A Fuzzy APRI for Hepatic Fibrosis Prediction and Chronic Liver Disease Classification

Wellington Cardia[1], **cardia@gmail.com**
Ernesto Araujo[1,2,3,4,5], **ernesto.araujo@{lit.inpe.br, unifesp.br}**
[1]Hospital Municipal Dr. José de Carvalho Florence, HJCF
Av. Saigiro Nakamura, 800
12.220-280, Av. Saigiro Nakamura, 800, São José dos Campos, SP, Brazil

[2]Associação Paulista para o Desenvolvimento da Medicina (SPDM)
R. Napoleão de Barros, 715
04.024-002, São Paulo, SP, Brazil

[3]Universidade Federal de São Paulo, UNIFESP
Health Informatics Department, DIS
R. Botucatu, 862
04.023-062, São Paulo, SP, Brazil

Instituto Nacional de Pesquisas Espaciais, INPE
[4]Integration and Testing Laboratory, LIT
[5]Space Engineering and Technology, ETE
Av. Astronautas, 1758
12.227-010, São José dos Campos, SP, Brazil

Abstract. *A Fuzzy Model for Hepatic Fibrosis Prediction and Chronic Liver Disease is proposed in this paper. This model is obtained by using a non-invasive, serological, and Aspartate Aminotransferase-to-Platelet Ratio Index (APRI) approach. The chronic liver disease is a high prevalence and incidence clinical condition with large etiological spectrum in which alcohol and C virus are the main agents. The clinical events related to the chronic liver disease are usually severe and associated to its complications such as portal hypertension, liver insufficiency degree, and the hepatocellular carcinoma development. The fibrosis degree achieved by Metavir classification has been the standard for antiviral therapy indication for chronic liver C hepatitis. This procedure has the disadvantage of risks due to the anaesthesical and pneumoperitonium complications and massive visceral lesions as well as visceral perforation risks by trocarters introduction. In order to overcome these problems serological and noninvasive markers have been studied to evaluate their accuracy in the liver fibrosis degree prediction. Nevertheless, the information obtained by employing these markers not always allow to decide upon safety and appropriate prophylaxis or therapeutic approaches due to the inherent imprecision and vague borderlines. In order to deal with this sort of problem, fuzzy set theory and fuzzy logic is employed as the basic component yielding a Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index (FAPRI). Experimental data are employed to generate this FAPRI model by using the Adaptive Neural Fuzzy Inference System (ANFIS). This paper demonstrates the richness of fuzzy set theory, fuzzy logic, and approximate reasoning in medicine and health care through the proposed FAPRI when employed for Hepatic Fibrosis Prediction and Chronic Liver Disease classification.*

Keywords: *Neuro-Fuzzy System, Aspartate Aminotransferase-to-Platelet Ratio Index (APRI), Hepatic Fibrosis Prediction, Chronic Liver Disease, Decision Support System*

1. INTRODUCTION

The chronic liver disease is a high prevalence and incidence clinical condition with large etiological spectrum in which alcohol and C virus are the main agents (World Health Organization, 2000). The clinical events related to the chronic liver disease are usually severe and associated to its complications such as portal hypertension, liver insufficiency degree, and the hepatocellular carcinoma development (Thomas and Seeff, 2005). The liver fibrosis severity caused by the action of aggressive agents and inflammatory process present high relevance since it is an important factor related to the complication development such as esophagus gastric varices and hepatocellular carcinoma (Thomas and Seeff, 2005). The fibrosis degree achieved by Metavir classification has been the standard for antiviral therapy indication for chronic liver C hepatitis (Strader et al., 2007).

The gold standard for evaluating liver fibrosis in dealing with histopathological studies is carried out by blind or laparoscopic liver biopsy. This procedure has the disadvantage of risks due to the anaesthesical and pneumoperitonium complications and massive visceral lesions as well as visceral perforation risks by trocarters introduction (Piccinino et al., 1986; Garcia-Tsao and Boyer, 1993).

Some serological and noninvasive markers have been studied to evaluate their accuracy in the liver fibrosis degree prediction in order to find out less invasive and risky alternatives to patients (Oberti et al., 1997; Imbert-Bismut et al.,

2001). Most of studies concerning these markers have been performed in chronic liver C disease due to its prevalence as well as the necessity of evaluating the degree of liver fibrosis for therapeutic planning (Cales et al., 2005). The liver fibrosis degree predictability limitations of these markers are concerned to the cut-off determination from what it is possible to guarantee the fibrosis severity.

The information obtained by employing these serological and noninvasive markers to fibrosis severity prediction not always allow to physicians decides upon safety and appropriate prophylaxis or therapeutic approaches. These results are most of time decided in imprecise and vague borderlines generating, thus, an imperfect evaluation and decision.

In order to deal with this sort of problem, fuzzy set theory and fuzzy logic is employed in this paper as the basic component for describing the while the artificial neural network is used for tuning the parameters of the model. Despite well established in the Computational/Artificial Intelligence field, this method has presented to be useful in medicine and health care as demonstrated by handing experimental and actual data.

The proposed approach exploits the advantages of the existing medical parameters and methods meanwhile allows to incorporate the advantage of evaluating the subjectivity and worth the linguistic information (Zadeh, 1965, 1996) as well as the ability of learning of artificial neural network.

2. LIVER FIBROSIS SEVERITY EVALUATION

The chronic liver C hepatitis is a great problem in public health treatment reaching more than 200 millions of human beings worldwide (World Health Organization, 2000). Despite the successful therapeutic approach with peguilate interferon and ribavirin that drive to a sustained virological response in more than 50% of patients, only a minority is eligible for this treatment (Pawlotsky, 2005). Most of patients present the risk of progressive fibrosis that may lead to cirrhosis and its complications such as hepatocelular carcinoma, severe portal hypertension, and end stage liver disease. In order to estimate prognostics and therapeutic decisions the accuracy of liver fibrosis degree is a priority conduct (McHutchison et al., 2006).

