An Introduction to Parameterized IFAM Models with Applications in Prediction

Peter Sussner[1] Rodolfo Miyasaki[2] Marcos Eduardo Valle[3]

1.Department of Applied Mathematics, University of Campinas
Campinas, State of São Paulo, Brazil
2.Credit Planning and Management Department, Banco Itaú
São Paulo, Brazil
3.Department of Mathematics, University of Londrina
Londrina, Brazil
Email: sussner@ime.unicamp.br, rodolfo.miyasaki@itau.com.br, valle@uel.br

Abstract— *Fuzzy associative memories (FAMs) and, in particular, the class of implicative fuzzy associative memories (IFAMs) can be used to implement fuzzy rule-based systems. In this way, a variety of applications can be dealt with. Since there are infinitely many IFAM models, we are confronted with the problem of selecting the best IFAM model for a given application. In this paper, we restrict ourselves to a subclass of the entire class of IFAMs, namely the subclass of IFAMs that are associated with the Yager family of parameterized t-norms. For simplicity, we speak of the class of Yager IFAMs. In this setting, we formulate the problem of choosing the best Yager IFAM for a given application as an optimization problem. Considering two problems in time series prediction from the literature, we solve this optimization problem and compare the performance of the resulting Yager IFAM with the performances of other fuzzy, neural, neuro-fuzzy, and statistical techniques.*

Keywords— Fuzzy associative memory, implicative fuzzy associative memory, Yager family of parameterized t-norms, time-series prediction, hydroelectric plant, monthly streamflow prediction.

1 Introduction

Implicative fuzzy associative memories (IFAMs) belong to the class of *fuzzy morphological associative memories* (FMAMs) [1, 2]. The theoretical background for FMAMs can be found in *fuzzy mathematical morphology* (FMM) [3, 4]. More precisely, FMAMs can be viewed as fuzzy neural networks whose neurons perform elementary operations of fuzzy mathematical morphology [2].

Recently, Sussner and Valle [5] showed that many well-known fuzzy associative memory (FAM) models such as the FAMs of Kosko, Junbo et al., Liu, and Bělohlavék [6, 7, 8, 9] represent particular instances of FMAMs. In this paper, we focus on the FMAM subclass of IFAMs. Upon presentation of an input pattern $\mathbf{x}$, an IFAM model performs max-T products, where T is a t-norm, at every node. A certain continuous t-norm determines a particular IFAM model. For example, the Lukasiewicz IFAM is associated with the Lukasiewicz t-norm. Learning in IFAM models occurs by forming the R-implication of the underlying t-norm and by computing a matrix product that corresponds to the Bandler-Kohout superproduct [10, 11]. In the context of fuzzy associative memories, we speak of *implicative fuzzy learning*. As we have pointed out in previous articles, implicative fuzzy learning is based on the greatest solution of a fuzzy relational inequality [1, 12, 13]. In the general setting of FMAMs, implicative fuzzy learning generalizes to *learning by adjunction*, a recording strategy that was derived from a concept in mathematical morphology known as the duality relationship of adjunction [2, 14].

We have successfully applied IFAMs to some problems in time series forecasting, in particular a problem of forecasting the demand of man power in steel manufacturing industry in the state of West Bengal, India, and a problem of stream flow prediction in a hydroelectric plant in southeastern Brazil [5, 15]. In these prediction problems, the Lukasiewicz IFAM exhibited the best performance compared to a variety of other FAM models and to a number of statistical, neural, fuzzy, and neuro-fuzzy approaches from the literature. The goal of this paper is to optimize our results by considering not only a finite number of IFAMs but an infinite number of IFAM models.

To this end, we introduce the class of Yager IFAMs, a class of parameterized IFAM models depending on a single parameter. We simply define a Yager IFAMs as an IFAM that is associated with a Yager t-norm [16]. Recall that a Yager t-norm is of the form $T^d(x, y) = 1 - \{1 \wedge [(1-x)^d + (1-y)^d]^{1/d}\}$, where the symbol $\wedge$ denotes the minimum operation and where $d > 0$.

