\|^2 + C \left(\nu \epsilon + \frac{1}{l} \sum_{i=1}^l \xi_i \right) \tag{4}$$

The solution of this minimization problem for obtaining the weight vectors \bar{w} is found by the standard optimization procedure for a problem with inequality restrictions when applying the conditions of Kuhn-Tucker to the dual problem. The main advantage of introducing parameter $\nu \in [0-1]$ is that it makes it possible to control the error fraction and the number (or fraction) of SVs with only one normalized parameter.

To solve a nonlinear regression problem it is sufficient to substitute the dot product between two independent original variables $\bar{x}_i \cdot \bar{x}_j$ by a kernel function $k(\Phi(\bar{x}_i) \cdot \Phi(\bar{x}_j))$. This function carries out the dot product in a space of higher dimension such that it ensures the linearity of the regression function in the new space via a nonlinear map Φ . Several functions may be used as kernel, such as the Gaussian $k(\bar{x}_i, \bar{x}_j) = \exp(-\|\bar{x}_i - \bar{x}_j\|^2 / (2\sigma^2))$ or the polynomial function $k(\bar{x}_i, \bar{x}_j) = (\bar{x}_i \cdot \bar{x}_j)^p$. In this way the nonlinear regression equation is given by equation 5.

$$f(\bar{x}) = k(\bar{w}, \bar{x}) + b \tag{5}$$

To implement the recurrence in the SVM we used external feedback from the delayed outputs ($v(t-i)$ =CBFV), and the present input ($p(t)$ =MABP) and past time instants ($p(t-j)$) are considered, as shown in figure 2.

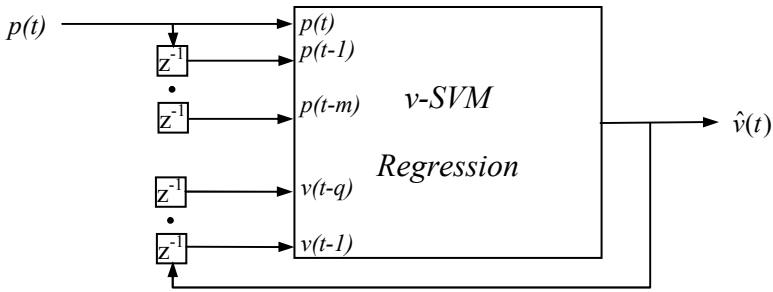


Fig. 2. Regression SVM with external recurrences for modeling the ABP-CBFV relation

Neural Networks

To implement the neural models we used the ideas that are introduced in the work of Nerrand et al. [15], where it is shown that a static neural network with external recurrences is equivalent to the neural networks with internal recurrences. The resultant model is similar to that shown in fig. 2 when the SVM is replaced by a static neural network.

The training algorithms for the static network correspond to variations of the backpropagation method, which is carried out by approximations of second order gradient descent methods.

The two methods used correspond to algorithms of the quasi-Newton type which avoid the direct calculation of the Hessian matrix. The One Step Secant method was used first; it uses an approximation for calculating the search direction of the best descending slope as the combination of the slopes calculated in previous steps. The second method corresponds to the algorithm of Levenbert-Marquardt, which uses a multiplication of Jacobian matrices to estimate the Hessian matrix.

Linear Model

As reference model, we implemented the most widely used linear method in the autoregulation literature, which is that of A-T [9]. This method estimates the CBFV from a second order linear differential equation that uses two state variables, x_1 and x_2 , as shown in equations 6.

$$\begin{aligned}
 NP(t) &= \frac{p(t)}{1 - CCP} ; \quad x_1(t) = x_1(t-1) + \frac{NP(t-1) - x_2(t-1)}{f \times T} \\
 x_2(t) &= x_2(t-1) + \frac{x_1(t-1) - 2 \times D \times x_2(t-1)}{f \times T} ; \quad \hat{v}(t) = 1 + dP(t) - K \times x_2(t)
 \end{aligned}
 \tag{6}$$

where f is the sampling frequency and CCP refers to the Critical Closing Pressure. $NP(t)$ is the change in the mean MABP normalized from a reference value. The model’s parameters are given by K , D and T , where K is a parameter that represents the system’s autoregulatory gain, D is the damping factor, and T is a time constant.