The liver biopsy is the golden standard procedure to collect samples of liver tissue to evaluate all sort of chronic liver disease. In viral hepatitis, in special C and B hepatitis, the biopsy allows to evaluate the degree of necroinflammatory activity and, most important, the degree of liver fibrosis (Strader et al., 2007). The liver biopsy procedure may be carried out in a blind or laparoscopic manner. While the first one is characterized by being relative simple, safe, and with lower complication risks the second manner is more complex and encompasses higher possibility of complications. Furthermore, a fragment of biopsy must achieve a size larger than $1 - 1.5$ cm because if lower than that it does not permit to the pathologist an appropriate diagnosis. Even being adequate, another important aspect concerning the size of fragment of biopsy is that it represents a small amount when compared to the total area of liver surface and so may not be representative of the actual condition of the organ. Additionally, discrepancies analysis from different pathologists may occur due to distinct experiences and inherent subjectivities present in this process. This procedure has the disadvantage of risks due to the anaesthesical and pneumoperitonium complications and massive visceral lesions as well as visceral perforation risks by trocarters introduction. The necessity of verifying the evolution in time of the liver tissue exposes patients to risk the number of interventions that are required for biopsies. Despite the alternatives for reducing this kind of approach, the cost is another limiting factor for performing such an activity (Bravo et al., 2001; Ishak et al., 1995).

In the last years, noninvasive diverse markers have been proposed to predict with accuracy the degree of hepatic fibrosis (Sebastiani and Alberti, 2006). Transitory hepatic elastography is a prominent technique based on ultrasound but expensive and restricted to specialized centers (Sandrin et al., 2003). There are serological markers, as glycoprotein (hyaluronic acid – also called hyaluronan or hyaluronate –, laminin, YKL-40), collagen, metalloproteinase and its inhibitors, cytokines (transforming growth factor beta), that have demonstrated to be associate to fibrosis and, due to that, have been used in methods for evaluate the hepatic fibrosis. Only to mention few, examples are Fibrometer, Hepascore, Fibrospect (Rockey and Bissell; Patel et al., 2004). Even though, these markers are much cheaper than those which involve hepatic biopsies, they are still expensive and not available in the most centers of treatment and diagnosis.

In contrary, markers that are indirectly related to the fibrosis such as biochemical parameters, platelet measurement, alpha-2-Macroglobulin are widely available and incorporated in tests as Fornsindex, Fibrosis Probability Index e o Fibrotest (Forns et al., 2002; Sud et al., 2004). Nevertheless, these approaches present limitations as adequate validation, difficulty in differentiating intermediary classification of fibrosis, and high cost.

Due to the limited availability and high cost of serological fibrosis markers the Aspartate Aminotransferase-to-Platelet Ratio Index (APRI) was developed and validate in (Wai et al., 2003). APRI is computed as:

$$APRI = \frac{AST_p}{AST_L \times Plat} \times 100 \tag{1}$$

where AST_P is the Aspartate Aminotransferase of the patient in [U/L], AST_L is normal upper limit of the Aspartate Aminotransferase in the laboratory in [U/L], and $Plat$ is the Platelet count in (10^9/L).

The advantage of this index is lower cost of laboratory tests that are commonly employed to evaluate the chronic liver disease. The severity fibrosis identification is grouped into two classes in a range from 0 to 10. The first class with negative

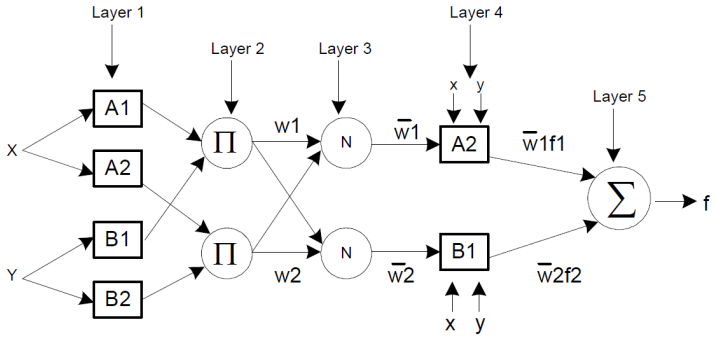


Figure 1. ANFIS Structure.

predictive value of 86 % are those indexes with score lower than 0.5 while in the other class is the positive predictive value of 88 % for those scores higher than 1.5. Based on these predictive values, the authors in (Wai et al., 2003) advocate that the APRI is able to avoid 50 % of the biopsies in the patients. Nevertheless, subsequent studies has shown controversial results mostly due to differences in prevalence of severe fibrosis that change with population and the laboratorial reference levels of AST (Wai et al., 2003).

Among the existing fibrosis diagnosis, APRI is the simplest. It has initially shown high degree of accuracy in identifying severe fibrosis and cirrhosis in patients of chronic C hepatitis. In the other hand, it must be taken into account the limitations in the laboratorial parameters in medical practice given by cut-off points for establishing (or separating) what belongs to one class from another, for instance, from normal to abnormal. These cut-off points disables a certain degree from one class to another not allowing overlapping of categories. Moreover, clinical reasoning are excluding by taking into account isolated variables and discarding others. In doing so, variables are isolated from others even though when it is required that the variables work together, and so with prejudice in the analysis.

In order to overwhelm these problems, this paper employs the fuzzy set theory for allowing in dealing with the inherent imperfection and subjectivity of clinical facts and fuzzy logic for permitting aggregate simultaneously diverse variables in the same system. Nevertheless, since fuzzy systems does not carry learning attributes, artificial neural network based on backpropagation, multilayer perceptron architecture is used for given such a ability to the resulting fuzzy systems.