Thus, we can formulate an optimization problem depending on the parameter d that consists in minimizing the distance between the target pattern and the pattern produced by the IFAM corresponding to the parameter d. As a result of this optimization process, we obtain the Yager IFAM (or the set of Yager IFAMs) that provides the best fit with respect to the training data.

The paper is organized as follows. After presenting a brief introduction to IFAMs, we introduce the Yager class of parameterized IFAMs in Section 3. Subsequently, we discuss the aforementioned prediction problems and determine the Yager IFAMs that minimize the mean squared distance between the outputs produced by the memory cues $\mathbf{x}^\xi$ and the desired outputs. We finish the paper with some concluding remarks and suggestions for further research.

2 A Brief Introduction to Implicative Fuzzy Associative Memories

Implicative Fuzzy Associative Memories (IFAMs) can be defined in terms of certain matrix products, namely the max-T product and the min-I product, where T is a t-norm and I is a fuzzy implication [2]. We will make use of the following

ISBN: 978-989-95079-6-8

t-norms:

$$T_M(x,y) = x \wedge y\,, \qquad \text{(Minimum)} \qquad (1)$$
$$T_P(x,y) = x \cdot y\,, \qquad \text{(Product)} \qquad (2)$$
$$T_L(x,y) = 0 \vee (x+y-1)\,, \qquad \text{(Lukasiewicz)} \qquad (3)$$
$$T_W(x,y) = \begin{cases} x \wedge y, & x \vee y = 1\,, \\ 0, & x \vee y < 1\,. \end{cases} \qquad \text{(Nilpotent Min.)} \qquad (4)$$

Reverse fuzzy implications can be easily derived from fuzzy implications. Recall the following popular implications:

$$I_M(x,y) = \begin{cases} 1, & x \leq y \\ y, & x > y \end{cases}\,, \qquad \text{(Gödel)} \qquad (5)$$
$$I_P(x,y) = \begin{cases} 1, & x \leq y \\ y/x, & x > y \end{cases}\,, \qquad \text{(Product)} \qquad (6)$$
$$I_L(x,y) = 1 \wedge (y-x+1)\,. \qquad \text{(Lukasiewicz)} \qquad (7)$$

Let $A \in [0,1]^{m \times p}$ and $B \in [0,1]^{p \times n}$. The max-T product of A and B is given by the matrix $C = A \circ B$ and the min-I product of A and B is given by the matrix $D = A \circledast B$, where C and D are defined as follows (note that the min-I product of A and B can also be viewed as a Bandler-Kohout superproduct [10, 11]).

$$c_{ij} = \bigvee_{k=1}^{p} C(a_{ik}, b_{kj}) \text{ and } d_{ij} = \bigwedge_{k=1}^{p} I(b_{kj}, a_{ik})\,. \qquad (8)$$

A FAM is a fuzzy neural network that is designed to store a fundamental memory set, i.e., a set of associations $\left\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, \ldots, p\right\}$, where $\mathbf{x}^\xi \in [0,1]^n$ and $\mathbf{y}^\xi \in [0,1]^m$. This task can be achieved by means of a synaptic weight matrix $W \in [0,1]^{m \times n}$. Let $X = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^p] \in [0,1]^{n \times p}$ denote the matrix whose columns are the input patterns and let $Y = [\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^p] \in [0,1]^{m \times p}$ denote the matrix whose columns are the output patterns. We suggested the following rule for synthesizing the weight matrix W of an IFAM [1, 12, 13]:

$$W = Y \circledast X^T\,. \qquad (9)$$

Note that (9) depends on the choice of a fuzzy implication. In an IFAM model, we require that the underlying fuzzy implication I is the R-implication of a continuous t-norm T. In other words, the corresponding fuzzy implication I must be adjoint to a continuous t-norm T [2, 4]. For example, the pairs (I_M, T_M), (I_P, T_P), and (I_L, T_L) represent pairs of adjoint operators.

