Fuzzy Optimization and Decision Making, 3, 243–259, 2003

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

MOHIT KUMAR mohit.kumar@etechnik.uni-rostock.de
Institute of Occupational and Social Medicine, Faculty of Medicine, University of Rostock, D-18055 Rostock, Germany

REGINA STOLL regina.stoll@med.uni-rostock.de
Institute of Occupational and Social Medicine, Faculty of Medicine, University of Rostock, D-18055 Rostock, Germany

NORBERT STOLL norbert.stoll@etechnik.uni-rostock.de
Institute of Automation, Department of Electrical Engineering and Information Technology, University of Rostock, D-18119 Rostock-Warnemnde, Germany

Abstract. This study considers the problem of Robust Fuzzy approximation of a time-varying nonlinear process in the presence of uncertainties in the identification data using a Sugeno Fuzzy System while maintaining the interpretability of the fuzzy model during identification. A recursive procedure for the estimation of fuzzy parameters is proposed based on solving local optimization problem that attempt to minimize the worst-case effect of data uncertainties on approximation performance. To illustrate the approach, several simulation studies on numerical examples are provided. The developed scheme was applied to handle the vagueness, ambiguity and uncertainty inherently present in the general notion of a Medical Expert about Physical Fitness based on a set of various Physiological parameters measurements.

Keywords: fuzzy-modelling, nonlinear constrained optimization, uncertainty, robustness, regularization

1. Introduction

For the approximation of ill-defined and complex processes, fuzzy systems have been considered as appropriate tools because of their capability of handling uncertainty based on set of fuzzy if-then rules derived from expert domain knowledge (Zadeh (1973)). However, in some situations this expert domain knowledge may not be sufficient to design fuzzy model due to lack of knowledge or problems due to different biases of human experts, so there is a need of learning of fuzzy inference system from data (Jang (1993), Nauck and Kruse (1997), Nauck and Kruse (1998)). Hence we need to develop some methods of identiffcation of parameters of the fuzzy system which approximate the complex process to a desired level of accuracy. As the data used in identiffcation contains generally vagueness and ambiguity, so there is need of developing robust algorithms for learning from imprecise data (Burger et al (2002)). Since, the process may be time varying, so the learning mechanism should be adaptive enough to cope up with the time varying characteristics of the process. Automatic construction or tuning of these fuzzy systems from example data has been widely explored for Sugeno type Fuzzy Inference System (Babuska (2000), Bodenhofer and Bauer (2000), Espinosa and Vandewalle (2000), Setnes et al (1998)), since they are supposed to ideally combine simplicity with good analytical properties (Takagi and Sugeno (1985)). In this study, we consider the subject of deterministic robustness of adaptive fuzzy identification problem in which the perturbations in the data are deterministic and bounded. The objective is to alleviate the worst-case effect of the uncertainties on estimation performance.

The proposed algorithm was applied to model or approximate the physical fitness with an interpretable fuzzy expert system tuned by the experience (comment) of a Medical Expert and various physiological parameters measurements. This Fuzzy Expert System is intended to be capable of not only approximating the functional relationship of physical fitness with various physiological parameters but also to be capable of handling uncertainties in the opinion of Medical Experts (due to different experiences and biases of their mind) and handling different variations of this functional relationship with different patients. Hence the identification of such a expert system requires the use of some robust and adaptive identification (with respect to different patients) in a very natural way.

2. Problem Formulation

Let us consider the problem of tuning a Sugeno controller ($Fs : Xs \rightarrow Y$), mapping n-dimensional input space ($X = X_1 \times X_2 \times .. \times X_n$) to one dimensional real line, consisting of K different rules. The ith rule is in the the form:
If X_1 *is* A_{i1} *and* X_2 is A_{i2} ... *and* X_n *is* A_{in} *then* $y = \alpha_i$; for all $i = 1, 2, .., K$, Where A_{i1}, $A_{i2}, .., A_{in}$ are non-empty fuzzy subsets of $X_1 \times X_2 \times .. \times X_n$ respectively such that membership function $\mu_{A_{ij}} : X_j \rightarrow [0, 1]$ fulfill $\sum_{i=1}^{K} \Pi_{j=1}^{n} \mu_{A_{ij}}(x_j) > 0$ for all $x_j \in X_j$. The values i real numbers. So we have

$$F_s(x_1, x_2, .., x_n) = \frac{\sum_{i=1}^{K} \alpha_i \prod_{j=1}^{n} \mu_{A_{ij}}(x_j)}{\sum_{i=1}^{K} \prod_{j=1}^{n} \mu_{A_{ij}}(x_j)} \tag{1}$$

where $x_j \in [a_j, b_j]$ for all $j = 1, .., n$.

For the common mathematical formulation of various classes of membership functions, we consider a knot sequence (θ), such that shape of membership function depends on the elements of sequence θ which partition the universe of each input variable ($x_i \in [a_i, b_i]$) in to P_i linguistic terms. To elaborate the construction of membership functions based on knot sequence (θ), two examples one for trapezoidal and other for Gaussian membership functions are provided below.

Trapezoidal membership curves let $\theta = (t_1^1, .., t_1^{2P1-2}, t_2^1, .., t_2^{2P2-2}, .., t_n^1, .., t_n^{2Pn-2}) \in \mathbb{R}^L$, such that for i^{th} input, $a_i \leq t_i^1 \leq ..... \leq t_i^{2Pi-2} \leq b_i$ holds for all $i = 1, .., n$. So P_i

membership curves for i^t input ($A_{1i}, A_{2i}, \ldots, A_{P_i i}$) can be defined as:

$$A_{1i}(x_i, \theta) = \begin{cases} 1 & \text{if } x_i \in [a_i, t_i^1] \\ \frac{-x_i + t_i^2}{t_i^2 - t_i^1} & \text{if } x_i \in [t_i^1, t_i^2] \\ 0 & \text{otherwise} \end{cases}$$

