Fuzzy Evaluation of Heart Rate Signals for Mental Stress Assessment

Mohit Kumar, Matthias Weippert, Reinhard Vilbrandt, Steffi Kreuzfeld, and Regina Stoll

***Abstract*—Mental stress is accompanied by dynamic changes in autonomic nervous system (ANS) activity. Heart rate variability (HRV) analysis is a popular tool for assessing the activities of autonomic nervous system. This paper presents a novel method of HRV analysis for mental stress assessment using fuzzy clustering and robust identification techniques. The approach consists of 1) online monitoring of heart rate signals, 2) signal processing (e.g., using the continuous wavelet transform to extract the local features of HRV in time-frequency domain), 3) exploiting fuzzy clustering and fuzzy identification techniques to render robustness in HRV analysis against uncertainties due to individual variations, and 4) monitoring the functioning of autonomic nervous system under different stress conditions. Our experiments involved 38 physically fit subjects (26 male, 12 female, aged 18–29 years) in air traffic control task simulations. The subjective rating scores of mental workload were assessed using NASA Task Load Index. Fuzzy clustering methods have been used to model the experimental data. Further, a robust fuzzy identification technique has been used to handle the uncertainties due to individual variations for the assessment of mental stress.**

***Index Terms*—Heart rate variability (HRV), continuous wavelet transform, NASA-task load index, fuzzy clustering, robustness, fuzzy identification.**

I. Introduction

ASSESSMENT of mental stress under different workload conditions is a recurrent issue in many engineering and medicine fields. Although mental stress cannot be measured directly, the physiological response of an operator can be interpreted to assess the level of mental stress. Several physiological parameters (like electroencephalograph, blood pressure, heart rate, heart rate variability (HRV), electrodermal activity (EDA), event-related potentials, and electromyograh) have been found sensitive toward any changes occurring in the mental stress level of an operator. Cardiac activity is the most common physiological measure for the assessment of mental workload [1]. An electrocardiogram (ECG) is a cardiac measure that shows sensitivity towards variations in workload [2]. Heart rate variability (a measure of electrocardiographic activity) has been widely accepted in the literature for the assessment of mental workload (see, e.g., [2]–[8]).

Manuscript received October 6, 2005; revised January 18, 2006, April 2, 2006, and May 22, 2006. This work was supported by the Center for Life Science Automation, Rostock, Germany.

M. Kumar and S. Kreuzfeld are with the Center for Life Science Automation, D-18119 Rostock, Germany.

M. Weippert, R. Vilbrandt, and R. Stoll are with the Institute of Preventive Medicine, Faculty of Medicine, University of Rostock, D-18055 Rostock, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2006.889825

The autonomic nervous system (ANS), which controls cardiac muscle, is of interest for stress detection. ANS is concerned with the regulation of heart rate, blood pressure, breathing rate, body temperature, and other visceral activities. The ANS activity is divided into two branches: sympathetic and parasympathetic, which influence the sinus node of the heart, thereby modulating heart rate. The sympathetic activity is primarily related to the preparation of body for stressful situations by *boosting of energy*. On the other hand, parasympathetic activity (most active under restful situations) counterbalances the effects of the sympathetic activity and restores the body to a resting state. Under normal situations, there is a balance between these two activities. However, during stress this balance will be altered and an analysis of heart rate signals could be potentially used to detect this alteration in system balance [9]–[14].

HRV is a measure of the variability in heart rate, i.e., variability in interbeat interval (IBI), which is defined as the time in milliseconds between consecutive R waves of an electrocardiogram. The IBI series (R-R intervals) can be analyzed using some mathematical theories (e.g., fast Fourier transform, wavelet theories, chaos) to assess ANS activities. The analysis of HRV provides a theoretical framework for ANS assessment by identifying the sympathetic and parasympathetic activities. The analysis of HRV in frequency domain could provide various information about cardiovascular control [15], [16]. Overall spectra of human HRV can be divided into three main frequency zones: below 0.04 Hz is very low frequency (VLF), between 0.04 and 0.15 Hz is low frequency (LF), and between 0.15 and 0.5 Hz is high frequency (HF). The LF is affected by both the sympathetic and parasympathetic activities, and the HF is found to be dominated by the parasympathetic activity [17]. The VLF is related to factors like temperature, hormones, etc. [18]. The ratio of the LF to HF power has been associated with the sympathovagal balance [18].

An emerging body of literature seems to suggest that HRV analysis can be potentially used to measure mental stress. However, a practical problem, which is so far not well addressed in the literature, is to derive some form of mathematical (quantitative) relations between parameters of ANS activity and mental stress. The problem can be formally stated as follows.

Problem 1: Given a 3-min IBI series (R-R intervals), estimate the level of mental stress on a scale ranging from 0 to 100 by monitoring the functioning of autonomic nervous system using HRV analysis.

The solution of above problem not only would provide a physiological interpretation of a so-called *mental stress* but also has the direct applications in engineering fields like adaptive automation and man–machine interface design. The major difficulties in solving Problem 1 are the following.

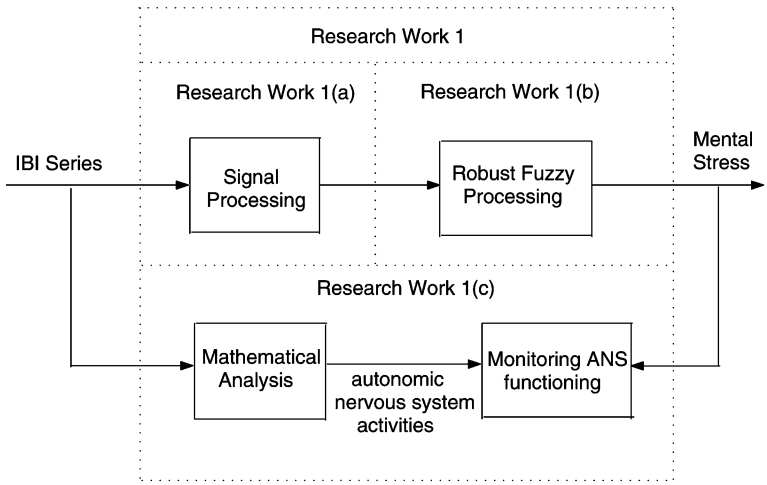

Fig. 1. Three components of research work.

1) HRV signal is nonstationary, i.e., it is characterized by time variations in its frequency components.
2) The alterations in HRV signal due to a change in the level of mental stress (these alterations would be used to detect the corresponding change) are subjective to individuals. The changes, which result in the HRV pattern of individuals due to a change in stress level, are different among individuals. This difference in behavior among the individuals may arise due to different body conditions, gender, age, physical fitness, emotional states, and so on.

To overcome the first difficulty, spectral components of HRV can be determined in the time-frequency domain using a wavelet transform [19]. Moreover, the logarithmic resolution of the wavelet transform enables the finer estimation of very low frequencies components. To resolve the second difficulty, some of the approaches coming from the area of neural network have been used (see [20] and references therein). This is done by combining HRV analysis with a neural network such that individual variations are learnt by a neural network. However, neural network–based approaches are like a "black box," since these do not provide a human-understandable insight into relationships between ANS activities and level of mental stress.

Fuzzy inference systems based on fuzzy set theory of Zadeh [21] are considered suitable for dealing with many real-world problems (which are characterized by complexities, uncertainties, and a lack of knowledge of the governing physical laws). The fuzzy systems not only are capable of learning complex input–output relationships but also offer additionally the advantage of providing an insight into relationships (i.e., interpretability). Thus, an identification of input–output mappings using a fuzzy model has become a common practice in literature; see, e.g., [22]–[29]. Also, in [30], a fuzzy inference system was used for interpreting ANS activities to estimate stress quantitatively.

The solving of Problem 1 would require, with respect to the state of art, a research work.

Research Work 1: This work involves development of a model (preferably fuzzy because of its uncertainties handling and interpretability advantages) robust enough to individual variations for establishing relationships between 3 min HRV analysis and mental stress (on a scale of 0–100). Once this model has been developed, the relationships between autonomic nervous system activities and mental stress could be established.

Fig. 1 shows the three components of Research Work 1. Research Work 1(a) deals with the extraction of IBI signal features in time-frequency domain using continuous wavelet transform. These features of IBI signal are interpreted to estimate the level of mental stress in component Research Work 1(b). In component Research Work 1(c), the sympathetic and parasympathetic activities of ANS are assessed using some mathematical analysis of IBI signal, and thus the functioning of ANS is monitored under different stress conditions. It is not difficult to realize that Research Work 1(b) is a bottleneck and plays the most important role in deciding the accuracy of the analysis. From a mathematical point of view, Research Work 1(b) is simply a problem of robust identification of a fuzzy model using input–output data in the presence of data uncertainties and modelling errors. Thus, we present a brief review of the literature in the field and identify the need of further research.

When it comes to the development of fuzzy models, a large number of fuzzy identification techniques have been developed using ad hoc approaches, neural networks, genetic algorithms, clustering techniques, and Kalman filtering [22], [24], [26], [28], [29], [31]–[40]. Most of the online fuzzy identification techniques in the literature use gradient-descent-based algorithms (such as backpropagation) for the adjustment of nonlinear fuzzy model parameters. However, in the presence of data uncertainties and modelling errors, gradient-descent-based techniques are not suitable due to their nonrobust nature. We believe that while studying fuzzy identification methods, the following issues should be addressed.

1) Fuzzy model identification is an ill-posed problem [41]. The identification of not only linear parameters (consequents) but also the nonlinear parameters (antecedents) of fuzzy model should mathematically take into account data uncertainties and modelling errors.
2) The identification procedure should be an online method.

3) The identification procedure should not require a priori knowledge of upper bounds, statistics, and distribution of data uncertainties and modelling errors.
4) The identification procedure should preserve the interpretability (a key property) of the fuzzy models.

These issues have been consider previously, e.g., in [23], [25], and [42]–[49]. To the authors' best knowledge, only the approach of [50] addresses all of the above concerns.

In this paper, our concern is to identify a fuzzy model that establishes the mappings between the clusters of input space and the desired output. To do so, we follow the proposed min-max approach of [51]. However, we apply to our problem a modified method based on a result of [50] to achieve fast convergence and low misadjustment error. This paper is organized as follows. In Section II, Research Work 1 will be formulated in a mathematical framework. Our research results on the topic will be presented in Section III. Afterwards, we discuss future research plans and offer concluding remarks.

II. Mathematical Formulations for a Fuzzy Based Approach to Stress Assessment

This section formulates mathematically the three components of Research Work 1, i.e., IBI signal processing, robust fuzzy processing, and assessing autonomic nervous system activities.

A. IBI Signal Processing

Many studies on the spectral analysis of HRV are based on fast Fourier transform (FFT) [12]. However, as discussed in previous section, IBI signal is nonstationary and should be analyzed in time-frequency domain. The typical methods of time-frequency decomposition (e.g., short-time Fourier transform, Wigner–Ville, continuous wavelet transform) divide the time span of the signal into successive windows and then assess the spectrum in each window. Thus, the changes in spectral components of the signal between successive windows provide a time variation of the spectral components. The length of the windows determines the time resolution. That is, the short windows provide a high time resolution and the long windows provide a low resolution. In any method of time-frequency decomposition, the time resolution (Δt) and the frequency resolution (Δf) obey the Heisenberg uncertainty principle: $\Delta t \Delta f \geq 2\pi$. The continuous wavelet transform uses a variable window length such that $\Delta t \propto 1/f$. This results in a high-frequency resolution in the lower frequency range of HRV spectra and thus enables a finer estimation of lower frequency content [52]. Fourier transform expresses a signal as a sum of sines and cosines waves, while continuous wavelet transform expresses a signal as a sum of *wavelets*. The wavelets are defined by the wavelet function $\psi(t)$. Mathematically, the continuous wavelet transform of a signal $s(t)$ at scale a and position b is defined by

$$W_\psi^s(a,b) = \frac{1}{\sqrt{a}} \int s(t)\psi\left(\frac{t-b}{a}\right) dt. \tag{1}$$

For the analysis of HRV in time-frequency domain, we choose our wavelet function $\psi(t)$ equal to the fifteenth derivative of the complex Gaussian function

$$f(t) = C_{15} e^{-it} e^{-t^2}$$

C_{15} is such that $\|f^{15}\|^2 = 1$, where f^{15} is the fifteenth derivative of f. Fig. 2(a) shows the plot of real and imaginary parts of our wavelet function. This wavelet function could be associated to a purely periodic signal of frequency 1 Hz. To see this, consider the FFT of wavelet function $\psi(t)$ in Fig. 2(b), that indicates 1 Hz to be the leading dominant frequency of $\psi(t)$. Thus, a frequency associated with $\psi(t/a)$ would be equal to $1/a$. This property of wavelet function is helpful in the interpretation of continuous wavelet transform. $W_\psi^s(a,b)$ can be interpreted as an estimate of the contribution of frequencies in a band around $1/a$ at time around b. Thus, the powers in frequency bands of signal (e.g., LF, HF), around time b, can be assessed by integrating $|W_\psi^s(a,b)|^2$ over the desired frequency band.

Our aim, in this section, is to extract the features of HRV in time-frequency domain. A 3-min continuous wavelet analysis of HRV in frequency range, say, 0.01–0.5 Hz (i.e., $a_i \in [2,100], b_j \in [0,3]$), will result in a two-dimensional matrix $[W_\psi^s(a_i,b_j)]$. It is not convenient to use this matrix as input of the fuzzy model. Moreover, the physiological interpretation of the matrix is unknown. Therefore, we suggest the following parameters for our analysis:

$$p_1 = \int_{0.01}^{0.04} M_W(1/f)\, df \tag{2}$$

$$p_2 = \int_{0.04}^{0.15} M_W(1/f)\, df \tag{3}$$

$$p_3 = \int_{0.15}^{0.5} M_W(1/f)\, df \tag{4}$$

where $M_W(a) = (1)/(n_b+1)\sum_{j=0}^{n_b} |W_\psi^s(a,b_j)|^2, b_j$ is the time index such that $b_0 = 0$ and $b_{n_b} = 3$ min.