For every maneuver produced in the $p(t)$, a corresponding model velocity $\hat{v}(t)$ is predicted, which can be compared with the actual velocity $v(t)$. To choose the velocity

given by the model, 10 different responses by the differential equation are generated, corresponding to 10 different values of parameters K , T and D . The combination of parameters to generate the $\hat{v}(t)$ outputs has been tabulated previously [9]. The model's response signal is chosen as the curve that shows the largest correlation (or smallest error) with the actual velocity signal $v(t)$.

This method has the advantage that the 10 tabulated responses for the combination of parameters correspond to an index that measures the quality of the autoregulation (ARI index) of the subjects who undergo the thigh cuff techniques. The ARI index classifies the subjects on a 0 to 9 scale in which 0 represents the complete absence of autoregulation and 9 refers to perfect autoregulation.

2.3 Selection of Parameters and Statistical Analysis

The main objective is to compare the different methods in terms of their predictive capacity, with the purpose of obtaining a method that generates models that represent adequately the ABFS. To evaluate the methods' predictions, the data were divided into two disjoint groups of similar size. One group consisted of odd-numbered maneuvers (1 3 5 for subjects with 6 maneuvers) and the other group contained the even-numbered maneuvers (2 4 6 for subjects with 6 maneuvers), in this way trying to avoid any bias due to physiological accommodation between one maneuver and another [11]. Therefore, a "balanced" cross-validation process takes place that consists in training with even-numbered maneuvers and performing the test with odd-numbered maneuvers, and vice versa. The model chosen corresponds to the training that produced the best correlation (or the smallest error) with the test set.

For the nonlinear models we have not made any a priori assumptions with respect to a general structure for autoregulation. Consequently, every parameter of the model must be fitted for each of the subjects; thereby obtaining what we call models "parametrized by subject".

In the case of the SVM one must seek the best combination of MABP and CBFV delays and of parameters v , C and σ for each subject, and similarly for the case of the neural networks, where in addition to fitting the delays, the optimum number of neurons in the hidden layer must be established.

To measure the performance of the models, the correlation coefficient r as well as the mean square error (MSE) were used. The latter is defined according to equation 7.

$$MSE = \frac{\sum_{i=1}^N (v(i) - \hat{v}(i))^2}{N} \quad (7)$$

In order to compare different model results in terms of error and correlation, Wilcoxon's non-parametric sign test was used, and two results were considered significantly different when $p < 0.05$.

In addition to evaluating the predictive capacity of each method, it is important to analyze the variability of the autoregulatory indices generated by each method. To calculate the autoregulatory indices in the nonlinear cases, it is possible to enter a step signal into the trained models and to fit the A-T equations [9] to obtain the

corresponding autoregulatory index (ARI). To analyze the variability of this index, we simply calculated the coefficient of variation ($CV = \sigma / \bar{x}$).

It is also important to find out if the nonlinear models succeed in determining whether the model is able to capture the physiological dynamics of the ABFS, or if the model has only managed a numerical fit for the data. One way of determining whether this is or is not the case is to input an ABP step signal that simulates a thigh cuff technique to an already trained model, and examine the CBFV response signal.

3 Results

The three methods were applied to each of the 15 subjects, generating a parametrized model by subject in the case of the nonlinear models.

To build a model by subject using the SVM, from 1 to 10 delays were tested in the CBFV (which corresponds to the order of the model) and from 1 to 14 delays in the MABP. Parameter ν was evaluated from 0.1 through 1 in intervals of 0.1. Parameter σ was tested from the units to the hundreds in an incremental way. Parameter C was increased by powers of two [16], and the best results were obtained when C reached higher values, deciding to leave it at infinity. This means that the solution is concentrated on minimizing the empirical risk, without considering the complexity of the model.

For the SVM, parameter ν varied from 0.1 to 0.9, with a predominance of 0.9 values. Parameter σ varied between 5 and 60, with a predominance of a value of 15.

In the case of the neural networks, the same delayed ranges used for the SVMs were applied to MABP and CBFV, and for the neurons of the hidden layer a range of 2 to 20 was used. When this parameter combination was applied to the subject population, the following results were obtained: the most frequent numbers of neurons in the hidden layer were 4 and 6; the delays in CBFV varied between 1 and 3, with 1 being the most frequent; the delays in MABP varied between 4 and 10, with 7 as the most frequent. The best results in the neural networks were obtained when the training method of Levenbert-Marquardt was used. Table 1 shows the correlation and MSE results for both the training and the test.