$$A_{ji}(x_i, \theta) = \begin{cases} \frac{x_i - t_i^{2j-3}}{t_i^{2j-2} - t_i^{2j-3}} & \text{if } x_i \in [t_i^{2j-3}, t_i^{2j-2}] \\ 1 & \text{if } x_i \in [t_i^{2j-2}, t_i^{2j-1}] \\ \frac{-x_i + t_i^{2j}}{t_1^{2j} - t_i^{2j-1}} & \text{if } x_i \in [t_i^{2j-1}, t_i^{2j}] \\ 0 & \text{otherwise} \end{cases}$$

$$A_{P_i i}(x_i, \theta) = \begin{cases} \frac{x_i - t_i^{2P_i-3}}{t_i^{2P_i-2} - t_i^{2P_i-3}} & \text{if } x_i \in [t_i^{2P_i-3}, t_i^{2P_i-2}] \\ 1 & \text{if } x_i \in [t_i^{2P_i-2}, b_i] \\ 0 & \text{otherwise} \end{cases}$$

The Figure 1 shows an example with the choice of antecedent parameters as: $a_i = 0$, $t_i^1 = 1$, $t_i^2 = 2$, $t_i^3 = 3$, $t_i^4 = 4$ and $b_i = 5$.

Gaussian membership curves let $\theta = (t_1^1, .., t_1^{P1-2},\ t_2^1, .., t_2^{P2-2}, ..,\ t_n^1, .., t_n^{Pn-2}) \in \mathbb{R}^L$, such that for $i^t h$ input, $a_i \leq t_i^1 \leq \ldots.. \leq t_i^{Pi-2} \leq b_i$ holds for all $i = 1, .., n$. So P_i membership curves for $i^t h$ input ($A_{1i}, A_{2i}, \ldots, A_{P_i i}$) can be defined as:

$$A_{1i}(x_i) = e^{-(x_i - a_i)^2}$$
$$A_{ji}(x_i, \theta) = e^{-(x_i - t_i^j)^2}$$
$$A_{P_i i}(x_i) = e^{-(x_i - b_i)^2}$$

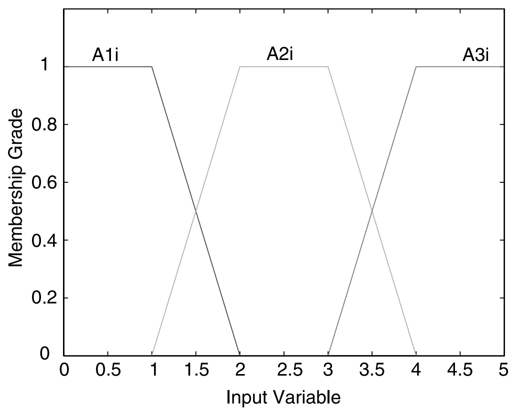

Figure 1. Tapezoidal membership curves.

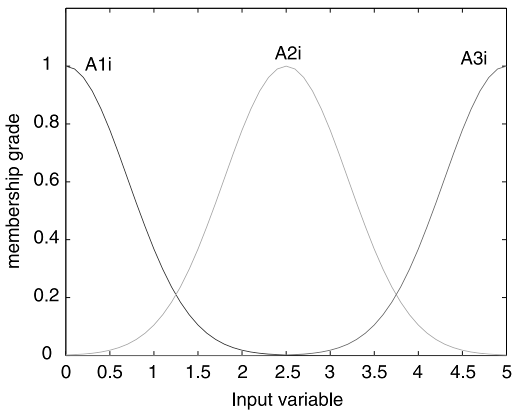

Figure 2. Gaussian membership curves.

The Figure 2 shows an example with the choice of antecedent parameters as: $a_i = 0$, $t_i^1 = 2.5$ and $b_i = 5$.

Total number of possible K rules depends on the number of membership curves for each input i.e. $K = \Pi_{i=1}^{n} P_i$, where P_i is the number of membership curves defined over i^{th} input. Depending upon the choice of membership curves equation (1) can be rewritten as function of θ.

$$F_s(x_1, x_2, .., x_n) = \sum_{j=1}^{K} \alpha_j B_j(x_i, x_2, .., x_n; \theta)$$

Lets introduce the following notation i.e.
Let $\alpha = [\alpha_j]_{j=1,2,\ldots,K}$; $x = (x_1, x_2, \ldots, x_n)$; $B = [B_j(x;\theta)]_{j=1,\ 2,\ldots,K}$.
Now above equation can be written as

$$F_s(x) = B^T(x, \theta)\alpha \tag{2}$$

Our objective is the adaptive identification of parameters α and θ from noisy measurements $[x(k), y(k)]$.

We suppose that (α_k^*, θ_k^*) is the true value of unknown fuzzy parameters at k^{th} instant of time, which approximate the time-varying process at that instant of time, then we have

$$y(k) = (B(x(k), \theta_k^*) + \Delta B)^T \alpha_k^* \tag{3}$$

where ΔB denotes the uncertainty in regression vector B $(x(k), \theta_k^*)$.

For the given uncertain model (3), we attempt to estimate the parameters (α_k^*, θ_k^*) in some robust manner. We assume that we already have estimate at time k, say (α_k, θ_k), and want to estimate $(\alpha_{k+1}, \theta_{k+1})$ for time $k + 1$ by solving an optimization (min-max) problem:

$$\min_{\{\alpha_{k+1}, \theta_{k+1}\}} \max_{\{\Delta B\}} J(\alpha_{k+1}, \theta_{k+1})$$

The cost function $J(\alpha_{k+1}, \theta_{k+1})$ is defined by

$$\begin{aligned} J(\alpha_{k+1}, \theta_{k+1}) = {} & |y(k) - (B(x(k), \theta_{k+1}) + \Delta B)^T \alpha_{k+1}|^2 \\ & + \beta_1 ||\alpha_{k+1} - \alpha_k||^2 + \beta_2 ||\theta_{k+1} - \theta_k||^2 \end{aligned}$$