It can be seen that parameters p_1, p_2, and p_3 assess in VLF, LF, and HF bands, respectively, the average power of the signal during 3 min, and thus are physiologically relevant. As an illustration, consider a 3-min IBI series extracted from the ECG. The R-R intervals were used to produce an HR signal in beats per minute, i.e., HR$= (60\,000)/$(R-R interval). Fig. 3 shows the analysis using the above defined complex wavelet function by plotting HR signal, color-plot of wavelet transform over scale-position plane, and average spectrum of the signal over 3 min.

B. Robust Fuzzy Processing

Now, we are concerned to analyze the parameters p_1, p_2, and p_3 using an interpretable fuzzy model (that is robust enough to individual variations) to estimate the level of mental stress on a scale of 0–100. This is done by first partitioning the input parameters space into different clusters and then establishing the fuzzy mappings between clusters and mental stress level. However, such a fuzzy model could not be completely constructed from a priori knowledge of experts because of the complexities and uncertainties (due to individual variations). Thus, the fuzzy

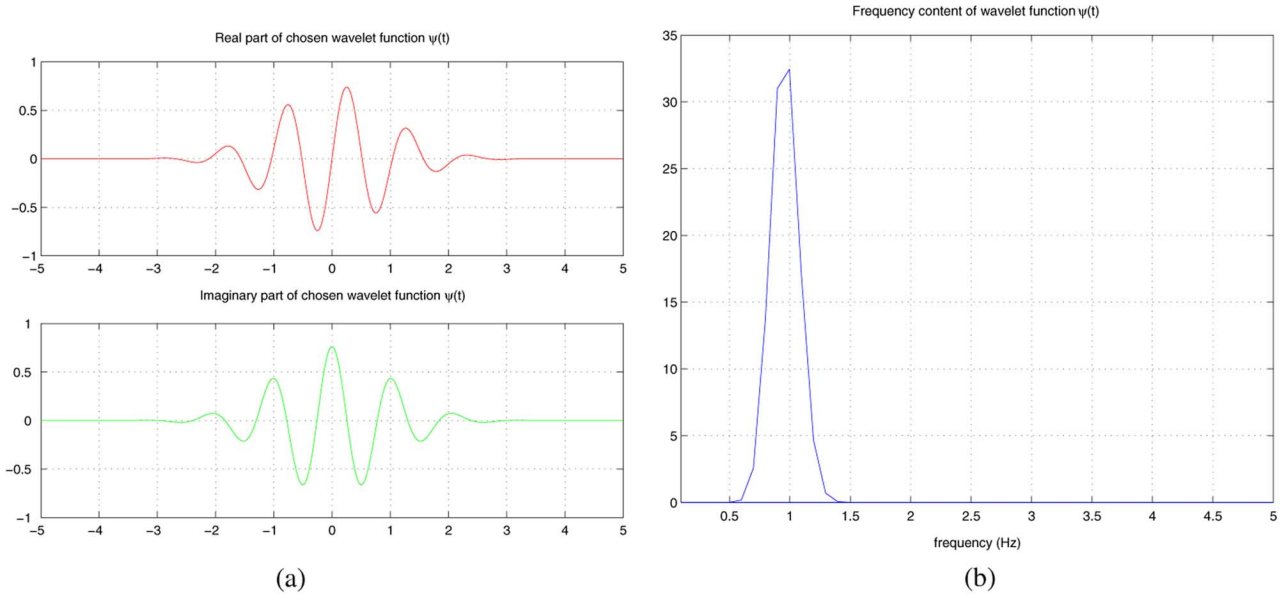

Fig. 2. A complex wavelet function. (a) Plot of the $\psi(t)$ and (b) FFT of the $\psi(t)$.

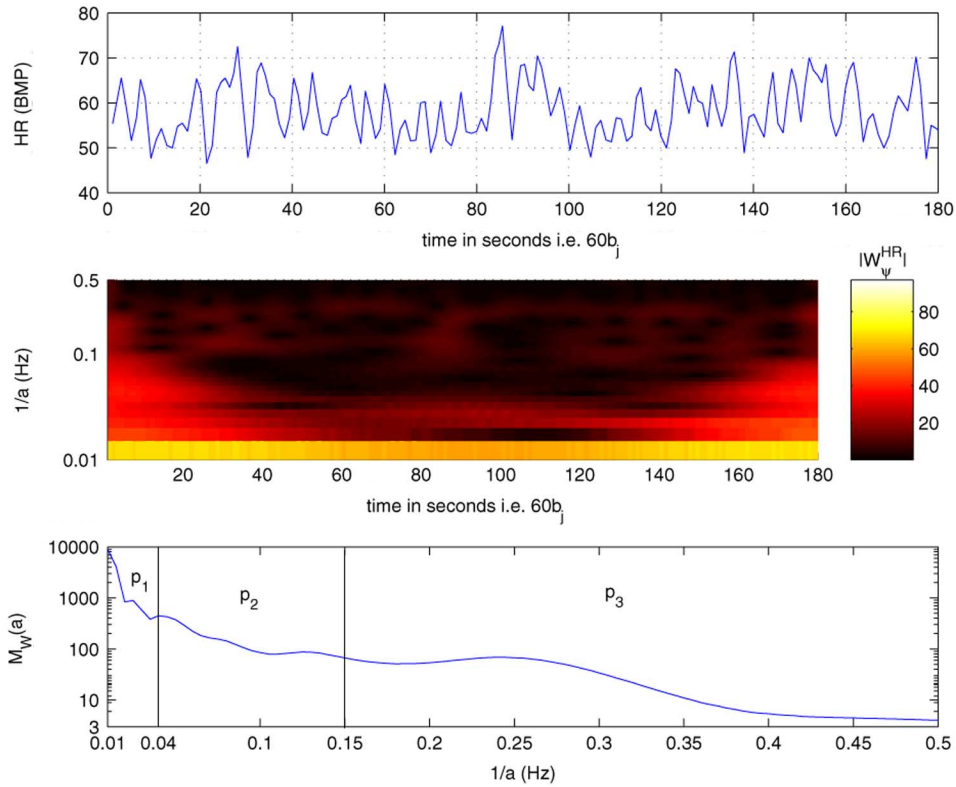

Fig. 3. HRV features in time-frequency domain.

model must be identified using input–output data. The inputs to the model are the parameters p_1, p_2, and p_3, while output is the quantification of mental stress. One can easily realize that the output identification data for individuals cannot be generated accurately, since the method to assess mental stress on a numerical scale would always involve an uncertainty in the assessment. In our experiments, we use NASA Task Load Index (TLX) for assessing the subjective rating score of mental workload [53]. The NASA Task Load Index is a method for providing an overall workload score based on a weighted average of ratings on six subscales (mental demands, physical demands, temporal demands, own performance, effort, and frustration). We assume that subjective mental workload is an uncertain measure of the mental stress level. That is

$$\text{subjective mental workload} = \text{mental stress} + \text{uncertainty}.$$

Hence, the input identification data consists of physiological response measuring parameters (e.g., p_1, p_2, p_3) and the output identification data is the subjective rating score of mental workload, as shown in Fig. 4. A major concern of this paper is to

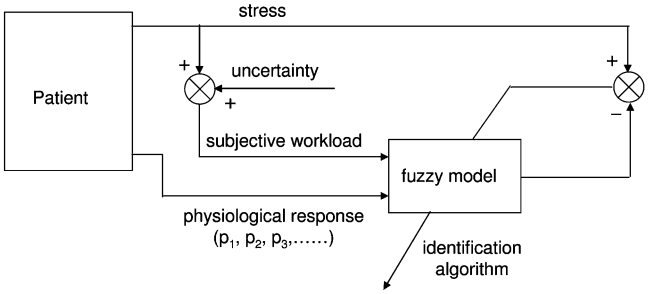

Fig. 4. The identification of a fuzzy model to estimate stress using physiological response parameters.

identify the fuzzy model. For this purpose, we first introduce a clustering based fuzzy model (as in [51]) and then formulate mathematically our problem of fuzzy identification.

1) A Clustering Based Fuzzy Model: Let us consider a Sugeno type fuzzy inference system ($F_s : X \rightarrow Y$), mapping n-dimensional input space ($X = X_1 \times X_2 \times \cdots \times X_n$) to a one-dimensional real line, consisting of K different following rules:

If x belongs to a cluster having centerc_i then $y = \alpha^i$

$i = 1, 2, \ldots, K$, where $x \in R^n$ is an n-dimensional input vector, $c_i \in R^n$ is the center of the ith cluster, and the values $\alpha^1, \ldots, \alpha^K$ are real numbers. Let $A_i(x)$ denote a multivariable membership function $A_i : X \rightarrow$[0,1] that represents the degree of membership of input vector $x \in X$ to the ith cluster. The different fuzzy rules can be aggregated as

$$F_s(x) = \frac{\sum_{i=1}^{K} \alpha^i A_i(x)}{\sum_{i=1}^{K} A_i(x)}. \tag{5}$$

For the fuzzy partition of input space X into K different clusters by the method of fuzzy c-means (FCM), the membership function $A_i(x)$ must satisfy (see [54])

$$\sum_{x \in X} \sum_{i=1}^{K} A_i^{\tilde{m}}(x) \|x - c_i\|^2 \rightarrow \text{Minimum} \sum_{i=1}^{K} A_i(x) = 1$$

where $\tilde{m} > 1$ is the *fuzzifier* and $\| \cdot \|$ denotes the Euclidean norm. The membership function that minimizes the above objective function for a given choice of cluster centers $\{c_i\}_{i=1}^{K}$ follows as

$$\text{FCM}_i(x, c_1, \ldots, c_K) = \begin{cases} \dfrac{1}{\sum_{j=1}^{K} \left(\dfrac{\|x - c_i\|^2}{\|x - c_j\|^2} \right)^{\frac{1}{\tilde{m}-1}}} & \text{for } x \in X \setminus \{c_j\}_{j=1,\ldots,K} \\ 1 & \text{for } x = c_i \\ 0 & \text{for } x \in \{c_j\}_{j=1,\ldots,K} \setminus \{c_i\}. \end{cases} \tag{6}$$

Fig. 5 shows a one-dimensional example of membership function $\text{FCM}_i(\cdot)$ for three different clusters with cluster centers at 20, 50, and 80 for $\tilde{m} = 2$. The membership functions constructed using (6) has an inherent limitation of being nonconvex, as seen from Fig. 5(a). That is, the points lying far away from the cluster center may be assigned more membership to the cluster than the points lying closer. The origin of the problem is the constraint that membership values' sum is equal to unity. In [55], a *possibilistic approach* for C-means clustering was introduced that relaxes the unit sum constraint on the membership values so that $A_i(x)$ better reflects the typicality of x to the ith cluster. This approach adopts the following objective function:

$$\sum_{x \in X} \sum_{i=1}^{K} \left[A_i^{\tilde{m}}(x) \|x - c_i\|^2 + (1 - A_i(x))^{\tilde{m}} \delta_i \right]$$

where δ_i is a positive scale parameter. The suggestions for choosing δ_i have been provided in [56]. Another approach, called the noise clustering method, has been introduced by Davé in [57] to deal with the noisy data. This approach considers noise a separate cluster such that membership of x to the noise cluster is defined as $1 - \sum_{i=0}^{K} A_i(x)$ and the noise prototype is always at the same distance from every point in the data set, say, δ. The objective function in the noise clustering approach is

$$\sum_{x \in X} \sum_{i=1}^{K} A_i^{\tilde{m}}(x) \|x - c_i\|^2 + \sum_{x \in X} \delta^2 \left(1 - \sum_{i=0}^{K} A_i(x) \right)^{\tilde{m}}.$$

The different clustering methods can be derived by modifying the possibilistic and noise clustering approaches. As an example, consider a clustering criterion that assumes a noise cluster outside each data cluster. We seek to minimize

$$\begin{aligned} &J_c(A_i(x), c_1, \ldots, c_K) \\ &\quad = \sum_{x \in X} \sum_{i=1}^{K} [A_i(x) \|x - c_i\|^2 \\ &\qquad + \{1 + A_i(x) \log A_i(x) - A_i(x)\} \delta_i] \end{aligned}$$

where the second term in the objective function is intended to be a noise cluster. The term $\{1 + A_i(x) \log A_i(x) - A_i(x)\}$ may be interpreted as the degree to which x does not belong to the ith cluster and thus the membership of x to the noise cluster. If the distance of x to the cluster center c_i is greater than $\sqrt{\delta_i}$, then the minimization of $J_c(\cdot)$ forces a small value of $A_i(x)$ and a large value of membership of x_i to the noise cluster. Therefore, one of the strategies may be to set δ_i equal to the distance of nearest cluster center from c_i, i.e.,

$$\delta_i = \min_j \|c_j - c_i\|^2.$$

Setting $(\partial J_c(A_i(x), c_1, \ldots, c_K))/(\partial A_i(x)) = 0$ leads to the following expression for optimal membership function:

$$\text{RC}_i(x, c_1, \ldots, c_K) = \exp\left(-\frac{\|x - c_i\|^2}{\delta_i} \right). \tag{7}$$

Fig. 5(b) shows the membership functions [(7)] for three different clusters with centers at 20, 50, and 80. Although the membership functions' shape is convex, the assignment of membership values to a cluster is totally independent of the location of other cluster centers. As an illustration, consider a point 50 (the center of second cluster) in Fig. 5(b) that has been assigned the