There are infinitely many IFAM models. In the special cases where the fuzzy implications occurring in (9) are I_M, I_P, and I_L, we speak of the Gödel IFAM, the Goguen IFAM, and the Lukasiewicz IFAM, respectively.

Once the recording phase has been completed, the IFAM weight matrix can be applied to an arbitrary input pattern $\mathbf{x} \in [0,1]^n$. If θ denotes a threshold or bias vector that is given by the entry-wise minimum over all $\mathbf{y}^\xi$, where $\xi = 1, \ldots, p$, then we obtain the following output pattern $\mathbf{y} \in [0,1]^m$:

$$\mathbf{y} = (W \circ \mathbf{x}) \vee \boldsymbol{\theta}\,, \qquad \text{where} \qquad \boldsymbol{\theta} = \bigwedge_{\xi=1}^{p} \mathbf{y}^\xi\,. \qquad (10)$$

Here, the t-norm T occurring in the max-T product is such that its R-implication I is associated with the product "$\circledast$" of (9). In other words, the t-norm T of (10) and the fuzzy implication I of (9) are adjoint [2, 4].

3 The Yager Class of Parameterized IFAMs

Note that a particular IFAM model is determined by a continuous t-norm. An entire subclass of IFAM models is given by the class of Yager t-norms or intersections which are obviously continuous. Recall that the class of Yager t-norms represents a family of parameterized t-norms given by the following equation [16]:

$$T^d(x,y) = 1 - \left\{1 \wedge [(1-x)^d + (1-y)^d]^{1/d}\right\},\ d > 0\,. \qquad (11)$$

For $d = 1$, the Yager t-norm coincides with the Lukasiewicz t-norm T_L. For $t \to 0$ the Yager t-norm approaches the nilpotent minimum T_W, which is the pointwise smallest t-norm. For $t \to \infty$ the Yager t-norm approaches the minimum t-norm T_M, the pointwise largest t-norm [17, 18].

For $d > 0$, the symbol I^d denotes the R-implication of T^d. The operator I^d can be computed as follows:

$$I^d(x,y) = 1 - \left\{0 \vee [(1-y)^d - (1-x)^d]\right\}^{1/d},\ d > 0\,. \qquad (12)$$

The resulting IFAMs are called Yager IFAMs. These parameterized IFAM models will be considered in the applications of the next section. In fact, we will tackle the problem of choosing the best Yager IFAM for a given application in terms of an optimization problem.

To this end, let us consider the fundamental memory set $\left\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, \ldots, p\right\}$ which can be viewed as the set of training patterns. Let W^d denote the synaptic weight matrix of the Yager IFAM with parameter $d > 0$. A possible approach towards finding the best Yager IFAM consists in minimizing the distances between $(W^d \circ^d \mathbf{x}^\xi) \vee \theta$ and $\mathbf{y}^\xi$ for $\xi = 1, \ldots, p$, where the symbol "$\circ^d$" denotes the max-T^d product. This optimization problem can be formulated as follows:

$$\begin{cases} \text{minimize} & \sum_{\xi=1}^{p} \left\| \mathbf{y}^\xi - [(W^d \circ^d \mathbf{x}^\xi) \vee \theta] \right\|_2 \\ \text{subject to} & d > 0 \end{cases}\,. \qquad (13)$$

If $\mathbf{y}^{\xi,d} = (W \circ^d \mathbf{x}^\xi) \vee \theta$ and Y^d is the matrix $Y^d = [\mathbf{y}^{1,d}, \ldots \mathbf{y}^{p,d}]$ then the optimization problem of (13) can be expressed in terms of the Frobenius norm as follows:

$$\begin{cases} \text{minimize} & \left\| Y - Y^d \right\|_F \\ \text{subject to} & d > 0 \end{cases}\,. \qquad (14)$$