Where β_1 and β_2 are the regularization parameters, which control the influence of the regularization terms. Define the quantities:

$$\alpha = \alpha_{k+1} - \alpha_k$$

$$e(\theta_{k+1}) = y(k) - (B(x(k), \theta_{k+1}))^T \alpha_k$$

$$u(k) = \Delta B^T (\alpha + \alpha_k)$$

We assume that uncertainty in data is bounded .i.e. $||\Delta B|| \leq \eta$ for some known scalar η. Let

$$\Phi(\alpha) = \eta ||\alpha + \alpha_\mathrm{k}||$$

Now the optimization problem can be formulated as

$$\min_{\{\alpha, \theta_{k+1}\}} \max_{|u(k)| \leq \Phi(\alpha)} J(\alpha, \theta_{k+1})$$

$$\begin{aligned} J(\alpha, \theta_{k+1}) = {} & |(B(x(k), \theta_{k+1}))^T \alpha - e(\theta_{k+1}) + u(k)|^2 + \beta_1 ||\alpha||^2 \\ & + \beta_2 ||\theta_{k+1} - \theta_k||^2 \end{aligned}$$

For any fixed values of parameters (α, θ_{k+1}), we define the following maximization problem:

$$C(\alpha, \theta_{k+1}) = \max_{|u(k)| \leq \Phi(\alpha)} |(B(x(k), \theta_{k+1}))^T \alpha - e(\theta_{k+1}) + u(k)|^2$$

$$C(\alpha, \theta_{k+1}) = \max_{|u(k)| \leq \Phi(\alpha)} R(\alpha, \theta_{k+1}, u(k))$$

For constant (α, θ_{k+1}), the cost function $R(\alpha, \theta_{k+1}, u(k))$ is convex in $u(k)$, so that the maximization over $u(k)$ is achieved at the boundary, $|u(k)| = \Phi(\alpha)$. Therefore the above constrained optimization problem can be reformulated by introducing a nonnegative Lagrange multiplier λ, as follows:

$$\max_{\{u(k),\lambda\}} \quad |(B(x(k),\theta_{k+1}))^T\alpha - e(\theta_{k+1}) + u(k)|^2 - \lambda(|u(k)|^2 - \Phi^2(\alpha))$$

Let λ^o and $u^o(k)$ be the optimal solution of above optimization problem obtained by differentiating the cost function with respect to u and λ which satisfy following equation:

$$(\lambda^o - 1)u^o(k) = (B(x(k),\theta_{k+1}))^T\alpha - e(\theta_{k+1})$$

Also $\lambda^o \geq 1$, for the hessian of cost function w.r.t $u(k)$ to be non positive-definite. So the $C(\alpha,\,\theta_{k+1})$ becomes as:

$$C(\alpha,\theta_{k+1}) = \left(1 + (\lambda^o - 1)^\dagger\right)\left((B(x(k),\theta_{k+1}))^T\alpha - e(\theta_{k+1})\right)^2 + \lambda^o\Phi^2(\alpha)$$

and

$$(\lambda^o - 1)^\dagger = \begin{cases} 0 & \text{if } \lambda^o = 1 \\ \frac{1}{\lambda^o - 1} & \text{otherwise} \end{cases}$$

Let us introduce a three-variable cost function $C(\alpha,\,\theta_{k+1},\,\lambda)$ as follows:

$$C(\alpha,\theta_{k+1},\lambda) = \left(1 + (\lambda - 1)^\dagger\right)\left((B(x(k),\theta_{k+1}))^T\alpha - e(\theta_{k+1})\right)^2 + \lambda\Phi^2(\alpha)$$

Hence, the original min-max problem is reduced to following min-min problem:

$$\min_{\{\alpha,\theta_{k+1}\}} \quad \min_{\lambda \geq 1} \quad J(\alpha,\theta_{k+1},\lambda)$$

$$J(\alpha,\theta_{k+1},\lambda) = C(\alpha,\theta_{k+1},\lambda) + \beta_1||\alpha||^2 + \beta_2||\theta_{k+1} - \theta_k||^2$$

3. Recursive Solution to Robust Adaptive Estimation Problem

To find minimum over α, gradient of $J(\alpha,\,\theta_{k+1},\,\lambda)$ with respect to α is set to zero, resulting in following equation

$$\left[\beta_1 I + \left(1 + (\lambda - 1)^\dagger\right)B(x(k),\theta_{k+1})B^T(x(k),\theta_{k+1})\right]\alpha + \frac{1}{2}\lambda\nabla\Phi^2(\alpha)$$
$$= \left(1 + (\lambda - 1)^\dagger\right)B(x(k),\theta_{k+1})e(\theta_{k+1})$$

Let α^o be the optimal solution obtained by solving above equation, which obviously is function of θ_{k+1} and λ, so we write

$$\alpha^o = \alpha^o(\theta_{k+1}, \lambda)$$
$$G(\theta_{k+1}, \lambda) = J(\alpha^o, \theta_{k+1}, \lambda)$$

Now the optimization problem is reduced to:

$$\min_{\theta_{k+1}} \min_{\lambda \geq 1} G(\theta_{k+1}, \lambda)$$

solving for α^o, we get

$$\alpha^o = \left[(\beta_1 + \lambda\eta^2)I + (1 + (\lambda - 1)^{\dagger})B(x(k), \theta_{k+1})B^T(x(k), \theta_{k+1})\right]^{-1} \left[(1 + (\lambda - 1)^{\dagger})B(x(k), \theta_{k+1})e(\theta_{k+1}) - \lambda\eta^2\alpha_k\right]$$

Applying matrix inversion lemma, we get

$$\alpha^o(\theta_{k+1}, \lambda) = -\gamma\alpha_k + \lambda\frac{\gamma B^T(x(k), \theta_{k+1})\alpha_k + e(\theta_{k+1})}{\beta(\lambda - 1) + \lambda||B(x(k), \theta_{k+1})||^2}B(x(k), \theta_{k+1})$$

where $\beta = \beta_1 + \lambda\eta^2$ and $\gamma = \eta^2\lambda/\beta$.