3. NEURO-FUZZY SYSTEM

The model used in this work is the well established hybrid system named Adaptive Neuro-Fuzzy Inference System (ANFIS) (Jang, 1993). One of the main advantages of fuzzy models is related to its capacity to mimic human reasoning allowing knowledge representation in the form of fuzzy IF-THEN rules and fuzzy sets theory. Fuzzy sets are appropriate to deal with uncertainty, imprecise measures and incomplete information. In contrary, one of the main drawbacks of fuzzy models is its lack of learning capacity. In the other hand, artificial neural network suppress this pitfall by allowing the system to learn with are low-level computational algorithms presenting learning capacity input-output examples (Lin and Lee, 1996). When working in synergy, fuzzy systems and Artificial Neural Networks compose a hybrid system that takes advantage of their individual advantages and produce a Neuro-Fuzzy system with capacities of learning, adaptation, optimization meanwhile is able to deal with uncertain, imprecise, vague information. In doing so, the system is able to generalize when dealing with large amounts of numerical data and with imperfect knowledge representation through fuzzy rules (J.C., 1992).

This neuro-fuzzy approach is effective in processing numerical data and presents distributed computational characteristic allowing that each node in the network to adjust its connections to obtain the best possible input-output mapping after learning from data. The neuro-fuzzy model may assume the fuzzy Takagi-Sugeno (T-S) model (Takagi and Sugeno, 1985) approach used in many problems of diverse areas. The T-S models may be represented by the following general form:

$$\text{Rs}^{(j)} : \text{IF} < x_1 \text{ is } A_1^j > \text{AND} \ldots \text{AND } < x_m \text{ is } A_m^j > \text{ THEN } y_j = f(\cdot)\,. \tag{2}$$

The <IF statements> defines the premise part that is featured as linguistic terms in the proposition form, $< x_i \text{ is } A_i^j >$, while the <THEN functions> constitutes the consequent part of the j-th rule of the fuzzy system. The vector $\mathbf{x} = [\mathbf{x_1}, \ldots, \mathbf{x_i}]^\mathbf{T}$ represents the i-th input vector of the premise, $\forall\, i = 1, \ldots, m$, and so, the dimensionality of the premise space. The terms A_i^j are linguistic labels of fuzzy sets. The j-th rule output, $y_j = f(\mathbf{x^j}, \mathbf{w^j})$, is usually function of the consequent input vector, $\mathbf{x} = [\mathbf{x_1^j}, \ldots, \mathbf{x_{q_j}^j}]^\mathbf{T}$, $\mathbf{w} = [\mathbf{w_1^j}, \ldots, \mathbf{w_{y_j}^j}]^\mathbf{T}$, that compose the consequent parameter set. One of the advantages of the TS model is that it does not contain defuzzification interface because it process and produces crisp data.

The firing strength of the j-th rule, $Rs^{(j)}$, represents its activation level and may, for instance, be chosen as the algebraic product:

$$w_j(\mathbf{z}) = \mathbf{w}_{\mathbf{A_1^j}}(\mathbf{x_1})\mathbf{w}_{\mathbf{A_2^j}}(\mathbf{x_2}) \ldots \mathbf{w}_{\mathbf{A_m^j}}(\mathbf{x_m}) \; . \tag{3}$$

A neuro-fuzzy model equivalent to the Takagi-Sugeno system is depicted in Fig. 1. This example has two inputs x, y, one output f and two rules. The ANFIS structure is composed by the following elements:

1) *Input Layer*: Computes the degree of relevancy of the inputs x,y with relation of the subgroups fuzzy that form the partition of x and y, or either, the process of fuzzification.

2) *Membership Layer*: Computes the degree of activation of each rule, with that degree the consequence of the rule is being taken care of. The function for this layer is a *T-norm* that uses the probabilistic form. In this, the outputs of the neurons given by Eq. (4) are equivalent to (3):

$$w_1 = \mu_{A_1}(x_1) \cdot \mu_{A_2}(x_2) \cdot \mu_{A_3}(x_3) \tag{4}$$

3) *Rule and Norm Layer*: Layer 3 is the degree of relevance of each rule, already normalized. Each point i calculates the reason for the firing strength of rule j for the sum of the firing strength of all the rules. The outputs of points this layer referring to Fig. 1 are:

$$\begin{aligned} \bar{w_1} &= w_1(w_1 + w_2 + w_3) \\ \bar{w_2} &= w_2(w_1 + w_2 + w_3). \end{aligned} \tag{5}$$

4) *Layer consequent*: Layer 4 contains the function of activation of the neurons, consequence part of the rules (Ci). It is calculated by the product of the normalized firing strength ($S_i \forall i = 1, 2, 3$) and the value of the consequence of the rule. The output values of each point of this layer are given by:

$$\begin{aligned} H_1 &= \bar{w_1} \cdot C_1 \\ H_2 &= \bar{w_2} \cdot C_2. \end{aligned} \tag{6}$$

5) *Output layer*: It computes the necessary output of the network as:

$$F = H_1 + H_2 \; . \tag{7}$$

Learning on a neural network consists of adjusting values in the synaptic connections. It can be made by means of a system specialist or through a learning algorithm (Rumelhart et al., 1987).

The parameters of membership functions are obtained by using the backpropagation algorithm achieving a supervised learning. This approach attempts to iteratively search a minimal error determined by the difference between the desired and actual measured outputs. The error signal is backward, then, of the output layer for each element of the previous intermediate layer that contributes directly to form the output in a feedforward manner. Nevertheless, each element of the intermediate layer just receives a portion of the signal of error total, proportional just to the relative contribution of each element in the formation of the original output. This process repeats, layer after layer, until each element of the network receives an error signal that describes its relative contribution to the total error. Based on this error, the weights of the connections are updated for each element allowing the neural network to converge all the patterns of the training group (Rumelhart et al., 1987).

The initial weights, the learning constant and momentum constant are among the most important factors determining the convergence of the backpropagation neural network (Lin and Lee, 1996; Wessels and Barnard, 1992). In each iteration of the learning method the parameters of the premises are fixed. This output is calculated from the linear combination of the parameters of the consequent part.

The parameters of the consequences are identified by the method Least Mean Square (LMS), which it carries through the adjustment of the coefficients that will be used in the synaptic weights during the stage of backpropagation. The error signals backward propagated to adapt the parameters of the premises, by means of the descending gradient (Rumelhart et al., 1987).