Table 1. Training and test correlations for the different methods that model the ABP-CBFV relation

Method	Training Correlation	Training MSE	Testing Correlation	Testing MSE
SVM	0.97	0.004	0.83	0.028
NN	0.85	0.043	0.79	0.066
A-T	0.77	0.074	0.70	0.094

To evaluate if there is a significant difference between the models, Wilcoxon’s test was applied to the correlations between the SVMs and the neural networks, showing that there is no significant difference, with $p=0.233$. When this test was applied to the SVMs and the linear model of A-T, the differences were significant, ($p=0.0001$).

Finally, the differences between the neural networks and the linear model were also significant ($p=0.0006$).

Figure 3 shows the average signals of the step response for the SVM and the neural network methods.

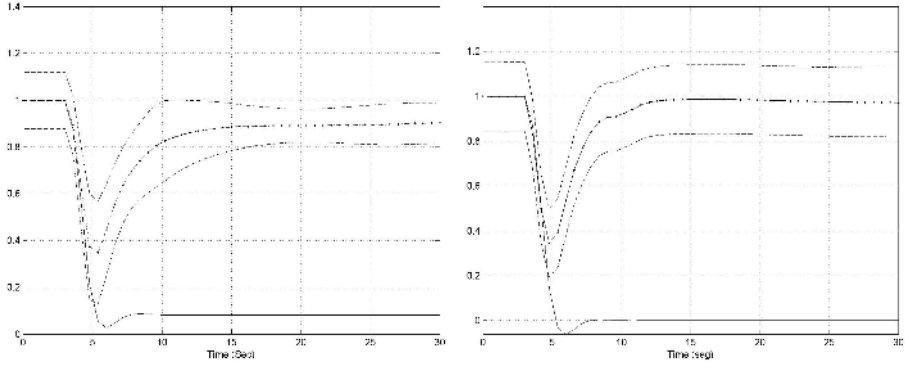


Fig. 3. CBFV response to the step signal for SVM (left) and Neural Networks (right). Solid line: step signal; dotted line: average CBFV response; dashed line: ± 1 standard deviation.

After applying a step signal to each nonlinear model, the method of A-T was used to obtain the ARI index of each subject represented by the model. In the case of the linear model of A-T, the ARI indices are obtained directly.

The mean ARI indices, standard deviations and coefficients of variation for all the subjects are shown in table 2.

Table 2. Mean values, standard deviations and coefficients of variation for the autoregulatory indices of the different methods

Method	Mean ARI	Standard Deviation	CV (%)
SVM	5.38	1.44	26.7
NN	6.30	2.52	40.0
A-T	4.39	1.47	33.5

4 Discussion and Conclusions

The correlation and error values in the test show that the method that best predicts, on the average, the CBFV responses for unknown cases is that of the SVMs, followed closely by neural networks. The same order, but with greater differences, is kept in the training sets.

The largest differences are found between the nonlinear and the linear methods, and they are significantly different when Wilcoxon's test is applied. When the

nonlinear methods are compared, even though there are differences between the average correlation values, they are not significant ($p > 0.05$).

Evaluating the predictive capacity by means of the errors made in the test set is not sufficient to infer that the model represents the ABFS adequately. It is also necessary to know if the model captures the physiological dynamics of the ABFS. When examining the responses of the CBFV to the step signal (figure 3), it can be seen that the responses of the nonlinear models correspond to normal responses, recovering the CBFV level in less than 6 s. It is also seen that the ranges of the standard deviations are different, with the neural networks showing the greatest deviations, which means less precision of the neural networks compared to the SVMs.

Table 2 shows that the highest mean ARI indices correspond to the neural networks, but they also have great variance, in agreement with the results of the step responses. ARI values greater than 6 are those closest to the actual values, because they correspond to normal autoregulation values, in agreement with the set of subjects in the study. But the variance that exists between these indices must also be analyzed, and in this case the smallest corresponds to the SVMs. A way of combining in a single measure the mean ARI values and their variance is by through the coefficient of variation, which in this case shows that SVMs are the best method. Reducing the variability of the ARI index is of great clinical importance, because this variability is the main limitation to use this index in measuring the autoregulation system [17].