$$G(\theta_{k+1}, \lambda) = \frac{\lambda}{\beta}\left[\frac{(\beta_1 B^T(x(k), \theta_{k+1})\alpha_k - \beta y(k))^2}{\beta(\lambda - 1) + \lambda||B(x(k), \theta_{k+1})||^2} + \beta_1\eta^2||\alpha_k||^2\right] + \beta_2||\theta_{k+1} - \theta_k||^2$$

For a given θ_{k+1}, $G(\theta_{k+1}, \lambda)$ can be minimized by finding numerically the roots of gradient of G with respect to λ. However from the reference AI-Naffouri and Sayed (2000), it is known that minimum value of $G(\theta_{k+1}, \lambda)$ occurs at a value of λ very close to 1. So with this approximation of $\lambda^o = 1$, we get the optimal value of $G(\theta_{k+1}, \lambda)$ as

$$G^o(\theta_{k+1}) = \frac{1}{\beta_1 + \eta^2}\left[\frac{(-\beta_1 e(\theta_{k+1}) - \eta^2 y(k))^2}{\beta_1 + \eta^2 + ||B(x(k), \theta_{k+1})||^2} + \beta_1\eta^2||\alpha_k||^2\right] + \beta_2||\theta_{k+1} - \theta_k||^2$$

and

$$\alpha^o(\theta_{k+1}) = -\gamma\alpha_k + \frac{y(k) - (1 - \gamma)B^T(x(k), \theta_{k+1})\alpha_k}{\beta_1 + \eta^2 + ||B(x(k), \theta_{k+1})||^2}B(x(k), \theta_{k+1})$$

where $\gamma = \eta^2/(\beta_1 + \eta^2)$. Finally the Robust adaptive estimation problem is reduced to

$$\min_{\theta_{k+1}} G^o(\theta_{k+1})$$

Moreover, we also want to preserve the interpretability of fuzzy system during learning. So the membership curves can be prevented to overlap by imposing some constraints on the position of knots, for instance, in case of trapezoidal membership curves the constraints can be formulated i.e. for all $i = 1, .., n$

$$\begin{array}{lll} t_i^1 - a_i & \geq & \epsilon_i \\ t_i^{j+1} - t_i^j & \geq & \epsilon_i \quad for\ all\ j = 1, 2, .., (2P_i - 3) \\ b_i - t_i^{2P_i - 2} & \geq & \epsilon_i \end{array}$$

These constraints can be formulated in term of a matrix inequality $c.\theta \geq h$, similar to the constrained problem formulated in reference Burger et al (2002). Constrained optimization problem is

$$\begin{array}{ll} min & G^o(\theta_{k+1});\ c.\theta_{k+1} \geq h \\ \theta_{k+1} & \end{array}$$

let

$$\Psi(\theta_{k+1}) = \frac{1}{(\beta_1 + \eta^2)} \cdot \frac{(-\beta_1 e(\theta_{k+1}) - \eta^2 y(k))^2}{(\beta_1 + \eta^2 + ||B(x(k), \theta_{k+1})||^2)} + \beta_2 ||\theta_{k+1} - \theta_k||^2$$

$$f(\theta_{k+1}) = \left[\frac{1}{(\beta_1 + \eta^2).(\beta_1 + \eta^2 + ||B(x(k), \theta_{k+1})||^2)} \right]^{1/2}$$

and

$$r(\theta_{k+1}) = \begin{bmatrix} -\beta_1 f(\theta_{k+1}) e(\theta_{k+1}) - \eta^2 f(\theta_{k+1}) y(k) \\ \sqrt{\beta_2}(\theta_{k+1} - \theta_k) \end{bmatrix} \epsilon\ R^{L+1}$$

The optimization problem finally can be rewritten as

$$\begin{array}{ll} min & ||r(\theta_{k+1})||^2 \quad ;\ c.\theta_{k+1} \geq h \\ \theta_{k+1} & \end{array}$$

This minimization problem comes out to be a constrained nonlinear regularized least squares optimization problem, which could be solved using Gauss-Newton method.

Let $r'(\theta)$ is the Jacobian matrix of vector r with respect to θ, determined by the method of finite-differences which is a full rank matrix, as a result of regularization and $s*$ be the unique solution of following constrained optimization problem solved by the algorithm suggested by Lawson and Hanson (1995).

$$s*(\theta) = arg\ min_s[||r(\theta) + r'(\theta)s||^2 c.s \geq h - c.\theta]$$

So the antecedent parameters(θ_{k+1}) can be estimated based on the solution of above problem by following equation

$$\theta_{k+1} = \theta_k + s*(\theta_k) \tag{4}$$

Note that, θ_{k+1} is not the global minimizer of $||r(\theta_{k+1})||^2$, however estimation by equation (4) allows the fuzzy system as local expert instead of global expert. These local mappings facilitate the minimal disturbance principle (Widrow and Lehr (1990)), which is particular important in, on-line learning.

After the antecedent parameters have been adapted by equation (4), α_{k+1} can be estimated by following equation

$$\alpha_{k+1} = (1-\gamma)\alpha_k + \frac{y(k) - (1-\gamma)B^T(x(k), \theta_{k+1})\alpha_k}{\beta_1 + \eta^2 + ||B(x(k), \theta_{k+1}||^2} B(x(k), \theta_{k+1}) \tag{5}$$

where $\gamma = \eta^2/(\beta_1 + \eta^2)$. Thus we see that parameters of a time varying fuzzy system can be estimated using equations (4) and (5) from uncertain data, with a prior knowledge of uncertainty parameter η.