4. FUZZY ASPARTATE AMINOTRANSFERASE-TO-PLATELET RATION INDEX (FAPRI)

The objective of this paper is to evaluate the accuracy of the APRI, AST, and Blood Platelet Counting as a predictor factor of for Hepatic Fibrosis by designing a Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index. The study took place at the Hospital Municipal Dr. José de Carvalho Florence in São José dos Campos, SP, Brazil. The project was accepted by the Ethics Committee of the Hospital Municipal Dr. José de Carvalho Florence and approved by the Research Ethics Committee (CEP – **C**omitê de **É**tica em **P**esquisa) of the Universidade de Taubaté.

The nonlinear fuzzy mapping obtained by FAPRI for APRI, AST, and Blood Platelet Counting are, respectively, shown in Fig. 2, 3, and 4. They are designed through ANFIS by using experimental data. The data employed in this paper is obtained in the medical record from the ambulatory center experts in the Hospital Municipal Dr. José de Carvalho

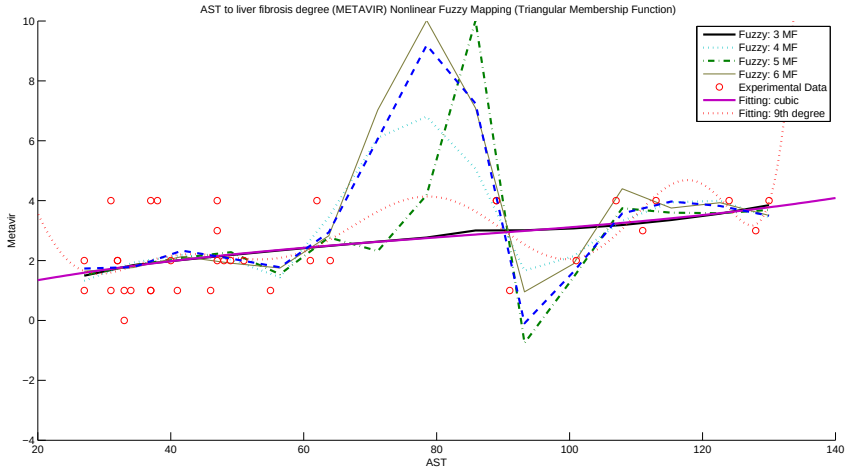
Table 1. Liver Fibrosis Degree obtained by FAPRI

	R^2	Median	Standard Deviation
AST-Gauss-3mf	0.295320859566904	**0.000000993881593**	1.064724652406316
AST-Gauss-4mf	**0.412491946956945**	**0.000000431521044**	0.972184099554713
AST-Gauss-5mf	0.398501205214649	0.000017124971219	0.983691644423898
AST-Trap-3mf	0.274802644676742	**0.000000379686326**	1.080114279388865
AST-Trap-4mf	0.301437026579527	**0.000000910538447**	1.060094016166904
AST-Trap-5mf	0.391776923909258	−0.000085064040712	0.989174806745589
AST-Tri-3mf	0.264645856481610	−0.000001267015477	1.087651777005659
AST-Tri-4mf	0.351799474654524	**0.000000990793449**	1.021165869987335
AST-Tri-5mf	**0.436268649148488**	0.000045243629633	0.952308575741628
AST-Bell-3mf	0.278258069102269	**0.000000442742457**	1.077537939496850
AST-Bell-4mf	0.398528262396305	0.000006503957065	0.983669519637254
AST-Bell-5mf	**0.473811282979354**	**−0.000000867352327**	0.920051986712636
APRI-Gauss-3mf	**0.494978147188034**	**−0.000000178282100**	0.901356691569082
APRI-Gauss-4mf	**0.584224011716374**	−0.000011107223398	0.817845567024001
APRI-Gauss-5mf	**0.626638682908897**	0.000020153615670	0.775008135757231
APRI-Trap-3mf	0.469516638612546	**−0.000000201856231**	0.923798994461731
APRI-Trap-4mf	0.507551621045788	**−0.000000388388547**	0.890065480086304
APRI-Trap-5mf	0.522849346676328	**−0.000000608802183**	0.876131637517565
APRI-Tri-3mf	0.449807810642573	0.000004889746386	0.940803256846231
APRI-Tri-4mf	0.537935677881712	0.000016141073274	0.862169826656511
APRI-Tri-5mf	0.541583303445630	**0.000000938455760**	0.858760015959966
APRI-Bell-3mf	0.513172053890248	**−0.000000200238979**	0.884971637409571
APRI-Bell-4mf	**0.577114753770571**	0.000001333377803	0.824808007532349
APRI-Bell-5mf	**0.623810799506599**	0.000016799309555	0.777937601567144
PLAQ-Gauss-3mf	0.581905282665095	0.000000334159794	0.820122906022205
PLAQ-Gauss-4mf	0.592721403043118	0.000001227262937	0.809445095067165
PLAQ-Gauss-5mf	**0.709616330650909**	−0.000049461782292	0.683482965382637
PLAQ-Trap-3mf	0.579828174563472	0.000000768674720	0.822157580654608
PLAQ-Trap-4mf	0.575610657476038	0.000000517790450	0.826273522728210
PLAQ-Trap-5mf	0.587744323580282	0.000000851928261	0.814375920422108
PLAQ-Tri-3mf	0.570734385252676	**−0.000000070741045**	0.831006943743467
PLAQ-Tri-4mf	0.574342280278346	−0.000000780305445	0.827507348010825
PLAQ-Tri-5mf	0.596982782054796	0.000014473013521	0.805199324875848
PLAQ-Bell-3mf	0.576657947112993	−0.000000280890797	0.825253371964020
PLAQ-Bell-4mf	0.586594898534638	0.000001064037279	0.815510425745168
PLAQ-Bell-5mf	0.594847576395197	0.000000738674100	0.807329500721423