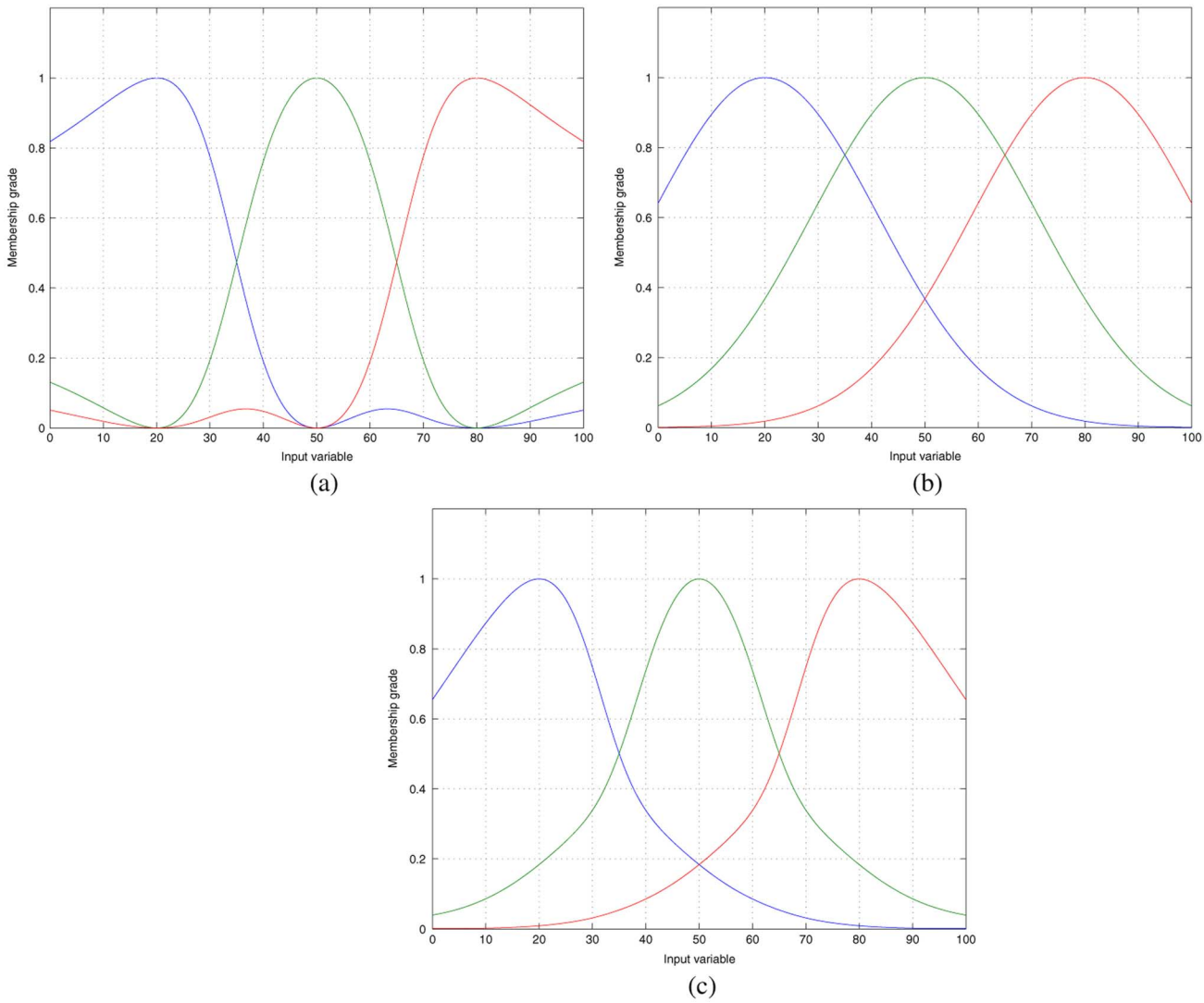

Fig. 5. Membership functions of three clusters in one-dimensional input space using: (a) (6), (b) (7), and (c) (8).

membership values more than 0.30 to the first and third clusters. This point, however, in case of FCM was assigned zero memberships to the first and third clusters (that was obviously the more natural assignment).

The membership functions of Fig. 5(a) and (b) can be combined by adopting a mixed clustering criterion [58], [59]. One way to do this is to assume the membership function A_i has two components A_{1i} and A_{2i} such that

$$A_i = \frac{A_{1i}^{\tilde{m}}}{2} + \frac{A_{2i}}{2}$$

where A_{1i}, A_{2i} minimizes the following constrained objective function:

$$\sum_{x \epsilon X} \sum_{i=1}^{K} \left[\left(A_{1i}^{\tilde{m}}(x) + A_{2i}(x) \right) \|x - c_i\|^2 + \{1 + A_{2i}(x) \log A_{2i}(x) - A_{2i}(x)\} \delta_i \right]$$
$$\sum_{i=1}^{K} A_{1i}(x) = 1.$$

Now, A_{1i} will be given by (6) and A_{2i} by (7). Thus

$$A_i(x, c_1, \ldots, c_K) = \frac{|\mathrm{FCM}_i(x, c_1, \ldots, c_K)|^{\tilde{m}}}{2} + \frac{\mathrm{RC}_i(x, c_1, \ldots, c_K)}{2}. \quad (8)$$

Fig. 5(c) shows a one-dimensional example of above membership function for three clusters (with centers at 20, 50, and 80) for $\tilde{m} = 2$. As seen from Fig. 5(c), (8) provides a compromise between the nonconvex nature of (6) and the independent behavior of (7) with respect to the location of cluster centers. If we denote

$$G_i(x, c_1, \ldots, c_K) = \frac{A_i(x, c_1, \ldots, c_K)}{\sum_{i=1}^{K} A_i(x, c_1, \ldots, c_K)}$$

then the output of Sugeno type fuzzy inference system for this shape of membership functions follows from (5) as

$$F_s(x) = \sum_{i=1}^{K} \alpha_i G_i(x, c_1, \ldots, c_K).$$

We introduce the following notations: $\alpha = [\alpha^i]_{i=1,\ldots,K} \in R^K, \theta = [c_1^T, \ldots, c_K^T]^T \in R^{Kn}$, and $G(x,\theta) = [G_i(x,\theta)]_{i=1,\ldots,K} \in R^K$. Now, the above expression can be rewritten as

$$F_s(x) = G^T(x,\theta)\alpha. \tag{9}$$

As indicated by the above expression, the output of a clustering-based fuzzy model is linear in consequents α and nonlinear in cluster centers θ.

2) Mathematical Formulation of Fuzzy Identification Problem: Let us consider a process described at the jth time-index as

$$y(j) = f(x(j)) + n_j$$

where n_j is the unknown uncertainty in output measurement $y(j)$. Our aim is to identify online a fuzzy model that approximates the unknown function f as closely as possible. At any time index j, let (α_j, θ_j) denote the identified fuzzy model parameters using identification data set $\{[x(0), y(0)], \ldots, [x(j), y(j)]\}$ and $\mathcal{F}$ the corresponding identification method. That is

$$(\alpha_j, \theta_j) = \mathcal{F}([x(0), y(0)], \ldots, [x(j), y(j)]).$$

The instantaneous identification error may be defined as

$$e_j = f(x(j)) - G^T(x(j), \theta_j)\alpha_j.$$

The fuzzy identification problem can be formally stated as follows.

Fuzzy Identification Problem 1: Given the identification data set $\{x(j), y(j)\}_{j=0}^k$, design an identification method $\mathcal{F}^*, \{\alpha_j, \theta_j\}_{j=0}^k = \mathcal{F}^*(\{x(j), y(j)\}_{j=0}^k)$ that, without making any statistical assumption and without requiring apriori knowledge about uncertainty signal n_j, minimizes the magnitude of identification error signal $|e_j|$ for all $j = 0, \ldots k$. That is

$$\sum_{j=0}^{k} |f(x(j)) - G^T(x(j), \theta_j)\alpha_j| \rightarrow \text{Minimum}.$$

The problem stated above cannot be solved directly, since uncertainty signal n_j is unknown. Therefore, we will consider some robust approaches to solve the problem.

C. Assessing Autonomic Nervous System Activities

Given the IBI signal and corresponding level of stress from Research Work 1(b) (see Fig. 1), our aim is to 1) assess the ANS activities and 2) study the effect of stress on ANS. We have seen that parameters p_1, p_2, and p_3 assess the power in VLF, LF, and HF bands, respectively. These powers have been largely used in the literature by associating them to ANS activities (i.e., LF is associated to sympathetic activity and HF to parasympathetic). Thus, the ANS activities can be assessed by defining sympathetic modulation index (SMI), vagal modulation index (VMI), and sympathovagal balance index (SVI) as

$$\text{SMI} = \frac{p_2}{p_2 + p_3}$$
$$\text{VMI} = \frac{p_3}{p_2 + p_3}$$
$$\text{SVI} = \frac{p_2}{p_3}. \tag{10}$$

However, LF/HF ratio (i.e., SVI) is not a true index of the sympathovagal balance, since LF is mediated not only by the sympathetic activity but also by the parasympathetic activity [60]. Some nonlinear methods, e.g., based on principal dynamic modes [61], could be used to separate the sympathetic and parasympathetic activities.

To visualize the effect of stress on ANS, we propose to identify the regions in ANS activity parameters space such that each region is characterized by some level of stress. For example, Fig. 10, as we will see, shows the identification of various stress regions on the two-dimensional ANS activity parameters space. The identification of stress regions should be independent of individual variations. Therefore, an important concern is to choose the suitable ANS activity parameters that provide robustness to individual variation, e.g., normalized increase in activities (as in Fig. 10).

III. Research Results

Our study involved:

1) a solution of the fuzzy identification problem 1 using a robust (min-max) approach;
2) computer simulations;
3) experiments with 39 individuals.

A. A Min-Max Solution to Fuzzy Identification Problem 1

By min-max solution, we mean a solution that minimizes the worst case effect of uncertainties on the performance of fuzzy parameters identification, and hence a min-max solution is robust to uncertainties.

Result 1: A min-max solution to fuzzy identification Problem 1 is to identify the fuzzy model parameters, for all $j = 0, \ldots, k$, as

$$\theta_j = \arg\min_{\theta} \psi_j(\theta) \tag{11}$$

$$\psi_j(\theta) = \frac{[y(j) - G^T(x(j),\theta)\alpha_{j-1}]^2}{1 + \mu\|G(x(j),\theta)\|^2} + \mu_\theta^{-1}\|\theta - \theta_{j-1}\|^2$$

$$\alpha_j = \alpha_{j-1} + \frac{\mu G(x(j),\theta_j)[y(j) - G^T(x(j),\theta_j)\alpha_{j-1}]}{1 + \mu\|G(x(j),\theta_j)\|^2} \tag{12}$$

where $\alpha_{-1} = 0, \mu, \mu_\theta$ are any positive constants, and θ_{-1} is an initial guess about cluster centers.

Proof: The result follows directly from [51]. ■

The above result suggests a robust method of fuzzy parameter learning that performs better than the standard clustering and gradient-descen-based learning methods in the presence of data uncertainties and modelling errors [51]. The parameters (μ, μ_θ) control the learning rate of the method. A higher learning rate leads to fast convergence and a higher misadjustment error, and vice versa. Thus, we like to use a variable or time-varying learning rate (i.e., $\mu(j)$ and $\mu_\theta(j)$) in recursions (11) and (12) to meet the requirements of fast convergence and low

misadjustment error. That is

$$\theta_j = \arg\min_{\theta} \psi_j(\theta) \tag{13}$$

$$\begin{aligned}\psi_j(\theta) &= \frac{[y(j) - G^T(x(j),\theta)\alpha_{j-1}]^2}{1+\mu(j)\|G(x(j),\theta)\|^2} + \frac{1}{\mu_\theta(j)}\|\theta - \theta_{j-1}\|^2 \\ \alpha_j &= \alpha_{j-1} \\ &\quad + \frac{\mu(j)G(x(j),\theta_j)[y(j) - G^T(x(j),\theta_j)\alpha_{j-1}]}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\end{aligned} \tag{14}$$

where $\alpha_{-1} = 0$. Let us for simplicity assume the ratio of antecedents learning rate to consequents learning rate at any time is constant, i.e.,

$$\frac{\mu_\theta(j)}{\mu(j)} = s_\theta$$

where $s_\theta > 0$ is a constant. Assume that input–output data $\{x(j), y(j)\}$ could be modelled as

$$y(j) = G^T(x(j),\theta_j)\alpha^* + n_j \tag{15}$$

where θ_j is given by (13) and α_j given by (14) is an estimate of some true vector α^*. Here, n_j includes measurement noise and modelling errors. For our analysis, we define the following error measures:

1) consequents-error vector $\tilde{\alpha}_j = \alpha^* - \alpha_j$;
2) a priori estimation error $e_a(j) = y(j) - G^T(x(j),\theta_j)\alpha_{j-1} = G^T(x(j),\theta_j)\tilde{\alpha}_{j-1} + n_j$;
3) a priori recursion error $G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}$;
4) a posteriori recursion error $G^T(x(j),\theta_j)\tilde{\alpha}_j$.

A possible strategy to choose values $\{\mu(j), \mu_\theta(j)\}$ for fast convergence and low misadjustment error is suggested in [50] for the identification of fuzzy models. The method of [50] for the identification of cluster centers and consequents is stated in Result 2.