However, we are more interested in the fuzzy approximation of a uncertain time-varying process. In this case exact fuzzy model (number of membership curves for each input) of the process and bounding value of uncertainty parameter may or may not be known. In such situations we neglect the uncertainties in regression vector by substituting zero as the value of η. so we have following adaptation equations in this case:

$$f(\theta) = \left[\frac{1}{\beta_1.(\beta_1 + ||B(x(k), \theta)||^2)}\right]^{1/2} \tag{6}$$

$$r(\theta) = \begin{bmatrix} -\beta_1 f(\theta) e(\theta) \\ \sqrt{\beta_2}(\theta - \theta_k) \end{bmatrix} \epsilon\ R^{L+1} \tag{7}$$

$$s*(\theta) = arg\ min_s[||r(\theta) + r'(\theta)s||^2\ c.s \geq h - c.\theta] \tag{8}$$

$$\theta_{k+1} = \theta_k + s*(\theta_k) \tag{9}$$

$$\alpha_{k+1} = \alpha_k + \frac{y(k) - B^T(x(k), \theta_{k+1})\alpha_k}{\beta_1 + ||B(x(k), \theta_{k+1})||^2} B(x(k), \theta_{k+1}) \tag{10}$$

4. Simulation Studies

To show the feasibility of our approach we consider the problem of adaptive identification of a time varying process from noisy measurements using a fuzzy inference system. Let the function to be identified is described by following nonlinear equation

$$y = f(x,p) = \frac{-10x}{2p + x^2} + p^2 \tanh(x) \qquad x\ \epsilon\ [-0.5, 2.5]$$

where the parameter $p(t)$ is time varying. Figure 3 shows $y = f(x, p)$ the functional relationship, when parameter p varies from 1 to 2. Let us chose a Sugeno type of fuzzy inference system with gaussian membership curves for the approximation of this time

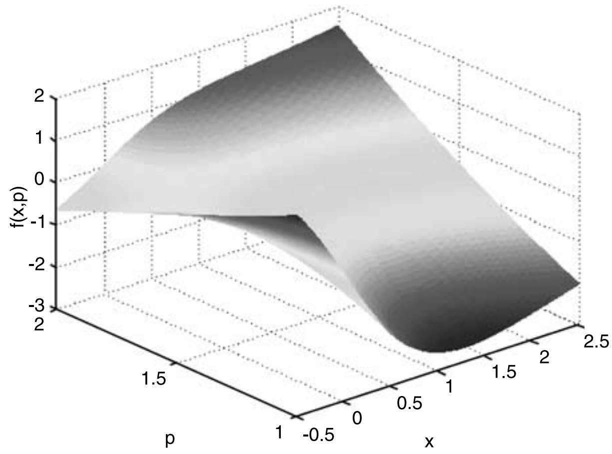

Figure 3. Unknown function to be identified.

varying function. Let us take 4 uniform distributed membership function over the range [−0.5, 2.5]. So initial guess about antecedent parameters is

$$\theta_0 = [-0.5, 0.5, 1.5, 2.5]$$

To preserve the interpretability of fuzzy system let us constrain our problem that minimum distance between two knots must be 0.1. The initial value α_o is taken equal to a zero vector.
We simulate the above function for $t = 0$ to $t = 15$ with $x(t)$ as

$$x(t) = -0.5 + |3 \sin 10t|$$

The sampling time $T = 0.01$ and a normally distributed random noise($n(t)$) with mean 0 and variance 0.01 is added to the output $y(t)$ shown in Figure 4. The regularization parameters $\beta_1 = 1$ and $\beta_2 = 10$ were taken for the simulations studies.

For the testing of the proposed scheme (i.e. equations (6−10)), following three cases of time varying parameter $p(t)$ were considered:

1. $p(t) = 1$.
2. $p(t) = \begin{cases} 1, & t \leq 7.5, \\ 2, & t > 7.5. \end{cases}$
3. $p(t) = 1 + \sin(0.1t)$.

For the first case when function to be identified is not time varying, Figure 5 shows the simulations results and Figure 6 shows the approximation error (difference between $y(t)$ and model output) surface over the time span of simulation.

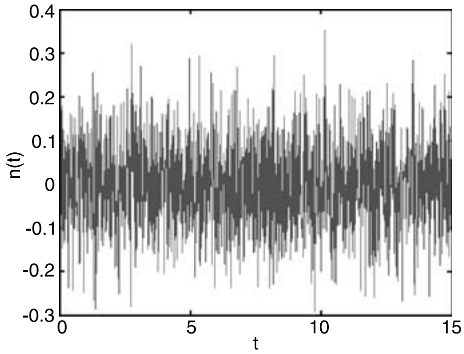

Figure 4. Random noise added to the measurements.

For the second case when function to be identified follows a step change at $t = 7.5$, Figure 7 shows the simulations results and Figure 8 shows the approximation error surface over the simulation time. We note that approximation error is maximum at $t = 7.5$, when $P(t)$ changes from 1 to 2. For the third case when function to be identified changes

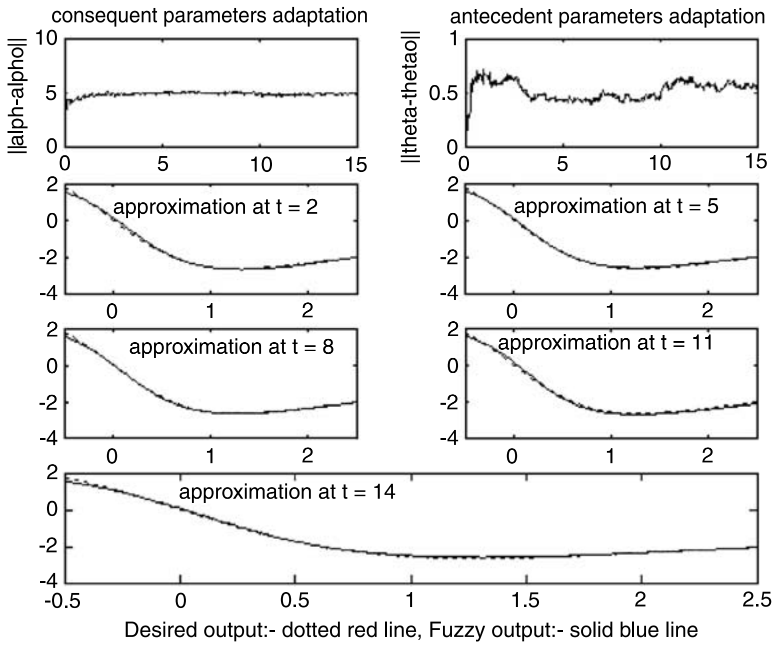

Figure 5. Robust estimation of nonlinear system with constant parameters.