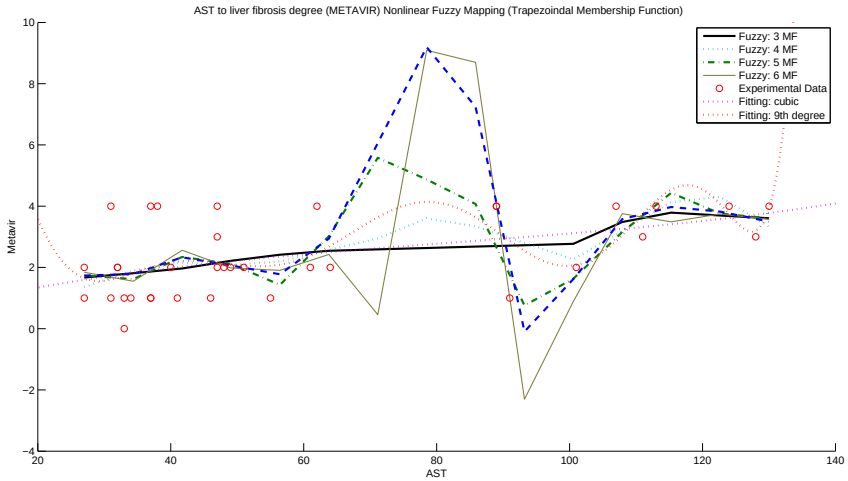
Florence. The database includes 36 patients with chronic liver C hepatitis with viral load detected by the *Polimerase Chain Reaction* (PCR). In parallel, hepatic biopsies was carried out to apply the Metavir scale in order to determine the necroinflammatory activity degree of the disease and the degree of hepatic fibrosis. Patients with alcohol intake (etilist) less than one year from biopsies were excluded as were those co-infected by the Human Immunodeficiency Virus (HIV) and/or Hepatitis B Virus (HBV).

The premise of the Takagi-Sugeno fuzzy model correspond to the (i) Aspartate Aminotransferase (AST) level, (ii) blood platelet count, and (iii) Aspartate Aminotransferase-to-Platelet Ratio Index (APRI) mapped into the consequent that is the liver fibrosis degree determined by the (iv) Metavir system. The input universe of discourses are parted into 3, 4, 5 and 6 membership functions associate to Triangular, Trapezoidal, and Gaussian shapes. The output universe of discourse is parted into four classes, i.e., four membership functions. The AST level and the blood platelet count are selected in a serum assessment performed in a period limited in three months, before or after the hepatic biopsies being carried out. The APRI is computed according to Eq. (1) derived from those two previous measures.

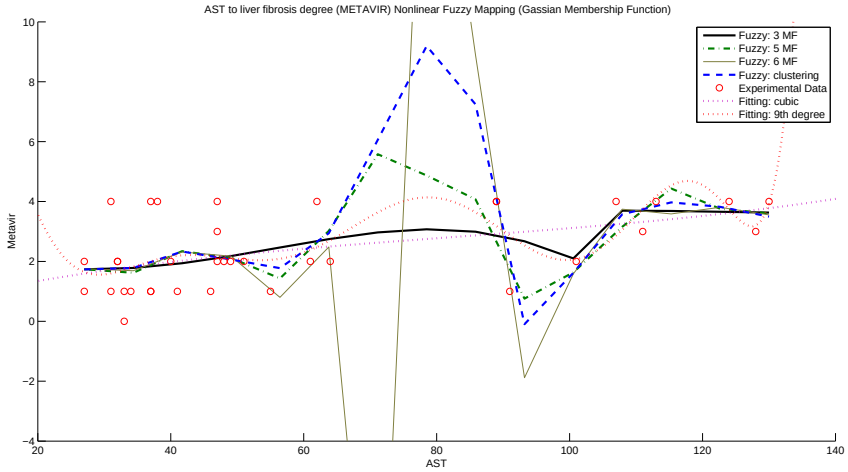
The final fuzzy mappings are compared to another fuzzy learning technique named subtractive clustering as well as to statistical cubic and higher order (9th and 5th) approaches. One of the main characteristic of the fuzzy subtractive clustering learning approach is the absence of determining the number of membership functions; they are automatically supplies by the algorithm. The AST to liver fibrosis degree (Metavir) nonlinear fuzzy mappings when using 3, 4, 5, or 6



(a) Triangular Membership Function with 3, 4, 5, and 6 Membership Function.

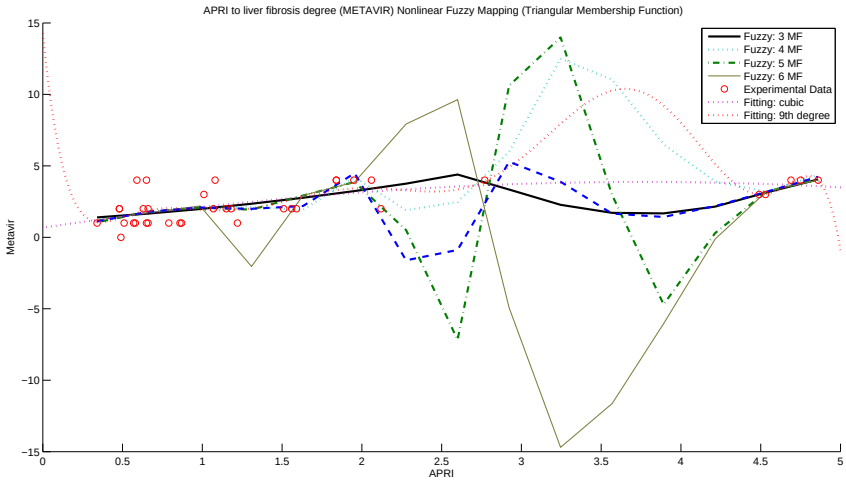


(b) Trapezoidal Membership Function with 3, 4, 5, and 6 Membership Function.