Result 2: The learning of fuzzy model parameters using a variable learning rate follows as

$$\theta_j = \arg\min_{\theta} \psi_j(\theta) \tag{16}$$

$$\begin{aligned}\psi_j(\theta) &= \frac{|y(j) - G^T(x(j),\theta)\alpha_{j-1}|^2}{1 + \frac{\|\hat{p}_j(\theta)\|^2}{C}} \\ &\quad + \left(\frac{s_\theta\|\hat{p}_j(\theta)\|^2}{C\|G(x(j),\theta)\|^2}\right)^{-1}\|\theta - \theta_{j-1}\|^2 \\ \hat{p}_j(\theta) &= \omega\hat{p}_{j-1} \\ &\quad + \frac{(1-\omega)[y(j) - G^T(x(j),\theta)\alpha_{j-1}]G(x(j),\theta)}{\|G(x(j),\theta)\|^2}\end{aligned} \tag{17}$$

$$\mu^*(j) = \frac{\|\hat{p}_j(\theta_j)\|^2}{C\|G(x(j),\theta_j)\|^2} \tag{18}$$

$$\begin{aligned}\alpha_j &= \alpha_{j-1} \\ &\quad + \frac{\mu^*(j)G(x(j),\theta_j)[y(j) - G^T(x(j),\theta_j)\alpha_{j-1}]}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}\end{aligned} \tag{19}$$

where $\hat{p}_{-1} = 0, \alpha_{-1} = 0$.

Proof: See Appendix 1. ∎

To study the stability of the learning method [(16)–(19)], we associate to the learning method a discrete-time nonlinear system of the form

$$\tilde{\alpha}_j = f_1(\tilde{\alpha}_{j-1}, n_j, x(j), y(j), \theta_{j-1}, \hat{p}_{j-1}) \tag{20}$$

where $\tilde{\alpha}_j$ is the state of the system that changes from $\tilde{\alpha}_{j-1}$ to $\tilde{\alpha}_j$ during a recursion of the algorithm (16)–(19), and f_1 is a continuous nonlinear function. Let $\xi_{G_j}(j-1)$ denote the *a priori recursion* error (i.e., $\xi_{G_j}(j-1) = G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}$) and $\xi_{G_j}(j)$ denote the *a posteriori recursion* error (i.e., $\xi_{G_j}(j) = G^T(x(j),\theta_j)\tilde{\alpha}_j$). This transition of the error from $\xi_{G_j}(j-1)$ to $\xi_{G_j}(j)$, during a recursion of the algorithm, is due to the transition of state in system (20) from $\tilde{\alpha}_{j-1}$ to $\tilde{\alpha}_j$. Therefore, $\xi_{G_j}(j)$ (a linear combination of the elements of state vector $\tilde{\alpha}_j$) is a modified state of the discrete-time system (20).

The algorithm (16)–(19) will be considered stable if, for a bounded magnitude of disturbances $\{n_j\}$, the errors $\{\xi_{G_j}(j)\}$ remain bounded and, for a small disturbance signal, the error becomes small no matter what the initial error is. Linking the algorithm stability to the stability of the nonlinear system (20), we consider the input-to-state stability (ISS) (see, e.g., [62]), where input corresponds to n_j and state corresponds to $\xi_{G_j}(j-1)$. The ISS approach also has been considered in [44] to suggest some stable fuzzy learning schemes.

Let us first recall some definitions and results (from [62]) concerning ISS for discrete-time nonlinear systems. A function $\gamma : R_{\geq 0} \to R_{\geq 0}$ is a $\mathcal{K}$-function if it is continuous, strictly increasing, and $\gamma(0) = 0$. A function $\beta : R_{\geq 0} \times R_{\geq 0} \to R_{\geq 0}$ is a $\mathcal{KL}$-function if, for each fixed $t \geq 0$, the function $\beta(\cdot, t)$ is a $\mathcal{K}$-function and, for each fixed $s \geq 0$, the function $\beta(s,\cdot)$ is decreasing and $\beta(s,t) \to 0$ as $t \to \infty$. A function $\tau : R_{\geq 0} \to R_{\geq 0}$ is a $\mathcal{K}_\infty$-function if it is a $\mathcal{K}$-function and $\tau(s) \to \infty$ as $s \to \infty$.

Definition 1: The learning algorithm in (16)–(19) will be considered stable from disturbances $\{n_j\}$ to errors $\{\xi_{G_j}(j)\}$ if there exists a $\mathcal{KL}$-function $\beta : R_{\geq 0} \times R_{\geq 0} \to R_{\geq 0}$ and a $\mathcal{K}$-function $\gamma : R_{\geq 0} \to R_{\geq 0}$ such that, for each $n_j \in L_\infty$ (i.e., $\max_j |n_j| < \infty$) and each $\xi_{G_0}(-1) \in R$, it holds that

$$|\xi_{G_j}(j)| \leq \beta\left(|\xi_{G_0}(-1)|, j\right) + \gamma(\max_j |n_j|).$$

The following theorem provides the stability condition.

Theorem 1: If there exists a continuous function $V : R \to R_{\geq 0}, \mathcal{K}_\infty$-functions τ_1, τ_2, τ_3 and a $\mathcal{K}$-function σ such that

$$\begin{aligned}&\tau_1(|s|) \leq V(s) \\ &\quad \leq \tau_2(|s|), \quad \forall s \in R \\ &V\left(\xi_{G_j}(j)\right) - V\left(\xi_{G_j}(j-1)\right) \\ &\quad \leq -\tau_3\left(\left|\xi_{G_j}(j-1)\right|\right) + \sigma(|n_j|)\end{aligned}$$

$\forall\, \xi_{G_j}(j-1) \in R, \forall\, n_j \in R$, then the learning algorithm, from $\{n_j\}$ to $\{\xi(j)\}$, is stable.

Proof: The result follows from [62]. ∎

Result 3: The learning algorithm in (16)–(19), from disturbances $\{n_j\}$ to errors $\{\xi_{G_j}(j)\}$, is stable.

Proof: It follows from (19) using $y(j) - G^T(x(j),\theta_j)\alpha_{j-1} = \xi_{G_j}(j-1) + n_j$, that

$$\begin{aligned}\tilde{\alpha}_j &= \tilde{\alpha}_{j-1} \\ &\quad - \frac{\mu^*(j)G(x(j),\theta_j)\left[\xi_{G_j}(j-1)+n_j\right]}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2} \\ \xi_{G_j}(j) &= \xi_{G_j}(j-1) \\ &\quad - \frac{\mu^*(j)\|G(x(j),\theta_j)\|^2\left[\xi_{G_j}(j-1)+n_j\right]}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}.\end{aligned}$$

Define a positive and radically unbounded function $V(s) = s^2$ and consider $\Delta V_j = V(\xi_{G_j}(j)) - V(\xi_{G_j}(j-1))$

$$\begin{aligned}\Delta V_j &= \left[\frac{1}{\left(1+\mu^*(j)\|G(x(j),\theta_j)\|^2\right)^2} - 1\right]|\xi_{G_j}(j-1)|^2 \\ &\quad + \left[\frac{\mu^*(j)\|G(x(j),\theta_j)\|^2}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}\right]^2 |n_j|^2 \\ &\quad - \frac{2\mu^*(j)\|G(x(j),\theta_j)\|^2\xi_{G_j}(j-1)n_j}{\left(1+\mu^*(j)\|G(x(j),\theta_j)\|^2\right)^2}.\end{aligned}$$

The above expression implies

$$\begin{aligned}\Delta V_j &\le \left[\frac{1}{(1+\mu^*(j)\|G(x(j),\theta_j)\|^2)^2} - 1\right]|\xi_{G_j}(j-1)|^2 \\ &\quad + \left[\frac{\mu^*(j)\|G(x(j),\theta_j)\|^2}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}\right]^2 |n_j|^2 \\ &\quad + \frac{2\mu^*(j)\|G(x(j),\theta_j)\|^2|\xi_{G_j}(j-1)||n_j|}{(1+\mu^*(j)\|G(x(j),\theta_j)\|^2)^2}.\end{aligned}$$

Using inequality $2|\xi_{G_j}(j-1)||n_j| \le |\xi_{G_j}(j-1)|^2 + |n_j|^2$, we have

$$\begin{aligned}\Delta V_j &\le \frac{\mu^*(j)\|G(x(j),\theta_j)\|^2|n_j|^2}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2} \\ &\quad - \frac{\mu^*(j)\|G(x(j),\theta_j)\|^2|\xi_{G_j}(j-1)|^2}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}.\end{aligned}$$

Substituting $\mu^*(j)$ from (18), the above inequality becomes

$$\begin{aligned}\Delta V_j \le &\frac{\|\hat{p}_j(\theta_j)\|^2}{C+\|\hat{p}_j(\theta_j)\|^2}|n_j|^2 \\ &- \frac{\|\hat{p}_j(\theta_j)\|^2}{C+\|\hat{p}_j(\theta_j)\|^2}|\xi_{G_j}(j-1)|^2.\end{aligned}$$

Defining a constant $\pi = \min_j \|\hat{p}_j(\theta_j)\|^2/(C+\|\hat{p}_j(\theta_j)\|^2), 0 \le \pi < 1$, it follows from the above inequality that

$$V\left(\xi_{G_j}(j)\right) - V\left(\xi_{G_j}(j-1)\right) \le -\pi|\xi_{G_j}(j-1)|^2 + |n_j|^2.$$

If we define $\mathcal{K}_\infty$-functions $\tau_1(s) = |s|\min(1,|s|), \tau_2(s) = s^2 + |s|, \tau_3(s) = \pi s^2$ and a $\mathcal{K}$-function $\sigma(s) = s^2$, then

$$\begin{aligned}&\tau_1(|s|) \le V(s) \le \tau_2(|s|) \\ &V\left(\xi_{G_j}(j)\right) - V\left(\xi_{G_j}(j-1)\right) \\ &\quad \le -\tau_3\left(|\xi_{G_j}(j-1)|\right) + \sigma(|n_j|)\end{aligned}$$

and thus by Theorem 1, the result follows. ∎

Now, we study the convergence properties of the learning algorithm. For this purpose, consider from (19) that

$$\tilde{\alpha}_j = \tilde{\alpha}_{j-1} - \frac{\xi_{G_j}(j-1) - \xi_{G_j}(j)}{\|G(x(j),\theta_j)\|^2}G(x(j),\theta_j).$$

Taking the squared norm on both sides, we have

$$\begin{aligned}\|\tilde{\alpha}_j\|^2 + \frac{|\xi_{G_j}(j-1)|^2}{\|G(x(j),\theta_j)\|^2} \\ = \|\tilde{\alpha}_{j-1}\|^2 + \frac{|\xi_{G_j}(j)|^2}{\|G(x(j),\theta_j)\|^2}. \quad (21)\end{aligned}$$

To study the convergence properties, assume that $n_j = 0$; then we have

$$\xi_{G_j}(j) = \frac{\xi_{G_j}(j-1)}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}$$

and (21) is reduced to

$$\begin{aligned}\|\tilde{\alpha}_j\|^2 &= \|\tilde{\alpha}_{j-1}\|^2 - \frac{|\xi_{G_j}(j-1)|^2}{\|G(x(j),\theta_j)\|^2} \\ &\quad + \left|\frac{1}{1+\mu^*(j)\|G(x(j),\theta_j)\|^2}\right|^2 \frac{|\xi_{G_j}(j-1)|^2}{\|G(x(j),\theta_j)\|^2}.\end{aligned}$$

Substituting $\mu^*(j)$ leads to

$$\begin{aligned}\|\tilde{\alpha}_j\|^2 &= \|\tilde{\alpha}_{j-1}\|^2 \\ &\quad - \left(1 - \left|\frac{C}{C+\|\hat{p}_j(\theta_j)\|^2}\right|^2\right)\frac{|\xi_{G_j}(j-1)|^2}{\|G(x(j),\theta_j)\|^2}.\end{aligned}$$

The above expression shows the convergence property in a sense that the squared norm of consequents-error vector $\tilde{\alpha}_j$ is a nonincreasing function of time index j. Further, it indicates that a smaller value of C should lead to a faster convergence and vice versa.

The proposed method for the identification of fuzzy model parameters can be implemented using a Gauss–Newton based algorithm provided in Appendix II.

B. Simulation Studies

The purpose of simulation studies is to compare the proposed approach with the standard techniques and to verify the fast convergence and low misadjustment error of the learning algorithm. Consider a nonlinear function

$$\begin{aligned}f(x_1,x_2) &= (1-x_1x_2)e^{-(x_1+x_2)^2} \\ &\quad - \cos(4x_1x_2) + \log(1+x_1x_2)\end{aligned}$$

where $x_1 \in [-0.9, 0.9]$ and $x_2 \in [-0.9, 0.9]$. The goal is to identify the function f using the identification data sequence

$$\left\{\begin{bmatrix}x_1(j)\\x_2(j)\end{bmatrix}, y(j)\right\}$$

generated according to

$$y(j) = f(x_1(j), x_2(j)) + n_j$$

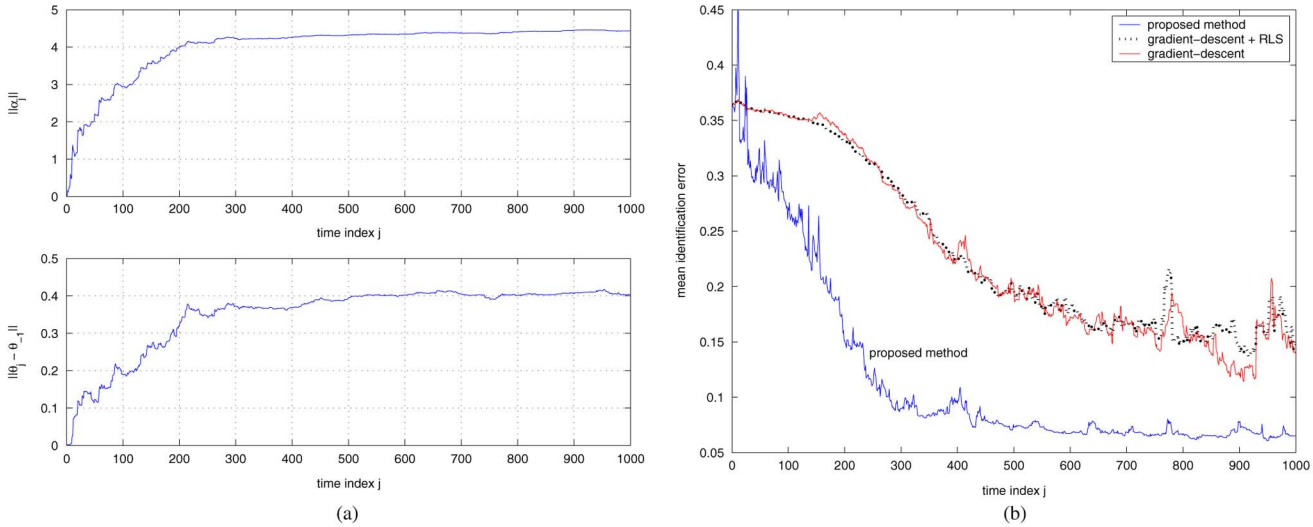

Fig. 6. The learning of fuzzy model parameters using noisy data. (a) Estimation of cluster centers and consequents. (b) Comparison with gradient-descent and a hybrid algorithm.

where n_j is a random signal normally distributed with zero mean, a variance of 0.01, and $\{x_1(j), x_2(j)\}$ are chosen from a uniform distribution on the interval [−0.9,0.9]. To filter the disturbance signal n_j from measurement signal $y(j)$ by approximating the function f using a fuzzy model, we employ following techniques:

1) standard gradient-descent for the identification of cluster centers and consequents;
2) a hybrid algorithm that uses gradient-descent for the identification of nonlinear cluster centers parameters and recursive least squares estimation (RLS) for linear consequents parameters;
3) the proposed method, i.e., (16)–(19).