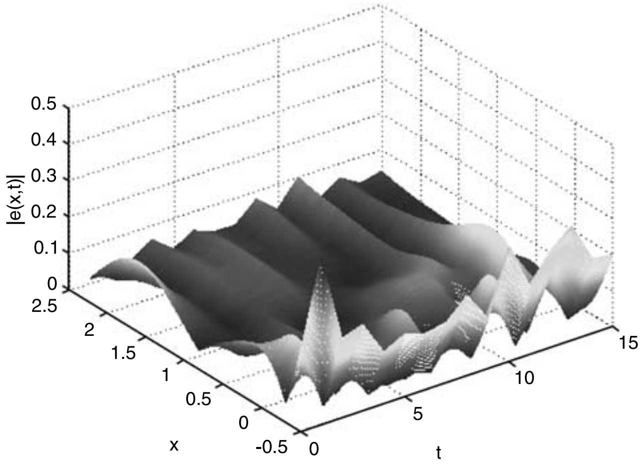

Figure 6. Robust estimation error surface with constant parameters.

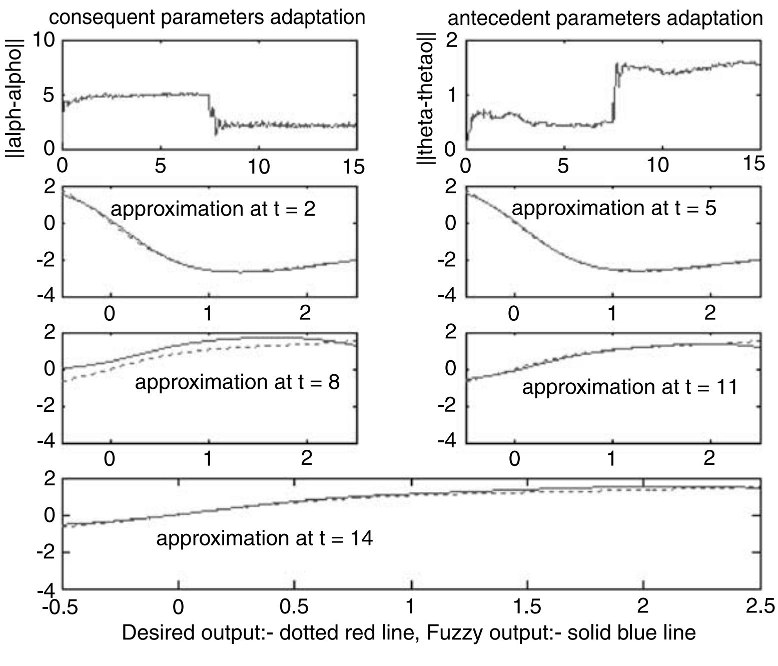

Figure 7. Robust estimation of nonlinear system with step changing parameters.

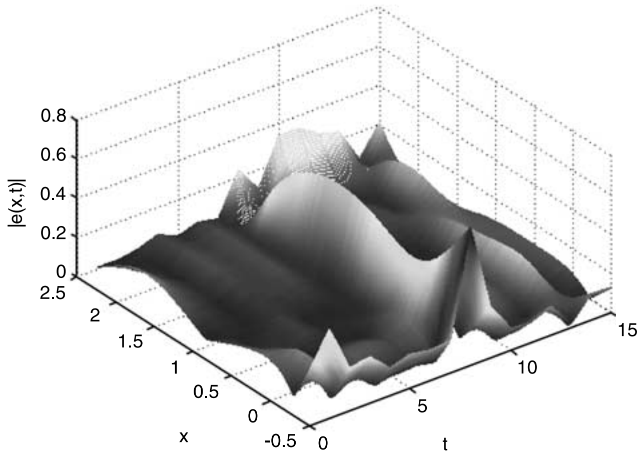

Figure 8. Robust estimation error surface for step changing parameters.

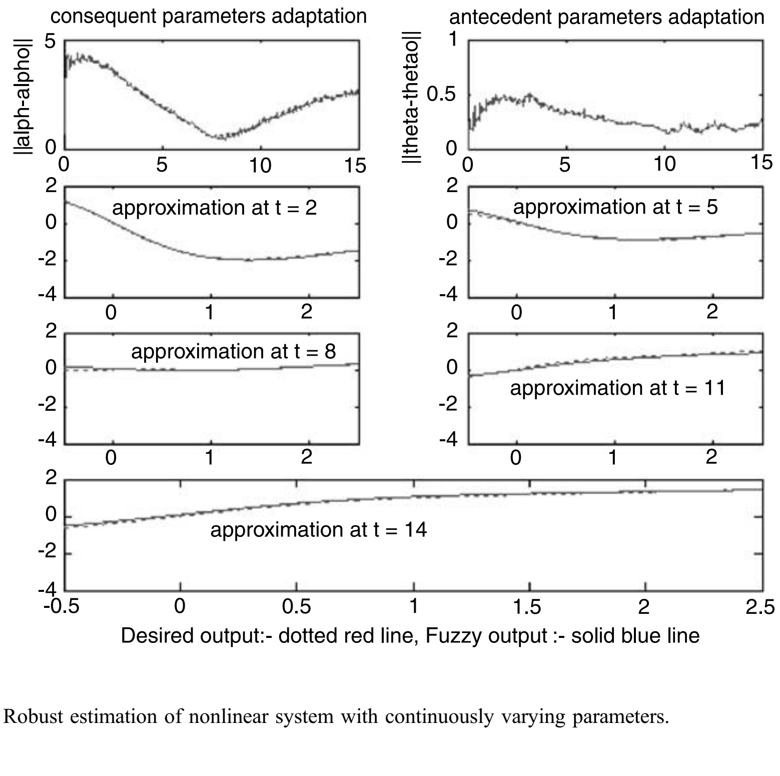

Figure 9. Robust estimation of nonlinear system with continuously varying parameters.