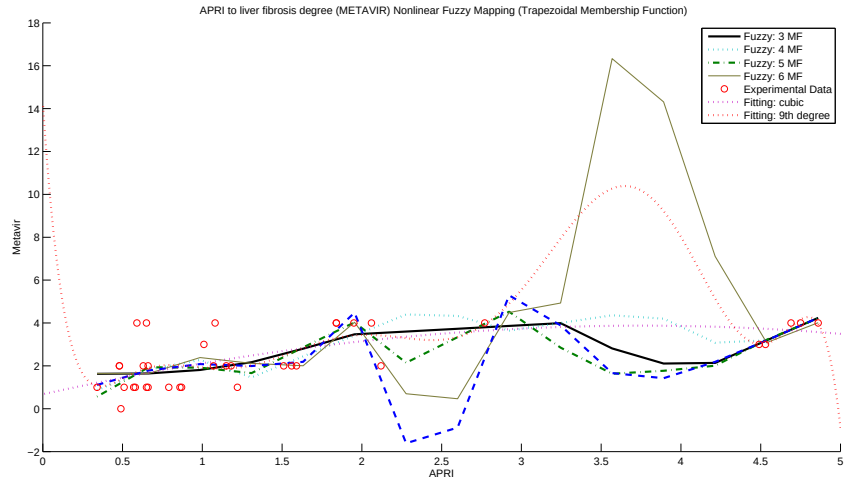


(c) Gaussian Membership Function with 3, 4, 5, and 6 Membership Function.

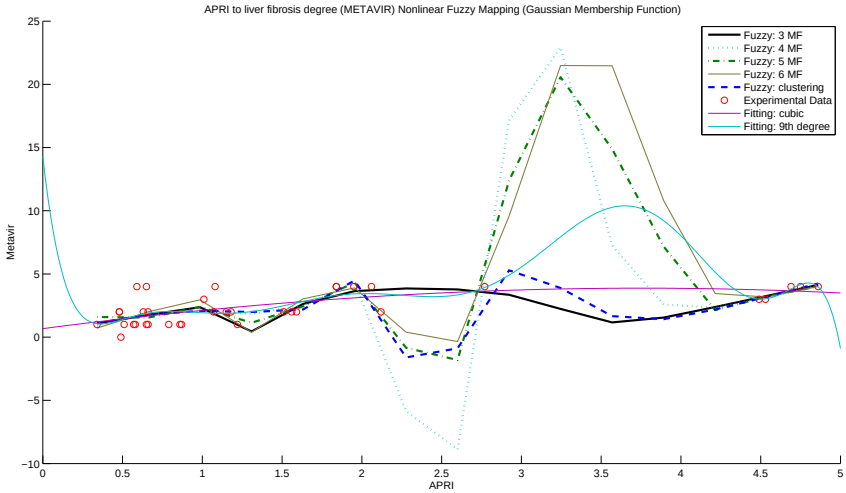
Figure 2. Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index (FAPRI) for AST to METAVIR Fuzzy Nonlinear Mapping with Triangular, Trapezoidal, and Gaussian Membership Function and with 3, 4, 5, and 6 Membership Function based on ANFIS compared to Fuzzy Clustering and Statistical Methods (cubic and 9th function).



(a) Triangular Membership Function with 3, 4, 5, and 6 Membership Function.

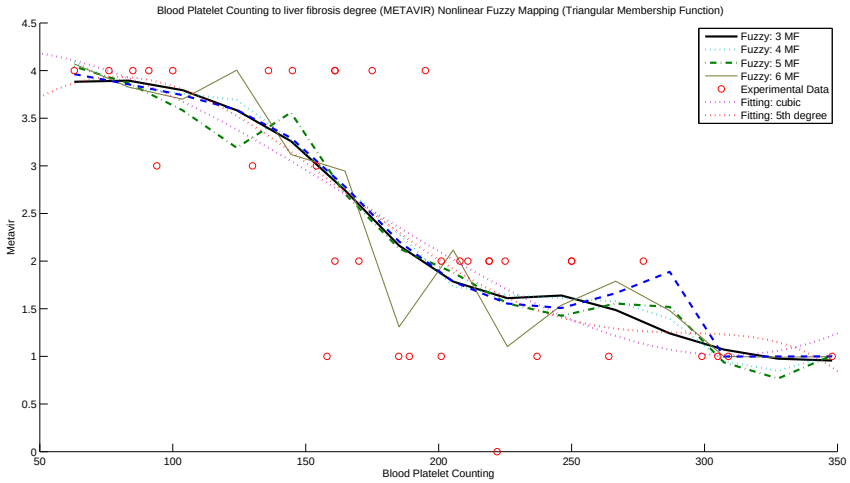


(b) Trapezoidal Membership Function with 3, 4, 5, and 6 Membership Function.

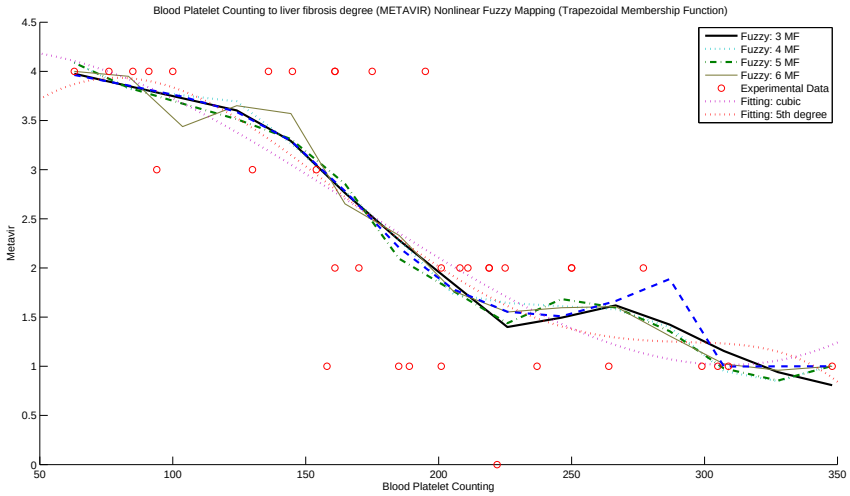


(c) Gaussian Membership Function with 3, 4, 5, and 6 Membership Function.