Let us choose a fuzzy model consisting of nine different rules (i.e., nine different clusters in the two-dimensional input space). The initial guess about cluster centers is taken as

$$[c_1^T, \ldots, c_5^T] = \begin{bmatrix} -0.9 & -0.9 & -0.9 & 0 & 0 \\ -0.9 & 0 & 0.9 & -0.9 & 0 \end{bmatrix}$$
$$[c_6^T, \ldots, c_9^T] = \begin{bmatrix} 0 & 0.9 & 0.9 & 0.9 \\ 0.9 & -0.9 & 0 & 0.9 \end{bmatrix}.$$

The simulation parameters were chosen as $\tilde{m} = 2, C = 0.005, \omega = 0.99$, and $s_\theta = 0.05$. To measure the identification performance, let us define the instantaneous value of mean identification error (MIE) as

$$\text{MIE}(j) = \frac{1}{100} \sum_{l=1}^{100} |f(x_1(l), x_2(l)) - G^T(x(l), \theta_j)\alpha_j|$$

where the points $\{x(l) = [x_1(l)x_2(l)]^T\}_{l=1,\ldots,100}$ are uniformly distributed in the two-dimensional input space. The learning of model parameters using (16)–(19) has been shown in Fig. 6(a) by plotting $\|\alpha_j\|$ and $\|\theta_j - \theta_{-1}\|$ with j. Fig. 6(b) shows, by plotting the instantaneous value of mean identification error, the better performance of our approach than the gradient-descent and hybrid method (i.e., gradient descent + RLS). It was observed that both the gradient-descent and hybrid methods show an improved performance for a step-size around 0.09. Thus, for the simulation studies, their step-size was manually tuned to a value of 0.085. Finally, the approximation of the unknown function by a clustering-based fuzzy model is shown in Fig. 7. This example has clearly shown the potential of the proposed approach in solving filtering and identification problems.

Let us now compare the following methods:

1) constant learning rate strategy [(11) and (12)] for $\mu = 0.5, \mu_\theta = 0.5 s_\theta, s_\theta = 0.05$;
2) constant learning rate strategy [(11) and (12)] for $\mu = 1.0, \mu_\theta = 1.0 s_\theta, s_\theta = 0.05$;
3) variable learning rate strategy [(16)–(19)] for $s_\theta = 0.05, C = 0.005$.

Fig. 8 shows, for the different strategies, the plot of expected mean identification error curve (i.e., $\{E[\text{MIE}(j)]\}$) from time index $j = 0$ to $j = 2000$. The expected mean identification curves have been obtained by averaging the curves over 100 independent experiments. Fig. 8(a) shows a faster convergence of $\mu = 1.0$ curve than the $\mu = 0.5$ curve. For a clear comparison of the misadjustment errors, a portion of the curves from $j = 500$ to $j = 2000$ has been shown separately in Fig. 8(b). As seen from Fig. 8(b), the misadjustment error for the $\mu = 1.0$ curve is higher than the $\mu = 0.5$ curve. The proposed method, as seen from Fig. 8(a) and (b), meets the conflicting requirements of fast convergence and low misadjustment error.

Regarding the simulation parameters, $\tilde{m} = 2, \omega = 0.99$, and $s_\theta = 0.05$ provide in general good results. However, the parameter C for a given problem should be tuned manually. It follows from (25) and (26) that C should be chosen as

$$C = \sigma_n^2 E\left[\frac{1}{\|G(x(j), \theta_j)\|^2}\right]$$

where $\sigma_n^2 = E|n_j|^2$. Therefore, C should be chosen proportional to σ_n^2 (i.e., proportional to the level of involved uncertainties).

The automatic tuning of parameter C, based on an adaptive estimation of uncertainty signal variance, would be a topic of our future research. As an illustration, consider that the random

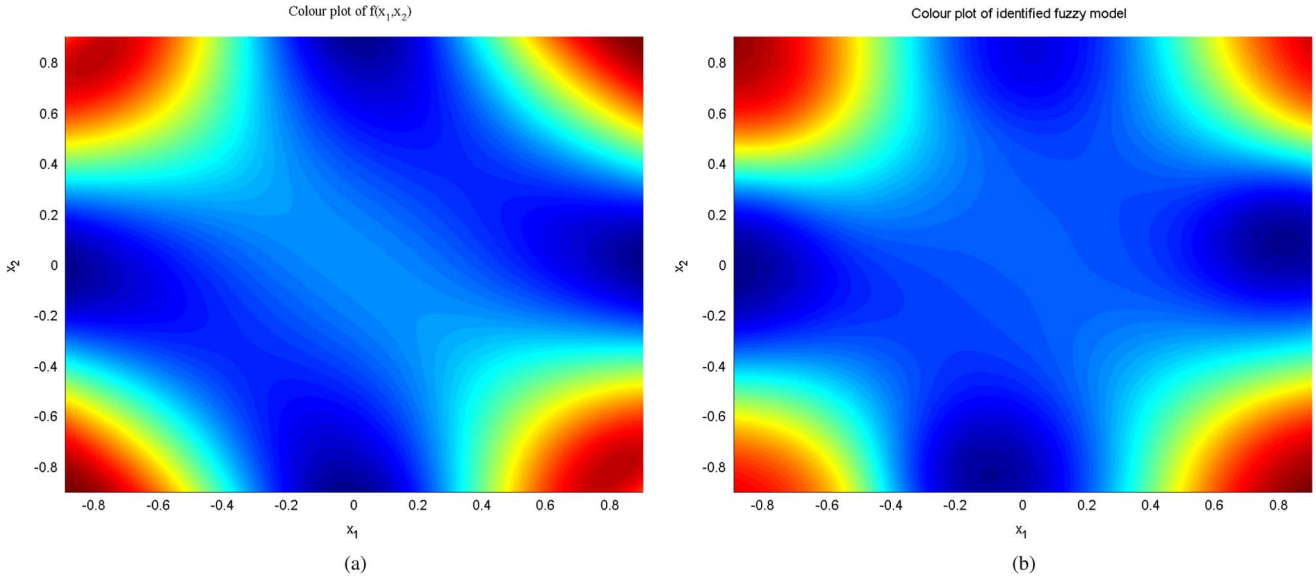

Fig. 7. Fuzzy identification of a nonlinear function. (a) Color plot of the function $f(x_1, x_2)$. (b) Color plot of the identified fuzzy model.

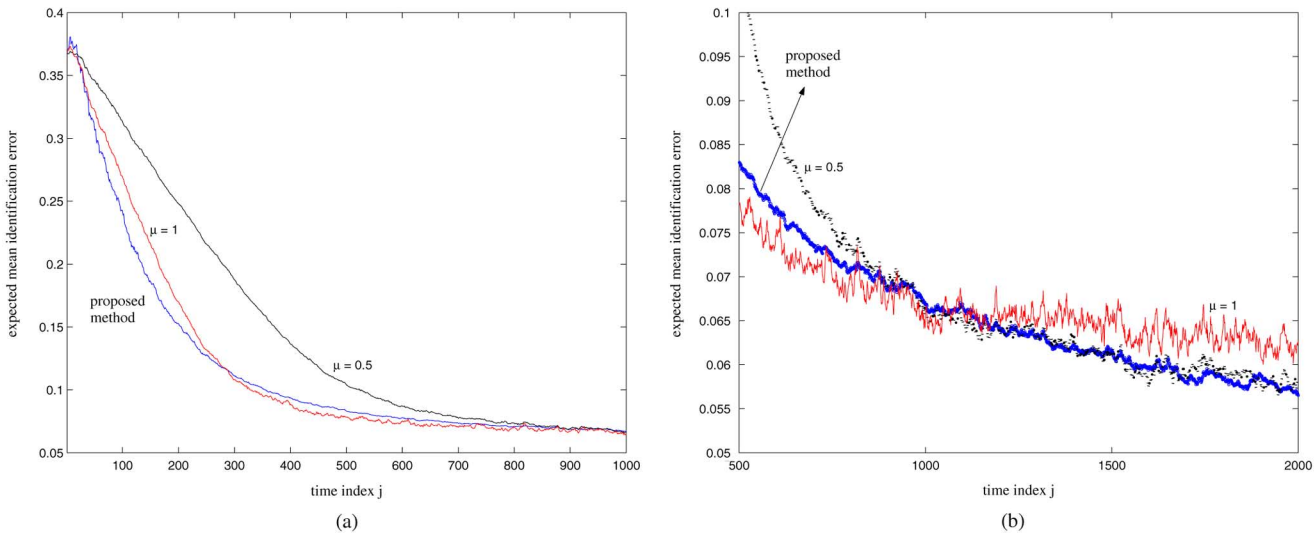

Fig. 8. Plot of expected mean identification error curve (averaged over 100 independent experiments). (a) From $j = 0$ to $j = 1000$. (b) From $j = 500$ to $j = 2000$.

variable $r_j = 1/\|G(x(j), \theta_j)\|^2$ has a uniform distribution on the interval $[1, K]$, since $1/K \leq \|G(x(j), \theta_j)\|^2 \leq 1$ (here, K is the number of rules in fuzzy model). Thus

$$E\left[\frac{1}{\|G(x(j), \theta_j)\|^2}\right] = \frac{1+K}{2}.$$

We are concerned with a model of the form

$$\hat{C}(j) = \hat{\sigma}_n^2(j)\frac{1+K}{2}$$

where $\hat{\sigma}_n(j)$ is the estimated variance of uncertainty signal and $\hat{C}(j)$ is the chosen value of parameter C at the jth time index. Alternatively, one could seek to render robustness in the fuzzy identification procedure against the choice of parameter C. Again, this would be a topic of our future work. We would study a fuzzy combination approach to address the issues related to the choice of C. The combination approach will be based on following ideas.

1) Independent fuzzy models are identified taking different values of parameter C.
2) During the identification of independent models, a fuzzy model is constructed to combine the independent models.
3) The construction of the combining fuzzy model is in such a way that the combination scheme provides simultaneously the benefits of the individual components.

C. Experiments

The experiments involve 38 healthy and physically fit volunteers (26 male, 12 female, aged 18–29 years). They were asked to refrain from using coffee and nicotine in the morning before and during the experiment. In our experiments, we used

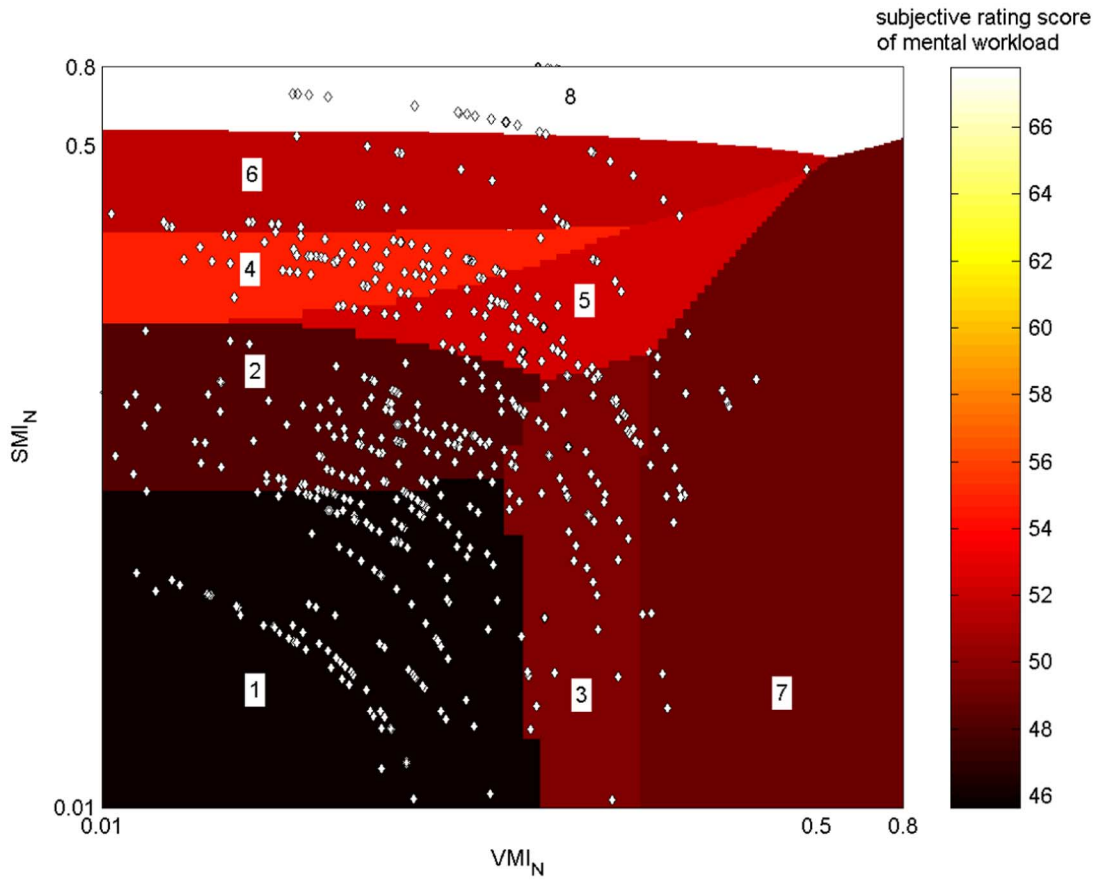

Fig. 9. Fuzzy c-means clustering of data.

a dual-task scenario including a dynamic control task and a secondary gauge-monitoring task. The mental stress among subjects was generated by the MultiTask simulator, a low-fidelity air traffic control (ATC) simulation, that has been used in various experiments in the context of air traffic control and adaptive automation (see, e.g., [63]).