The results of this work show that the nonlinear models are significantly better than the linear models for representing the ABFS. Similar approaches had only been tested with other types of ABFS signals, such as spontaneous pressure changes [12-16, 18], but so far, the available work with thigh cuff maneuvers had not succeeded in improving on the linear models [10-11].

The idea of using external recurrences in the SVMs for the treatment of biological signals has proved to be adequate. Although the SVM predictive correlation and error were not significantly different from the corresponding figures for the neural network (Table 1), the trend showed by the slightly higher numerical values of correlation and smaller error, together with the much smaller coefficient of variation (Table 2) for estimating the autoregulatory index of individual subjects, suggest that this approach should take precedence as the best technique for modelling cerebral autoregulation by non-linear methods, as assessed by the thigh cuff technique. Future work should be performed to compare the performance of SVMs with alternative approaches in other conditions, such as spontaneous fluctuations in ABP or changes induced by the Valsalva maneuver or CO₂ reactivity tests.

References

1. Raicharoen T, Lursinsap C, Sanguanbhokai P, Application of critical support vector machine to time series prediction. *Circuits and Systems*, 2003. Proceedings of ISCAS '03. **5** (2003) 741-4.
2. Wan-Zhao C, Chang-Chun Z, Wen-Xing B and Jun-Hua L. Chaotic time series prediction using mean-field theory for support vector machine. *Chinese Phys.* **14** (2005) 922-9.
3. Jankowski S and Oreziak A, Learning system for computer-aided ECG analysis based on support vector machines. *International Journal of Bioelectromagnetism* **5** (2003) 175-6.

4. Acir N and Guzelis C. Automatic spike detection in EEG by a two-stage procedure based on support vector machines. *Comput Biol Med.* **7** (2004) 561-75.
5. Newell D, Aaslid R, Lam A, Mayberg T, Winn R. Comparison of flow and velocity during autoregulation testing in humans. *Stroke* **25** (1994) 793-7.
6. Panerai R. Assessment of cerebral pressure autoregulation in humans - a review of measurement methods. *Physiological Measurement* **19** (1998) 305-38.
7. Panerai R, Evans D, Mahony P, Deverson S, Hayes P. Assessment of thigh cuff technique for measurement of dynamic cerebral autoregulation. *Stroke* **31** (2000)476-80.
8. Panerai RB, Dawson SL, Eames PJ and Potter JF. Cerebral blood flow velocity response to induced and spontaneous sudden changes in arterial blood pressure. *Am J Physiol* **280** (2001) H2162-H2174.
9. Tiecks F, Lam A, Aalid R, Newell D. Comparison of static and dynamic cerebral Autoregulation measurements. *Stroke.* **26** (1995) 1014-19.
10. Panerai R, Dawson S and Potter J. Linear and nonlinear analysis of human dynamic cerebral autoregulation. *Am J Physiol.* **227** (1999) H1089-H1099.
11. Panerai R, Chacón M, Pereira R and Evans D. Neural network modeling of dynamic Cerebral Autoregulation: assessment and comparison with established methods. *Med Eng & Phys* **26** (2004) 43-52.
12. Mitsis G, Zhang G, Levine BD, Marmarelis VZ. Modeling of Nonlinear Physiological Systems with fast and Slow Dynamics. II. Application to cerebral Autoregulation. *Ann. Biomedical Engineering* **30** (2002) 555-65.
13. Schölkopf B, Smola A, Williamson RC, and Bartlett PL. New support vector algorithms. *Neural Computation*, **12** (2000) 1207-45.
14. Vapnik V. *The Nature of Statistical Learning Theory*, Springer Verlag. New York, (1995).
15. Nerrand O, Roussel-Ragot P, Personnaz L. and Dreyfus G. Neural networks and non-linear adaptive filtering: unifying concepts and new algorithms. *Neural Comput.* **5** (1993) 165-99.
16. Frohlich H. and Zell A. Efficient parameter selection for support vector machines in classification and regression via model-based global optimization, *Neural Networks. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference* **3** (2005) 1431-6.
17. Panerai RB, Eames PJ and Potter JF, Variability of time-domain indices of dynamic cerebral Autoregulation. *Physiol. Meas.* **24** (2003) 367-81.
18. Chacón M, Blanco C, Panerai R and Evans D. Nonlinear Modeling of Dynamic Cerebral Autoregulation Using Recurrent Neural Networks. *LNCS, CIARP 2005* **3773** (2005) 205-13.