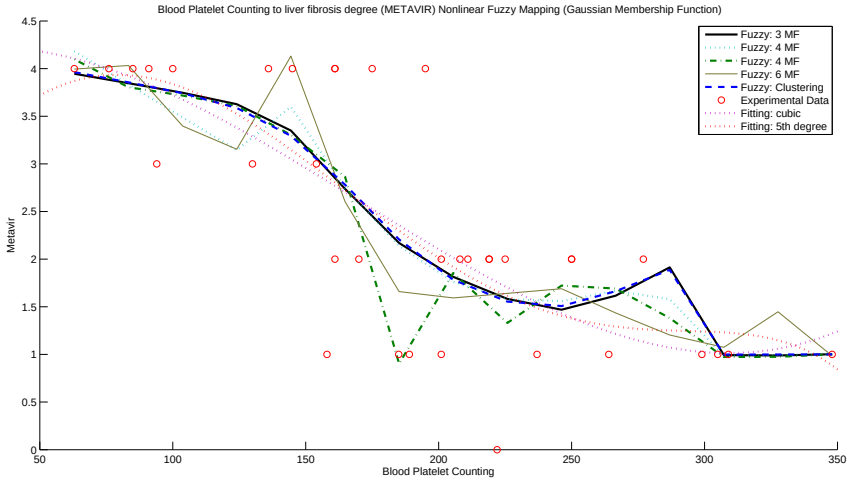
Figure 3. Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index (FAPRI) for APRI to METAVIR Fuzzy Nonlinear Mapping with Triangular, Trapezoidal, and Gaussian Membership Function and with 3, 4, 5, and 6 Membership Function based on ANFIS compared to Fuzzy Clustering and Statistical Methods (cubic and 9th function).



(a) Triangular Membership Function with 3, 4, 5, and 6 Membership Function.



(b) Trapezoidal Membership Function with 3, 4, 5, and 6 Membership Function.



(c) Gaussian Membership Function with 3, 4, 5, and 6 Membership Function.

Figure 4. Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index (FAPRI) for Blood Platelet Degree to METAVIR Fuzzy Nonlinear Mapping with Triangular, Trapezoidal, and Gaussian Membership Function and with 3, 4, 5, and 6 Membership Function based on ANFIS compared to Fuzzy Clustering and Statistical Methods (cubic and 5th function).

triangular membership functions are depicted in Fig. 2(a); 3, 4, 5, or 6 trapezoidal membership functions are depicted in Fig. 2(b); and 3, 4, 5, or 6 Gaussian membership functions are depicted in Fig. 2(c). The advantage of showing in separate graphics is to avoid the strong overlapping of fuzzy mappings. The same proceeding is carried out for APRI to liver fibrosis degree (Metavir) nonlinear fuzzy mappings when using 3, 4, 5, or 6 triangular membership functions are depicted in Fig. 3(a); 3, 4, 5, or 6 trapezoidal membership functions are depicted in Fig. 3(b); and 3, 4, 5, or 6 Gaussian membership functions are depicted in Fig. 3(c). Again, for blood platelet counting to liver fibrosis degree (Metavir) nonlinear fuzzy mappings when using 3, 4, 5, or 6 triangular membership functions are depicted in Fig. 4(a); 3, 4, 5, or 6 trapezoidal membership functions are depicted in Fig. 4(b); and 3, 4, 5, or 6 Gaussian membership functions are depicted in Fig. 4(c).

The performance criterion (fitness function) chosen for evaluate the relationship between the real output and the estimate output is the *Pearson multiple correlation coefficient index*. This coefficient represents the R^2 of TS fuzzy model as given by:

$$R^2_{training} = 1 - \frac{\sum_{k=1}^{Na} \left[y(k) - \widehat{y}(k)\right]^2}{\sum_{k=1}^{Na} \left[y(k) - \overline{y}\right]^2} \tag{8}$$

where Na is the total number of samples evaluated, and $\bar{y}(k)$ is the system real output. When $R(.)^2$ is close to unit a sufficient accurate model for the measured data of the system is found.

The Pearson multiple correlation coefficient indexes, R^2, the Median, and the Standard Deviation for estimated output are presented in Table 1. As it is possible to note, the fuzzy systems that is best associate to the liver fibrosis degree (Metavir) are those who have Gaussian membership functions. Five membership functions present, in general, best performance when analyzing their number. When three membership functions are used there is a smooth variation. Nevertheless, they are quite similar with fitting cubic function. Another observation that can be extracted from the results is that when taking into account isolated inputs the blood platelet counting (PLAQ) is a primordial element in diagnosis when compared to the AST and APRI. In contrary, AST present the worst contribution for the liver fibrosis degree.

When analyzing the AST, there is a concentration of data in the intervals from 20 to 70 and from 100 to 140. Due to this, the interval without data, from 70 to 100, the fuzzy mappings show a not smooth variation. Nevertheless, independent of the amount of membership functions and their shape, it is noticed that the Metavir may be associate to these two clusters of data in a growing manner. The Metavir is then directly associated to the AST value. The same characteristic is repeated when analyzing the APRI. Here, there is a concentration of data from 0.4 to about 2.2. The interval that there is no data there also is a not smooth variation. It is worth mentioning that this concentration may be explained by the fact that APRI is computed according to Eq. (1) and so strongly influenced by AST that, as mentioned, present a high concentration in the beginning of the available data. The analysis of blood platelet counting is characterized by a distribution of data in the input universe of discourse, although presents a reduced concentration in the output universe of discourse, given by the Metavir. The fuzzy mapping reproduced adequately the decaying and the inverse relationship between the PLAQ and the Metavir. It is interesting to note that the variation in data is best represented by the fuzzy system with five Gaussian membership functions what is closely related to the best Person coefficient computed as shown in Table 1.