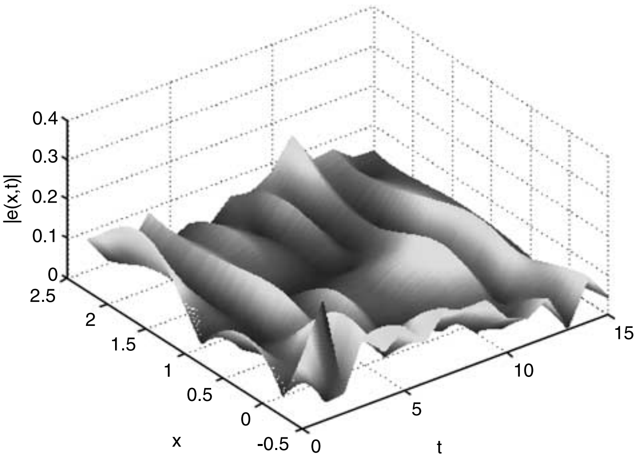

Figure 10. Robust estimation error surface for continuously varying parameters.

continuously in a sinusoidal manner, Figure 9 shows the simulations results and Figure 10 shows the approximation error surface over the simulation time.

The interpretability of fuzzy systems remain preserved during approximation, as a result of putting constraints on the membership curves.

5. A Fuzzy Expert System for Physical Fitness Approximation

This study deals with the application of Robust adaptive identification of Fuzzy Expert system in Medicine. It is clear that physical fitness relationship with various physiological parameters is not so easy to define even in linguistic terms due to inherent presence of vagueness, linguistic uncertainty, hesitation, measurement imprecision, natural diversity

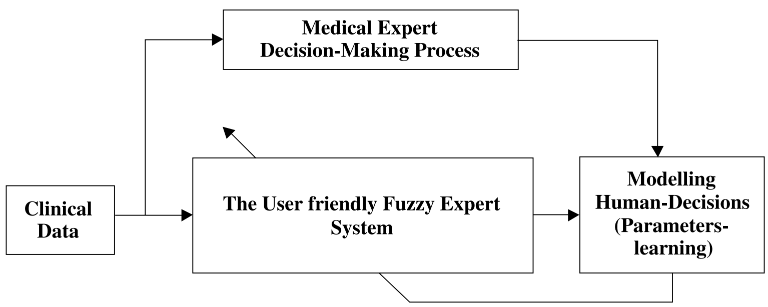

Figure 11. A fuzzy filter for physical fitness approximation.

and subjectivity in the opinion of individuals. So, there is a need of development of algorithms for the automatic construction of such a expert system as shown in Figure 11 from example data in some robust and adaptive manner.

The aim of the task is to approximate the physical fitness or to identify the functional relationship between Fitness and other physiological parameters with a fuzzy system and to filter out various uncertainties lying in the understanding of this relationship. Neural network based approaches for such type of approximation problems have been already considered (Väinämö et al (1996), Nauck and Kruse (1999), Väinämö et al (1998)), but without considering the robustness issues of adaptive identifications schemes. The above developed algorithm (i.e. equations (6–10)) was used to identify the fuzzy expert system. The real world clinical data consist of various physiological variables:- Body Mass Index (BMI), Body Fat percentage, absolute VO_2max, relative VO_2max and relative physical working capacity (PWC/170). This set of 5 parameters are the inputs of the fuzzy system and the output of fuzzy system is the quantification of physical fitness (opinion of Medical Expert). Our Medical Data consist of 160 patients data set and the comment on the physical fitness of these patients was made quantitatively in the range of zero to one by a Medical Expert. However we assume the level of uncertainty in the advice of Medical Expert for each patient to be equal to a random number chosen from a uniform distribution on the interval $[-0.1, 0.1]$ for the fitness scale ranging from zero to one.

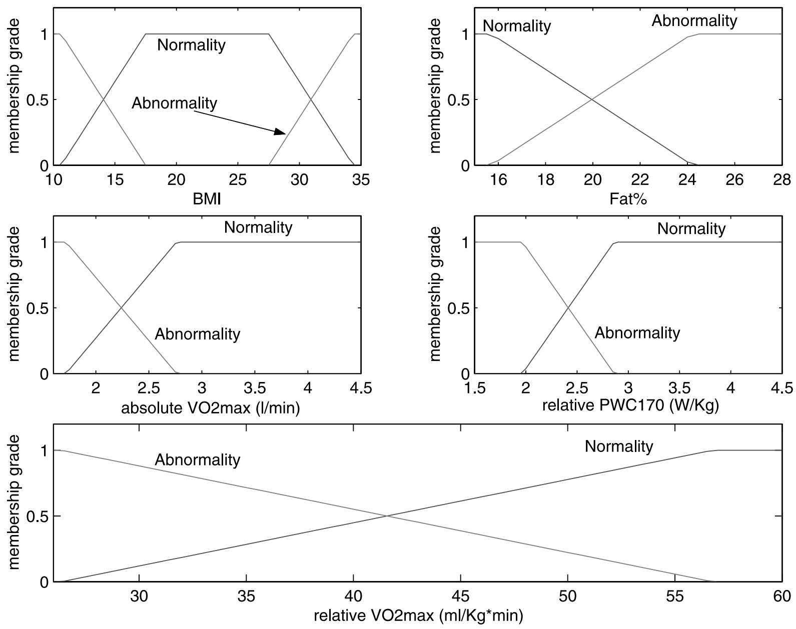

Figure 12. Shape of various membership curves after identification.