All subjects attended three 30-min ATC trials of different load and type of automation. During the experiments, IBI signal was recorded by a Polar S810i (Polar Electro Oy, Finland, time-resolution 1 ms) heart rate monitor. The NASA Task Load Index has been used to assess the subjective rating score of mental workload. For every 3-min IBI signal, parameters p_1, p_2, and p_3 were calculated using (2)–(4). Further, an index of vagal modulation and sympathetic modulation is calculated using (10). Also, for each subject, the base values (i.e., minimum values) of vagal modulation index and sympathetic modulation index have been observed. To avoid individual variations in the analysis of data, a normalized increase of vagal and sympathetic modulation indexes from base values has been taken, i.e.,

$$\mathrm{VMI}_N = \frac{\mathrm{VMI} - \mathrm{VMI}_b}{\mathrm{VMI}_b} \quad \text{and} \quad \mathrm{SMI}_N = \frac{\mathrm{SMI} - \mathrm{SMI}_b}{\mathrm{SMI}_b}.$$

Fig. 9 shows the partition of data into eight different clusters, marked as $1, 2, \ldots, 8$, by the method of fuzzy c-means. In Fig. 9, the points represent data and the color of a cluster corresponds to the average value of mental workload associated with the cluster. It can be seen in Fig. 9 by comparing the eighth cluster with others that a higher level of mental stress corresponds to an increased value of sympathetic modulation index. Several literature findings indicate that mental stress leads to an increase in sympathetic modulation index and a decrease in vagal modulation index; see, e.g., [9]–[11] and [64]. However, the results of Fig. 9 do not clearly demonstrate an increase in sympathetic modulation index and a decrease in vagal modulation index due to an increase in stress level. As an illustration, while moving along SMI-axis from cluster 1 to 8, the stress level does not always increases, since stress associated to the 4th cluster is higher than the 6th cluster. Also, moving along VMI-axis, there is no significant reduction in the stress level. The reason obviously is that the different individuals behave differently, and an average behavior in Fig. 9 may not represent the real facts. Therefore, we need to handle the uncertainties due to individual variations using a fuzzy inference system.

Let us assume that the mental stress is some unknown function of parameters $(\mathrm{VMI}_N, \mathrm{SMI}_N)$ that needs to be approximated by a fuzzy model using subjective mental workload score (which is an uncertain measure of the mental stress). That is

$$\text{workload sore} = \text{mental stress} + \text{uncertainty}$$
$$\text{mental stress} = f(\mathrm{VMI}_N, \mathrm{SMI}_N).$$

The aim is to filter out the uncertainty (arising due to individual variations) by identifying the unknown function $f(\mathrm{VMI}_N, \mathrm{SMI}_N)$. We find this problem as an application of the developed filtering technique in Result 2. Please note that

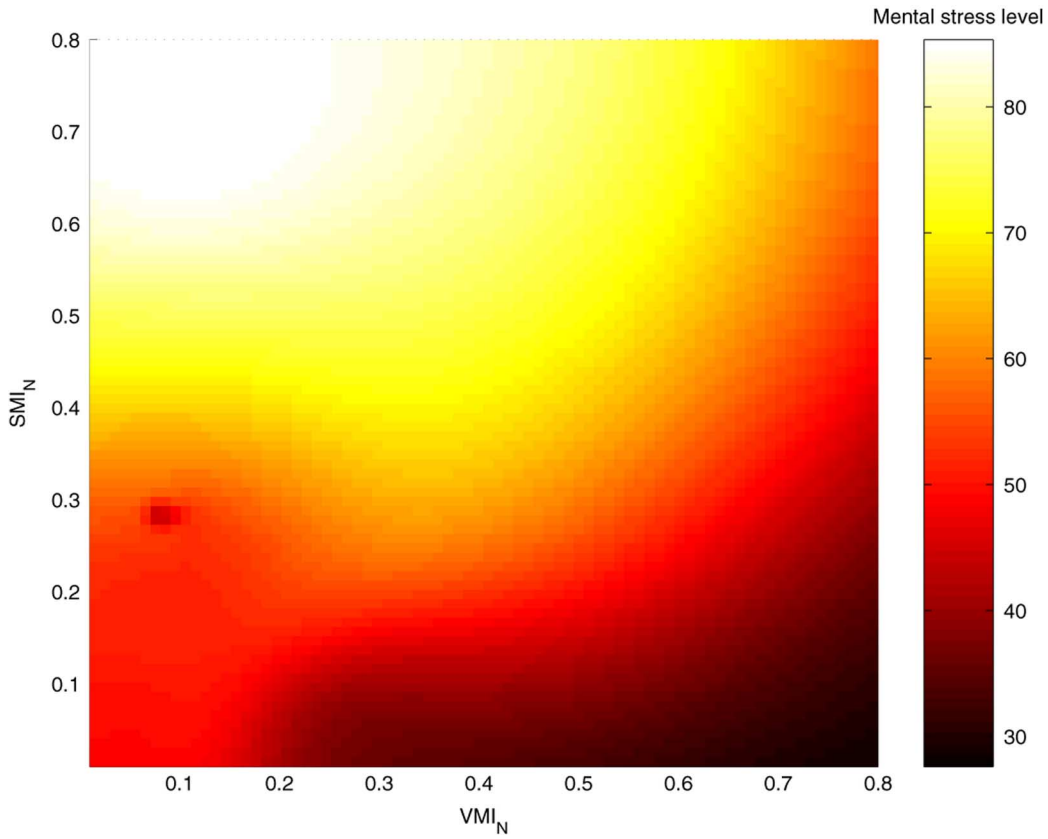

Fig. 10. Mental stress assessment by a fuzzy model identified using the proposed approach.

the problem here is different from standard fuzzy identification problem, where data could be divided into training and test data. The data here, being uncertain, cannot be used for the validation of the identified fuzzy model.

The proposed fuzzy filtering and identification method in (16)–(19) has been applied with $K = 8, \tilde{m} = 2, C = 5, \omega = 0.99$, and $s_\theta = 0.05$. Finally, the color plot of Fig. 10 shows the assessment of mental stress from parameters $(\text{VMI}_N, \text{SMI}_N)$ using the identified fuzzy model.

Fig. 10 clearly shows an increase of SMI_N and a decrease of VMI_N with an increase in the level of mental stress, and thus agrees completely with the previous literature findings and physiological understanding of mental stress that sympathetic activity is primarily related to the preparation of body for stressful situations and parasympathetic or vagal activity (most active under restful situations) counterbalances the effects of the sympathetic activity and restores the body to a resting state. By comparing the results of Figs. 9 and 10, one could easily appreciate the good performance of the identified fuzzy model from a physiological viewpoint that:

1) low sympathetic activity and high vagal activity $\rightarrow$ low stress.
2) high sympathetic activity and low vagal activity $\rightarrow$ high stress.

Hence, we see that the identified fuzzy model is capable of filtering out uncertainties in the assessment of mental stress. For the sake of comparison, let us investigate the potential of a linear model of the form

$$\text{stress} = \alpha_1 \text{SMI}_N + \alpha_2 \text{VMI}_N + \alpha_3$$

in filtering out the uncertainties by identifying the relationship between stress and $(\text{SMI}_N, \text{VMI}_N)$. To identify the parameters $(\alpha_1, \alpha_2, \alpha_3)$, we employ classical H^2 filtering i.e., RLS, and robust H^∞ filtering, i.e., normalized least mean squares algorithm (NLMS). A step-size of 0.005 was used for RLS and NLMS in identifying the model parameters till the algorithms nearly converged. Fig. 11 shows the plots of mental stress surface over $\text{SMI}_N - \text{VMI}_N$ plane drawn using the identified linear models. As seen from Fig. 11, the linear models, unlike the fuzzy model, could not capture the nonlinear relationships between physiological parameters and mental stress level. The linear models in Fig. 11, unlike fuzzy model, fail to provide any insight into the relationships between VMI_N and mental stress level.

We have seen that a fuzzy model, identified using the proposed approach, could handle the involved uncertainties (arising due to individual variation) in the assessment of mental stress based on the interpretation of physiological parameters. Now, we study two more nonlinear fuzzy filters for filtering out the uncertainties in mental stress assessment. The fuzzy filters are of same structure as considered in this text [i.e., (9)], however, the

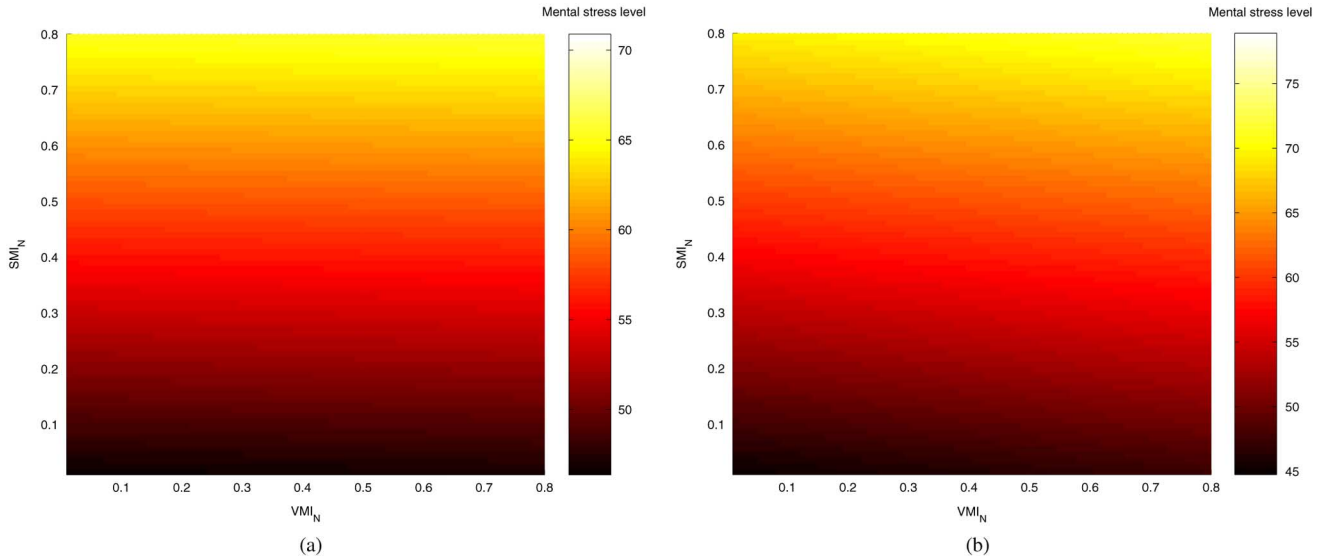

Fig. 11. Mental stress assessment using linear models. Model parameters identification using (a) RLS algorithm and (b) NLMS algorithm.

parameters are identified using standard techniques. The nonlinear parameters of the fuzzy models are identified using gradient descent learning

$$\theta_k = \theta_{k-1} - \mu \left(\frac{\partial E(\theta, k)}{\partial \theta} \right)_{\theta_{k-1}}$$
$$E(\theta, k) = \frac{1}{2}[y_k - G^T(x_k, \theta)\alpha_{k-1}]^2$$

where y_k is equal to subjective mental workload sore and $x_k = [\mathrm{VMI}_N(k)\ \mathrm{SMI}_N(k)]^T$ for the kth data index. Once θ_k has been computed, the data could be modelled as

$$y_k = G^T(x_k, \theta_k)\alpha^* + n_2(k)$$

where α^* is some true vector and n_2 includes uncertainties and modelling errors. For the identification of linear parameter α^*, we consider the following.

1) RLS algorithm

$$\alpha_k = \alpha_{k-1} + \frac{P_k G(x_k, \theta_k)(y_k - G^T(x_k, \theta_k)\alpha_{k-1})}{1 + G^T(x_k, \theta_k) P_k G(x_k, \theta_k)}$$
$$P_{k+1} = P_k - \frac{P_k G(x_k, \theta_k) G^T(x_k, \theta_k) P_k}{1 + G^T(x_k, \theta_k) P_k G(x_k, \theta_k)}, P_0 = \mu I$$

that minimizes the expected error energy

$$E\left[\sum_{j=0}^{k} |G^T(x_j, \theta_j)\alpha^* - G^T(x_j, \theta_j)\alpha_{j-1}|^2 \right]$$

assuming α^* and n_2 are zero-mean uncorrelated random variables with variances $\mu I \succ 0$ and unity, respectively.