5. CONCLUSION AND FUTURE WORK

A Fuzzy Aspartate Aminotransferase-to-Platelet Ratio Index (FAPRI) for Hepatic Fibrosis Prediction and Chronic Liver Disease is proposed to eliminate the disadvantage of the traditional approach. In this paper, the unidimensional FAPRI is obtained by using a non-invasive, serological approach. The fibrosis degree achieved by Metavir classification and its relationship with AST, APRI and blood platelet count are computed by using artificial neural network in the ANFIS architecture. Gaussian membership function is the shape that produced the best results. In general, five membership functions presented best accuracy. It was also possible to conclude that blood platelet count is the input that influences most the liver fibrosis degree while AST has not presented appropriate performance. Since APRI is computed from AST and blood platelet count the results are influenced by them. Despite the database is characterized by concentration in determined intervals in the respective universe of discourse, the results are shown really worthy and deserve to be better explored. Additionally, future work is driven to aggregate the distinct inputs and verify the influence of one in another.

6. REFERENCES

Bravo, A. A., Sheth, S. G., and Chopra, S. 2001. Liver biopsy. N Engl Journal Med Vol. 344, pp. 495–500.

Cales, P., Oberti, F., Michalak, S., Hubert-Fouchard, I., Rousselet, M. C., Konate, A., and et al. 2005. A novel panel of blood markers to assess the degree of liver fibrosis. Hepatology Vol. 42, pp. 1373–1381.

Forns, X., Ampurdanes, S., Llovet, J. M., Aponte, J., Quinto, L., Martinez-Bauer, E., and et al. 2002. Identification of chronic hepatitis c patients without hepatic fibrosis by a simple predictive model. Hepatology Vol. 36, pp. 986–992.

Garcia-Tsao, G. and Boyer, J. L. 1993. Outpatient liver biopsy: how safe is it? Ann Intern Med Vol. 118, pp. 150–153.

Imbert-Bismut, F., Ratziu, V., Pieroni, L., Charlotte, F., Benhamou, Y., and Poynard, T. 2001. Biochemical markers of liver fibrosis in patients with hepatitis c virus infection: a prospective study. Lancet Vol. 357, pp. 1069–1075.

Ishak, K., Baptista, A., Bianchi, L., Callea, F., Groote, J. D., Gudat, F., and et al. 1995. Histological grading and staging of chronic hepatitis. J Hepatol Vol. 22, pp. 696–699.

Jang, J.-S. R. 1993. Anfis: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man and Cybernetics Vol. 23, No. 3, pp. 665–685.

J.C., B. 1992. Guest editorial. IEEE Trans. On Neural Networks Vol. 3, No. 1, pp. 641.

Lin, C.-T. and Lee, C. S. G. 1996. Neural fuzzy systems: A neuro-fuzzy synergism to intelligent systems. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

McHutchison, J., Poynard, T., and Afdhal, N. 2006. Fibrosis as an end point for clinical trials in liver disease: a report of the international fibrosis group. Clin Gastroenterol Hepatol Vol. 4, pp. 1214–1220.

Oberti, F., Valsesia, E., Pilette, C., Rousselet, M. C., Bedossa, P., Aube, C., and et al. 1997. Noninvasive diagnosis of hepatic fibrosis or cirrhosis. Gastroenterology Vol. 113, pp. 1609–1616.

Patel, K., Gordon, S. C., Jacobson, I., Hezode, C., Oh, E., Smith, K. M., and et al. 2004. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. J Hepatol Vol. 41, pp. 935–942.

Pawlotsky, J. M. 2005. Current and future concepts in hepatitis c therapy. Semin Liver Dis Vol. 25, pp. 72–83.

Piccinino, F., Sagnelli, E., Pasquale, G., and Giusti, G. 1986. Complications following percutaneous liver biopsy. a multicentre retrospective study on 68.276 biopsies. Journal of Hepatology Vol. 2, pp. 165–173.

Rockey, D. C. and Bissell, D. M. Noninvasive measures of liver fibrosis. Hepatology Vol. 43, No. 2 Suppl 1, pp. S113–S120.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1987. Parallel distributed processing, volume 1 of Foundations. Ed. MIT Press, Cambridge, Massachusetts. D. E. Rumelhart, J. L. McClelland and the PDP group.

Sandrin, L., Fourquet, B., Hasquenoph, J. M., Yon, S., Fournier, C., Mal, F., and et al. 2003. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol Vol. 29, pp. 1705–1713.

Sebastiani, G. and Alberti, A. 2006. Noninvasive fibrosis biomarkers reduce but not substitute the need for liver biopsy. World J Gastroenterol Vol. 12, pp. 3682–3694.

Strader, D. B., Wright, T., Thomas, D. L., and Seeff, L. B. 2007. AALSD practice guideline: diagnosis, management, and treatment of hepatitis c. Hepatology Vol. 39, pp. 1147–1171.

Sud, A., Hui, J. M., Farrell, G. C., Bandara, P., Kench, J. G., Fung, C., and et al. 2004. Improved prediction of fibrosis in chronic hepatitis c using measures of insulin resistance in a probability index. Hepatology Vol. 39, pp. 1239–1247.

Takagi, T. and Sugeno, M. 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. on Systems, Man and Cybernetics Vol. 15, No. 1, pp. 116–132.

Thomas, D. L. and Seeff, L. B. 2005. Natural history of hepatitis c. Clin Liver Dis Vol. 9, pp. 383–398.

Wai, C. T., Greenson, J. K., Fontana, R. J., Kalbfleisch, J. D., Marrero, J. A., Conjeevaram, H. S., and Lok, A. S. 2003. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis c. Hepatology Vol. 38, pp. 518–526.

Wessels, L. and Barnard, E. 1992. Avoiding false local minima by proper initialization of connections. IEEE Trans. on Neural Networks Vol. 3, No. 6, pp. 899–905.

World Health Organization 2000. Hepatitis c – global prevalence. Wkly Epidemiol Rec Vol. 75, pp. 17–28.

Zadeh, L. A. 1965. Fuzzy sets. Information and Control Vol. 8, pp. 338–353.

Zadeh, L. A. 1996. Fuzzy logic = computing with words. IEEE Trans. on Fuzzy Systems Vol. 4, No. 2.

7. Responsibility notice

The author is the only responsible for the printed material included in this paper.