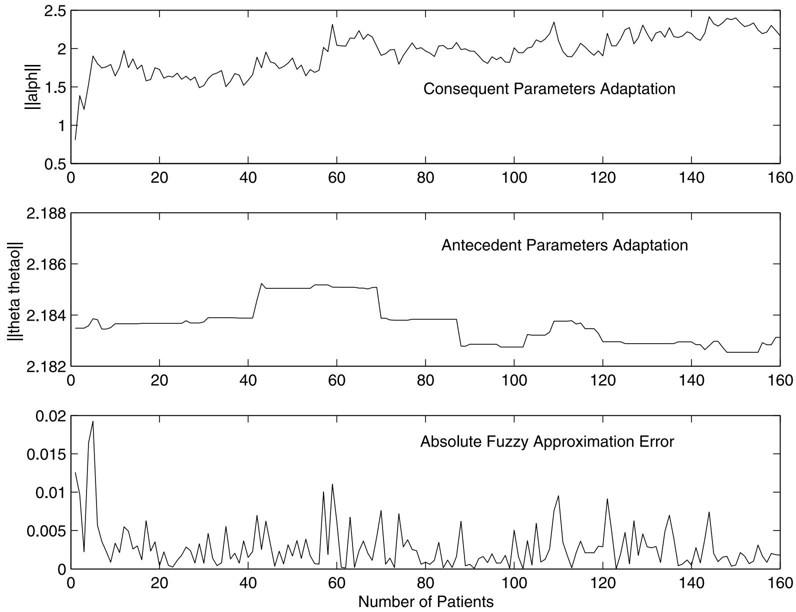

Figure 13. Robust adaptation of fuzzy system for filtering out the uncertainties.

The whole range of all input parameters has been divided into two membership curves with linguistic terms *Normality and Abnormality* resulting the fuzzy inference system to consist of 32 rules. To preserve the interpretability of fuzzy system, the membership curves were prevented to overlap by putting constraints in the training algorithm. The on-line adaptation (identification) of fuzzy system continued till 160 data sets. Figure 12 shows the shape of membership curves afterwards. The Figure 13 shows the adaptation and instant approximation error of the Fuzzy Expert System used to approximate physical fitness.

6. Conclusion

This studies outline an attempt to approximate any time varying process using an interpretable Fuzzy Inference System with bounded uncertainty in the identification data with or without requiring any *a priori* knowledge of a bound on the disturbance and noise. The feasibility of the proposed approach was verified by simulation studies and then used to identify a Fuzzy Expert System for Physical Fitness Approximation with Real World Medical Data.

The identified Fuzzy Expert System consisting of 32 rules can be analyzed critically to have insight of fitness and physiological parameters relationship from physiological point of view to have not only better understanding of this complex relationship but also to realize the Expert system more intelligent, which is the topic of our future work.

Acknowledgments

This work was supported by European Space Agency, under ESTEC Contract 14350/01/NL/SH, MAP Project AO-99-058.

References

AI-Naffouri, T. Y. and A. H. Sayed. (2000). "An Adaptive Filter Robust to Data Uncertainties," In *Proc. Allerton. Conference on Communication, Control and Computing*. Allerton, IL, 1175–1183 (October).

Babuska, R. (2000). "Construction of Fuzzy Systems-Interplay between Precision and Transparency," *Proc. ESIT 2000*. Aachen, 445–452.

Bodenhofer, U. and P. Bauer. (2000). "Towards an Axiomatic Treatment of Interpretability," In *Proc. IIZUKA2000*. Iizuka, 334–339 (October).

Burger, M., H. W. Engl, J. Haslinger, and U. Bodenhofer. (2002). "Regularized Data-Driven Construction of Fuzzy Controllers," *J. Inverse and Ill-posed Problems* 10(2002), 319–344.

Espinosa, J. and J. Vandewalle. (2000). "Constructing Fuzzy Models with Linguistic Integrity from Numerical Data-AFRELI Algorithm," *IEEE Trans. Fuzzy Systems* 8(5), 591–600 (October).

Jang, J.-S. Roger. (1993). "ANFIS: Adaptive-Network-Based Fuzzy Inference Systems," *IEEE Trans.Syst. Man Cybern* 23(3), 665–685.

Lawson, C. L. and R. J. Hanson. (1995). *Solving Least Squares Problems.* Philadelphia: SIAM Publications.

Nauck, D. and R. Kruse. (1997). "Function Approximation by NEFPROX," *Proc. Second European Workshop on Fuzzy Decision Analysis and Neural Networks for Management, Planning, and Optimization (EFDAN'97)*. Dortmund, 160–169.

Nauck, D. and R. Kruse. (1998). "A Neuro-Fuzzy Approach to Obtain Interpretable Fuzzy Systems for Function Approximation," *Proc. IEEE International Conference on Fuzzy Systems 1998 (FUZZ-IEEE'98)*. AK: Anchorage, 1106–1111 (May 4–9).

Nauck, D. and R. Kruse. (1999). "Obtaining Interpretable Fuzzy Classiffication Rules from Medical Data," *Artificial Intelligence in Medicine* 16, 149–169.

Setnes, M., R. Babuka, and H. B. Verbruggen. (1998). "Rule-Based Modeling: Precision and Transparency," *IEEE Trans. Syst. Man Cybern.Part C: Applications and Reviews* 28, 165–169.

Takagi, T. and M. Sugeno. (1985). "Fuzzy Identification of Systems and Its Applications to Modeling and Control," *IEEE Trans. Syst. Man Cybern* 15(1), 116–132.

Väinämö, K., S. Nissilä, T. Mäkikallio, M. Tulppo, and J. Röning. (1996). "Artificial Neural Network for Aerobic Fitness Approximation," *International Conference on Neural Networks (ICNN96)*. Washington DC, USA (June 3–6).

Väinämö, K., T. Mäkikallio, M. Tulppo, and J. Röning. (1998). "A Neuro-Fuzzy Approach to Aerobic Fitness Classiffication: A Multistructure Solution to the Context-Sensitive Feature Selection Problem," *Proc. WCCI '98*. Alaska, USA: Anchorage, 797–802 (May 4–9).

Widrow, B. and M. A. Lehr. (1990). "30 Years of Adaptive Neural Networks: Perceptron, Madline and Backpropagation," *Proceeding of the IEEE* 78(9), 1415–1422.

Zadeh, L. A. (1973). "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Trans. Syst. Man Cybern* 3(1), 28–44.