2) NLMS algorithm

$$\alpha_k = \alpha_{k-1} + \frac{\mu G(x_k, \theta_k)(y_k - G^T(x_k, \theta_k)\alpha_{k-1})}{1 + \mu \|G(x_k, \theta_k)\|^2}$$

that minimizes the maximum value of energy gain from disturbances to estimation error and ensures that

$$\frac{\sum_{j=0}^{k} |G^T(x_j, \theta_j)\alpha^* - G^T(x_j, \theta_j)\alpha_j|^2}{\mu^{-1}\|\alpha^* - \alpha_{-1}\|^2 + \sum_{j=0}^{k} |n_2(j)|^2} \leq 1.$$

Two different fuzzy models (one based on RLS and other on NLMS) were identified to model the relationships between mental stress and $(\mathrm{VMI}_N, \mathrm{SMI}_N)$. A step-size of $\mu = 0.01$ was taken for the identification algorithms, since it provided a good compromise between fast convergence and low misadjustment error. Fig. 12 shows the color plots of mental stress surface drawn using the two identified fuzzy models. A comparison among Figs. 10,–12 from the above-discussed physiological viewpoint shows that 1) the nonlinear fuzzy filters, as expected, perform better than linear filters and 2) the proposed method of fuzzy model identification performs better than the standard techniques. This can be seen, e.g., by comparing the upper right corner of Fig. 12(a) and (b) with that of Fig. 10. The stress level in the upper right corner, as in Fig. 10, should be higher than the level in the lower right corner.

IV. Conclusion and Future Work

This paper has suggested a new approach to stress assessment by interpreting the parameters of autonomic nervous system activity using a fuzzy model. The most important contribution of this paper is to quantify the concept of mental stress and to establish a direct functional relationship (independent of individual variations) between ANS activities and mental stress. It seems that the proposed fuzzy evaluation approach, combined with linear (or nonlinear) analysis of physiological parameters, may be helpful in various physiological modelling problems for the handling of uncertainties. The fuzzy evaluation approach, after being applied to mental stress assessment, motivates the start of new interesting research of physical and mental stress separation.

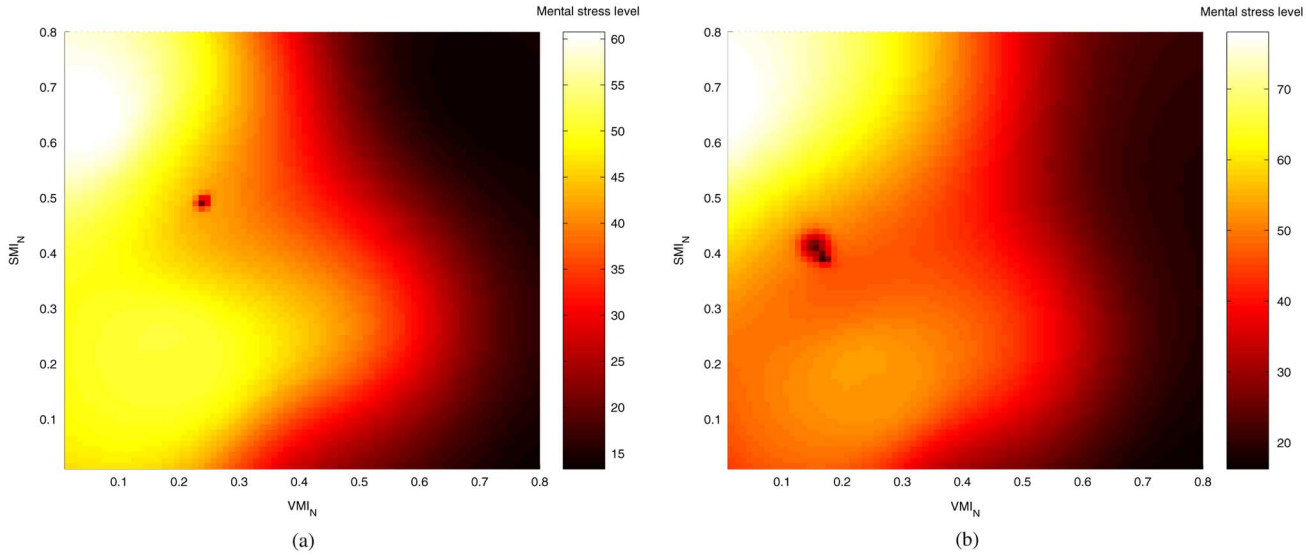

Fig. 12. Mental stress assessment using fuzzy models. Model parameters identification using (a) gradient descent and RLS and (b) gradient descent and NLMS.

Research Work 2: Development of a model (preferably fuzzy because of its uncertainties handling and interpretability advantages), robust enough to individual variations, for separating the physical and mental stress by interpreting the parameters of autonomic nervous system activity.

APPENDIX I

Recursion (14) can be written as

$$\tilde{\alpha}_j = \tilde{\alpha}_{j-1} - \frac{\mu(j)G(x(j),\theta_j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}e_a(j). \tag{22}$$

Squaring both sides, we get

$$\|\tilde{\alpha}_j\|^2 = \|\tilde{\alpha}_{j-1}\|^2 - 2\frac{\mu(j)G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2} + \left[\frac{\mu(j)\|G(x(j),\theta_j)\|e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\right]^2.$$

Taking expectations

$$E\|\tilde{\alpha}_j\|^2 = E\|\tilde{\alpha}_{j-1}\|^2 + E\left[\frac{\mu(j)\|G(x(j),\theta_j)\|e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\right]^2 - 2E\left[\frac{\mu(j)G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\right].$$

If we define

$$\Delta(\mu(j)) = 2E\left[\frac{\mu(j)G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\right] - E\left[\frac{\mu(j)\|G(x(j),\theta_j)\|e_a(j)}{1+\mu(j)\|G(x(j),\theta_j)\|^2}\right]^2$$

then

$$E\|\tilde{\alpha}_j\|^2 = E\|\tilde{\alpha}_{j-1}\|^2 - \Delta(\mu(j)).$$

In [65], an optimal value of step-size has been derived by maximizing $\Delta(\mu(j))$ at every time index j, since this will guarantee that the expected value of consequents-error norm will undergo the largest decrease from iteration $(j-1)$ to j. That is, to solve

$$\mu^*(j) = \arg\max_{\mu(j)>0} \Delta(\mu(j)). \tag{23}$$

To solve the maximization problem (23), let us define $\mu(j)$ using a constant $\overline{\mu}$ (where $0 < \overline{\mu} < 1$) such that

$$\mu(j) = \frac{\overline{\mu}}{\|G(x(j),\theta_j)\|^2(1-\overline{\mu})}, 0 < \overline{\mu} < 1$$

and

$$\Delta(\overline{\mu}) = 2\overline{\mu}E\left[\frac{G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}e_a(j)}{\|G(x(j),\theta_j)\|^2}\right] - \overline{\mu}^2E\left[\frac{|e_a(j)|^2}{\|G(x(j),\theta_j)\|^2}\right].$$

Instead of solving (23) directly, we reformulate our problem as

$$\mu^*(j) = \frac{\overline{\mu}^*}{\|G(x(j),\theta_j)\|^2(1-\overline{\mu}^*)}, \overline{\mu}^* = \arg\max_{0<\overline{\mu}<1} \Delta(\overline{\mu}).$$

Substituting $e_a(j) = G^T(x(j),\theta_j)\tilde{\alpha}_{j-1} + n_j, \Delta(\overline{\mu})$ becomes

$$\Delta(\overline{\mu}) = 2\overline{\mu}E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2 + G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}n_j}{\|G(x(j),\theta_j)\|^2}\right] - \overline{\mu}^2E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}+n_j|^2}{\|G(x(j),\theta_j)\|^2}\right].$$

Assuming that zero-mean noise sequence $\{n_j\}$ is independent, identically distributed and statistically independent of regres-

sion sequence $\{G(x(j),\theta_j)\}$, $\Delta(\overline{\mu})$ can be approximated as

$$\Delta(\overline{\mu}) \approx 2\overline{\mu}E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2}{||G(x(j),\theta_j)||^2}\right] - \overline{\mu}^2 E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2}{||G(x(j),\theta_j)||^2}\right] - \overline{\mu}^2\sigma_n^2 E\left[\frac{1}{||G(x(j),\theta_j)||^2}\right] \quad (24)$$

where $\sigma_n^2 = E|n_j|^2$. Maximizing (24) leads to

$$\overline{\mu}^* = \frac{E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2}{||G(x(j),\theta_j)||^2}\right]}{E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2}{||G(x(j),\theta_j)||^2}\right] + \sigma_n^2 E\left[\frac{1}{||G(x(j),\theta_j)||^2}\right]}$$

$$\mu^*(j) = \frac{E\left[\frac{|G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}|^2}{||G(x(j),\theta_j)||^2}\right]}{\sigma_n^2||G(x(j),\theta_j)||^2 E\left[\frac{1}{||G(x(j),\theta_j)||^2}\right]}.$$

If we define a vector $p_j = (G^T(x(j),\theta_j)\tilde{\alpha}_{j-1}G(x(j),\theta_j))/(||G(x(j),\theta_j)||^2)$, then

$$\mu^*(j) = \frac{E||p_j||^2}{\sigma_n^2||G(x(j),\theta_j)||^2 E\left[\frac{1}{||G(x(j),\theta_j)||^2}\right]}. \quad (25)$$

Thus, to compute the optimal learning rate $\mu^*(j)$, we first need to at least compute the term $E||p_j||^2$. To do this, we follow the approach of [65] to estimate p_j as

$$\hat{p}_j = \omega\hat{p}_{j-1} + (1-\omega)\frac{e_a(j)}{||G(x(j),\theta_j)||^2}G(x(j),\theta_j)$$

for a smoothing factor ω $(0 < \omega < 1)$, being motivated by the fact that

$$E[p_j] = E\left[\frac{e_a(j)}{||G(x(j),\theta_j)||^2}G(x(j),\theta_j)\right].$$

Thus, we propose to estimate $\mu^*(j)$ as

$$\mu^*(j) = \frac{||\hat{p}_j||^2}{C||G(x(j),\theta_j)||^2} \quad (26)$$

where C is a constant. Now, the estimation strategy in (13) and (14) for optimal learning rate $\mu^*(j)$ becomes (16)–(19).

APPENDIX II
A GAUSS–NEWTON BASED ALGORITHM

Given input–output data pairs $\{x(j),y(j)\}_{j=0}^N$, to compute the parameters

$$\theta_j = \arg\min_\theta \psi_j(\theta) = ||r(\theta)||^2$$

$$r(\theta) = \begin{bmatrix} \left[\frac{[y(j)-G^T(x(j),\theta)\alpha_{j-1}]^2}{1+\mu||G(x(j),\theta)||^2}\right]^{+1/2} \\ (\mu_\theta^{-1})^{1/2}(\theta-\theta_{j-1}) \end{bmatrix}$$

$$\alpha_j = \alpha_{j-1} + \frac{\mu G(x(j),\theta_j)[y(j)-G^T(x(j),\theta_j)\alpha_{j-1}]}{1+\mu||G(x(j),\theta_j)||^2}$$

$j = 0, 1, \ldots$, we suggest a Gauss–Newton based algorithm. The algorithm consists of the following steps.

1) Choose initial guess about cluster centers θ_{-1}, number of maximum epochs $E_{\max}, \alpha_{-1} = 0$, epoch count EC $= 0$, and data index $j = 0$.
2) If EC $< E_{\max}$
   a) if $j \le N$
      i) define

$$r(\theta) = \begin{bmatrix} \left[\frac{[y(j)-G^T(x(j),\theta)\alpha_{j-1}]^2}{1+\mu||G(x(j),\theta)||^2}\right]^{1/2} \\ (\mu_\theta^{-1})^{1/2}(\theta-\theta_{j-1}) \end{bmatrix}$$

      and let $s^*(\theta)$ be the unique solution of the following linear least squares problem:

$$s^*(\theta) = \arg\min_s[||r(\theta) + r'(\theta)s||^2]$$

      where $r'(\theta)$ is the Jacobian matrix of vector r with respect to θ, determined by the method of finite differences. The Jacobian $r'(\theta)$ is a full rank matrix, as a result of using regularization.
      ii) compute $\theta_j = \theta_{j-1} + s^*(\theta_{j-1})$.
      iii) compute

$$\alpha_j = \alpha_{j-1} + \frac{\mu G(x(j),\theta_j)[y(j)-G^T(x(j),\theta_j)\alpha_{j-1}]}{1+\mu||G(x(j),\theta_j)||^2}.$$

      iv) $j := j+1$ and go to step 2a).
   b) EC := EC $+ 1, \alpha_{-1} := \alpha_N, \theta_{-1} := \theta_N, j = 0$, and go to step 2.

The major computational complexity of the above algorithm lies in solving the linear least squares problem $\min_s[||r(\theta) + r'(\theta)s||^2]$. A Matlab code was developed to implement the algorithm, where the linear least squares problem was solved by computing the *Moore–Penrose pseudoinverse*. The average time in seconds required for one pass of the algorithm (i.e., step 2a) on a Pentium 2.53 GHz computer was approximately equal to (number of clusters)/(4000).

REFERENCES

[1] W. W. Wierwille and F. T. Eggemeier, "Recomendations for mental workload measurement in a test and evaluation environment," *Human Factors*, vol. 35, no. 2, pp. 263–281, 1993.

[2] A. F. Kramer, "Physiological metrics of mental workload: A review of recent progress," in *Multiple-Task Performance*, D. L. Damos, Ed. London, U.K.: Taylor and Francis, 1991, pp. 279–328.

[3] J. Aasman, G. Mulder, and L. J. M. Mulder, "Operator effort and the measurement of heart-rate variability," *Human Factors*, vol. 29, pp. 161–170, 1987.

[4] L. J. M. Mulder and G. Mulder, "Cardiovavscular reactivity and mental workload," in *The Beat-by-Beat Investigation of Cardiovascular Function*, O. Rompelman and R. I. Kitney, Eds. Oxford, U.K.: Oxford Univ. Press, 1987, pp. 216–253.

[5] E. J. Sirevaag, A. F. Kramer, C. D. Wickens, M. Reisweber, D. L. Strayer, and J. F. Grenell, "Assessment of pilot performance and mental workload in rotary wing aircraft," *Ergonomics*, vol. 36, pp. 1121–1140, 1993.

[6] A. J. Tattersall and G. R. Hockey, "Level of operator control and changes in heart rate variability during simulated flight maintenance," *Human Factors*, vol. 37, no. 4, pp. 682–698, 1995.

[7] G. F. Wilson, "Air-to-ground training missions: A psychophysiological workload analysis," *Ergonomics*, vol. 36, pp. 1071–1087, 1993.

[8] G. F. Wilson and F. T. Eggemeier, "Psychophysiological assessment of workload in multitask environments," in *Multiple-Task Performance*, D. L. Damos, Ed. London, U.K.: Taylor and Francis, 1991, pp. 329–360.

[9] A. Lindqvist, E. Keskinen, K. Antila, L. Halkola, T. Peltonen, and I. Valimaki, "Heart rate variability, cardiac mechanics, and subjectively evaluated stress during simulator flight," *Aviation Space Environ. Med.*, vol. 54, pp. 685–690, 1983.

[10] T. Kamada, S. Miyake, M. Kumashiro, H. Monou, and K. Inoue, "Power spectral analysis of heart rate variability in type As and type Bs during mental workload," *Psychosom. Med.*, vol. 54, pp. 462–470, 1992.

[11] M. Kollai and B. Kollai, "Cardiac vagal tone in generalised anxiety disorder," *Br. J. Psychiatr.*, vol. 161, pp. 831–835, 1992.

[12] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart rate variability. standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.

[13] W. A. Tiller, R. McCraty, and M. Atkinson, "Cardiac coherence: A new, noninvasive measure of autonomic nervous system order," *Alt. Therapies*, vol. 2, no. 1, pp. 52–65, 1996.

[14] J. K. Lenneman and R. W. Backs, "The validity of factor analytically derived cardiac autonomic components for mental workload assessment," in *Engineering Psychophysiology: Issues and Applications*, R. Backs and W. Boucsein, Eds. Mahwah, NJ: Lawrence Erlbaum, 2000, pp. 161–175.

[15] B. M. Sayers, "Analysis of heart rate variability," *Ergonomics*, vol. 16, no. 1, pp. 17–32, 1973.

[16] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Barger, and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat to beat cardiovascular control," *Science*, vol. 213, pp. 220–222, 1981.

[17] C. Yang and T. Kuo, "Assessment of cardiac sympathetic regulation by respiratory-related arterial pressure variability in the rat," *J. Physiol.*, vol. 515, pp. 887–896, 1999.

[18] A. M. Bianchi, L. T. Mainardi, C. Merloni, S. Chierchia, and S. Cerutti, "Continuous monitoring of the sympatho-vagal balance through spectral analysis," *IEEE Eng. Med. Bio.*, vol. 16, pp. 64–73, 1997.

[19] J. Morlet, G. Arens, I. Fourgeau, and D. Giard, "Wave propagation and sampling theory," *Geophysics*, vol. 47, pp. 203–236, 1982.

[20] O. Fukuda, Y. Nagata, K. Homma, and T. Tsuji, "Evaluation of heart rate variability by using wavelet transformation and a recurrent neural network," in *23rd Annu. Int. Conf. IEEE Eng. Med. Biol.*, Istanbul, Turkey, 2001.

[21] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 28–44, Jan. 1973.

[22] R. Babuška, *Fuzzy Modeling for Control*. Boston, MA: Kluwer Academic, 1998.

[23] M. Kumar, R. Stoll, and N. Stoll, "Robust solution to fuzzy identification problem with uncertain data by regularization—Fuzzy approximation to physical fitness with real world medical data: An application," *Fuzzy Opt. Decision Making*, vol. 3, no. 1, pp. 63–82, Mar. 2004.

[24] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 6, pp. 1414–1427, 1992.

[25] M. Kumar, R. Stoll, and N. Stoll, "Robust adaptive fuzzy identification of time-varying processes with uncertain data: Handling uncertainties in the physical fitness fuzzy approximation with real world medical data: An application," *Fuzzy Opt. Decision Making*, vol. 2, pp. 243–259, Sep.. 2003.

[26] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference systems," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, pp. 665–685, May 1993.

[27] M. Kumar, R. Stoll, and N. Stoll, "Regularized adaptation of fuzzy inference systems. Modelling the opinion of a medical expert about physical fitness: An application," *Fuzzy Opt. Decision Making*, vol. 2, Dec. 2003.

[28] D. Nauck and R. Kruse, "A neuro-fuzzy approach to obtain interpretable fuzzy systems for function approximation," in *Proc. IEEE Int. Conf. Fuzzy Syst. 1998 (FUZZ-IEEE'98)*, Anchorage, AK, May 1998, pp. 1106–1111.

[29] F. Herrera, M. Lozano, and J. Verdegay, "Generating fuzzy rules from examples using genetic algorithms," in *Proc. 5th Int. Conf. Inf. Process. Manag. Uncertainty Knowledge-Based Syst. (IPMU'94)*, Paris, France, Jul. 1994, pp. 675–680.

[30] P. Rani, J. Sims, R. Brackin, and N. Sarkar, "Online stress detection using psychophysiological signal for implicit human-robot cooperation," *Robotica*, vol. 20, no. 6, pp. 673–686, 2002.

[31] J. S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Upper Saddle River, NJ: Prentice-Hall, 1997.

[32] K. Nozaki, H. Ishibuchi, and H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets Syst.*, vol. 86, pp. 251–270, 1997.

[33] P. Thrift, "Fuzzy logic synthesis with genetic algorithms," in *Proc. 4th Int. Conf. Genetic Algorithms*, 1991, pp. 509–513.

[34] J. Liska and S. S. Melsheimer, "Complete design of fuzzy logic systems using genetic algorithms," in *Proc. 3rd IEEE Int. Conf. Fuzzy Syst.*, 1994, pp. 1377–1382.

[35] J. Abonyi, R. Babuska, and F. Szeifert, "Modified Gath-Geva fuzzy clustering for identification of takagi-sugeno fuzzy models," *IEEE Trans. Syst., Man, Cybern.*, pt. B, pp. 612–621, Oct. 2002.

[36] J. J. Shan and H. C. Fu, "A fuzzy neural network for rule acquiring on fuzzy control systems," *Fuzzy Sets Syst.*, vol. 71, pp. 345–357, 1995.

[37] A. Gonzalez and R. Pérez, "Completeness and consistency conditions for learning fuzzy rules," *Fuzzy Sets Syst.*, vol. 96, pp. 37–51, 1998.

[38] D. Simon, "Design and rule base reduction of a fuzzy filter for the estimation of motor currents," *Int. J. Approx. Reason.*, vol. 25, pp. 145–167, Oct. 2000.

[39] ——, "Training fuzzy systems with the extended Kalman filter," *Fuzzy Sets Syst.*, vol. 132, pp. 189–199, Dec. 2002.

[40] S. Wu and M. J. Er, "Dynamic fuzzy neural networks—Anovel approach to function approximation," *IEEE Trans. Syst., Man., Cybern. B*, vol. 30, pp. 358–364, 2000.

[41] M. Burger, H. Engl, J. Haslinger, and U. Bodenhofer, "Regularized data-driven construction of fuzzy controllers," *J. Inverse Ill-Posed Problems*, vol. 10, pp. 319–344, 2002.

[42] D. S. Chen and R. C. Jain, "A robust back propagation learning algorithm for function approximation," *IEEE Trans. Neural Netw.*, vol. 5, pp. 467–479, May 1994.

[43] W. Y. Wang, T. T. Lee, C. L. Liu, and C. H. Wang, "Function approximation using fuzzy neural networks with robust learning algorithm," *IEEE Trans. Syst., Man., Cybern. B*, vol. 27, pp. 740–747, Sep.. 1997.

[44] W. Yu and X. Li, "Fuzzy identification using fuzzy neural networks with stable learning algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 3, pp. 411–20, Jun. 2004.

[45] T. Johansen, "Robust identification of Takagi-Sugeno-Kang fuzzy models using regularization," in *Proc. IEEE Conf. Fuzzy Syst.*, New Orleans, LA, 1996, pp. 180–186.

[46] X. Hong, C. J. Harris, and S. Chen, "Robust neurofuzzy rule base knowledge extraction and estimation using subspace decomposition combined with regularization and d-optimality," *IEEE Trans. Syst., Man., Cybern. B*, vol. 34, no. 1, pp. 598–608, 2004.

[47] M. Kumar, R. Stoll, and N. Stoll, "SDP and SOCP for outer and robust fuzzy approximation," in *Proc. 7th IASTED Int. Conf. Artificial Int. Soft Comp.*, Banff, Canada, Jul. 2003.

[48] ——, "A robust design criterion for interpretable fuzzy models with uncertain data," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 314–328, Apr. 2006.

[49] ——, "Robust adaptive identification of fuzzy systems with uncertain data," *Fuzzy Opt. Decision Making*, vol. 3, no. 3, pp. 195–216, Sep.. 2004.

[50] M. Kumar, N. Stoll, and R. Stoll, "An energy-gain bounding approach to robust fuzzy identification," *Automatica*, vol. 42, no. 5, pp. 711–721, May 2006.

[51] M. Kumar, R. Stoll, and N. Stoll, "A min-max approach to fuzzy clustering, estimation, and identification," *IEEE Trans. Fuzzy Syst.*, vol. 14, pp. 248–262, Apr. 2006.

[52] M. B. Lotrič, A. Štefanovska, D. Stajer, and V. Urbančič-Rovan, "Spectral components of heart rate variability determined by wavelet analysis," *Physiol. Meas.*, vol. 21, pp. 441–457, 2000.

[53] S. G. Hart and L. E. Staveland, "Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam, The Netherlands: Elsevier, 1988.

[54] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

[55] R. Krishnapuram and J. M. Keller, "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 98–110, May 1993.

[56] R. Krishnapuram and J. M. Keller, "The possibilistic c-means algorithm: Insights and recommendations," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 385–393, Aug. 1996.

[57] R. N. Davé, "Characterization and detection of noise in clustering," *Pattern Recognit. Lett.*, vol. 12, no. 11, pp. 657–664, 1991.

[58] J. S. Zhang and Y. W. Leung, "Improved possibilistic c-means clustering algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 12, pp. 209–217, Apr. 2004.
[59] N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 13, pp. 517–530, Aug. 2005.
[60] M. S. Houle and G. E. Billman, "Low-frequency component of heart rate variability spectrum: A poor marker of sympathetic activity," *Amer. J. Physiol.*, vol. 276, pp. 215–223, 1999.
[61] Y. Zhong, H. Wang, K. H. Ju, K.-M. Jan, and K. H. Chon, "Nonlinear analysis of the separate contributions of autonomic nervous systems to heart rate variability using principal dynamic modes," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 255–262, Feb. 2004.
[62] Z. P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.
[63] M. P. Clamann, M. C. Wright, and D. B. Kaber, "Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing," in *Proc. 46th Annu. Meeting Human Factors Ergon. Soc.*, Santa Monica, CA, 2002, pp. 342–346.
[64] R. McCraty, M. Atkinson, W. Tiller, G. Rein, and A. D. Watkins, "The effects of emotions on short-term power spectrum analysis of heart rate variability," *Amer. J. Cardiol.*, vol. 76, no. 14, pp. 1089–1093, 1995.
[65] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size nlms and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, pp. 132–135, Feb. 2004.

Mohit Kumar received the B.Tech. degree in electrical engineering from the National Institute of Technology, Hamirpur, India, in 1999, the M.Tech. degree in control engineering from the Indian Institute of Technology, Delhi, in 2001, and the Ph.D. degree *(summa cum laude)* in electrical engineering from Rostock University, Germany, in 2004.

He was a Research Scientist with the Institute of Occupational and Social Medicine, Rostock, from 2001 to 2004. Currently, he is with the Center for Life Science Automation, Rostock. His research interests include robust adaptive fuzzy identification, fuzzy logic in medicine, and robust adaptive control.

Matthias Weippert received the first state examination for lectureship in biology and physical education from Rostock University, Germany, in 2003.

He was a Research Scientist with the Institute of Occupational and Social Medicine, Rostock, during 2004–2005. Since 2006, he has been with the Institute of Preventive Medicine, Rostock. His research interests include performance diagnostics and stress-strain-reactions in occupational and sports medicine.

Reinhard Vilbrandt received the university entrance diploma from CJD Rostock, Germany, in 1997. He received the intermediate diploma and the Dipl. Ing. degree from Rostock University, Germany, in 2000 and 2004, respectively, both in automation.

He is a Scientific Assistant with the Institute of Preventive Medicine, Rostock University. He was with the civil service during 1997–1998 at a residential home for the elderly. His thesis topic was "Mobile System for Flexible Recordation of Physical and Physco-Mental Load." He is responsible for the integration and development of mobile measurement into a central database.

Steffi Kreuzfeld received the Dr.Med. degree from Rostock University, Germany, in 2003.

She has been with the Institute of Occupational and Social Medicine since 2004.

Regina Stoll received the Dipl.-Med. degree, the Dr.Med. degree in occupational medicine, and the Dr.Med.Habil degree in occupational and sports medicine from Rostock University, Germany, in 1980, 1984, and 2002, respectively.

She is Head of the Institute of Preventive Medicine, Rostock. She is a Member of the Faculty of Medicine and a Faculty Associate with the College of Computer Science and Electrical Engineering, Rostock University. She is a member of the Adjunct Faculty in the Industrial Engineering Department, North Carolina State University. Her research interests include occupational physiology, preventive medicine, and cardiopulmonary diagnostics.