In the applications as fuzzy rule-based systems that are described below, an input fuzzy set $\mathbf{x} \in [0,1]^n$ is derived from a real-valued input vector $\mathbf{v}$ and the respective output fuzzy set $\mathbf{y} = (W^d \circ \mathbf{x}) \vee \theta$ is defuzzified yielding a real-valued output $s = defuzz(\mathbf{y})$. The training patterns are given in the form $(\mathbf{v}^1, s^1), \ldots, (\mathbf{v}^p, s^p)$. Let $s^{\xi,d}$ denote the defuzzification of $\mathbf{y}^{\xi,d}$, i.e. $s^{\xi,d} = defuzz(\mathbf{y}^{\xi,d})$ for all $\xi = 1, \ldots, p$. If we have $\mathbf{s} = (s^1, \ldots, s^p)^T$ and $\mathbf{s}^d = (s^{1,d}, \ldots, s^{p,d})^T$ then we choose to minimize the following expression instead of (14):

ISBN: 978-989-95079-6-8

$$\begin{cases} \text{minimize} & \left\|\mathbf{s} - \mathbf{s}^d\right\|_2 \\ \text{subject to} & d > 0 \end{cases} . \quad (15)$$

In other words, we minimize the Euclidean distance between the desired results $\mathbf{s}$ and the results produced by the Yager IFAMs for $d > 0$. To this end, we applied the routine FMINBND of MATLAB's Optimization Toolbox in order to determine a local minimum of the real-valued objective function $f(d) = \left\|\mathbf{s} - \mathbf{s}^d\right\|_2$ in an interval. Note that plotting the objective function $f : \mathbb{R} \to \mathbb{R}$ allows us to select an interval $[x_1, x_2]$ such that an application of FMINBND yields a candidate for a global minimum. The resulting parameter d yields the Yager IFAM that exhibits the best performance on the training data. The next sections describes applications of this strategy to two forecasting problems that can be found in the literature.

4 Applications of Yager IFAMs in Prediction

Fuzzy associative memories such as IFAMs can be used to implement mappings of fuzzy rules. In this case, a set of rules in the form of human-like IF-THEN conditional statements are stored. In this subsection, we consider two problems in time-series prediction. We use (15) to select the Yager IFAM that produces the least MSE on the training data.

4.1 Prediction of Manpower Requirement in Steel Manufacturing

Let us consider the problem of predicting the manpower requirement in steel manufacturing industry in the state of West Bengal, India [19]. Initially, we have five linguistic values A_i, $i = 1, \ldots, 5$ and a set of fuzzy conditional statements such as "If the manpower of year n is A_i, then that of year $n+1$ is A_j". The linguistic values A_i correspond to fuzzy sets. Table 1 shows the set of input-output pairs that we stored in a number of different FAM models including the Gödel, the Goguen, the Lukasiewicz, and the Yager IFAMs with $d > 0$.

If W is the synaptic weight matrix of an IFAM model and $\boldsymbol{\theta}$ is the threshold obtained after the learning process, then the predicted manpower of year $n+1$ is given by the following equation:

$$A_{n+1} = (W \circ A_n) \vee \boldsymbol{\theta} , \quad (16)$$

where A_n is the manpower of year n and $\circ$ is the max-t composition. Note that here we have $\boldsymbol{\theta} = (0,0,0,0,0)^T$ and therefore

$$A_{n+1} = W \circ A_n . \quad (17)$$

The fuzzy set A_n is computed by fuzzifying the numerical input value v corresponding to the manpower requirement of year n according to the method described by Choudhury et. al. [19]. We calculated the predicted value s of the required manpower for year $n+1$ by applying (17) to A_n and by defuzzifying the result A_{n+1} using the "mean of maxima" (MOM) scheme.

Choudhury et. al. used this problem to compare the average or mean percentage error (MPE) produced by Kosko's FAM [6, 20] and by the statistical methods ARIMA1 and ARIMA2 [21, 22]. In a recent paper [1], we included some IFAM models, in particular the Lukasiewicz IFAM, as well as the Lukasiewicz Generalized FAM (GFAM) of Chung and Lee [23] and the max-min FAM with threshold of Liu [8] in this comparison. The Lukasiewcz IFAM, which is closely related to the original morphological associative memory (MAM) [1, 24, 25], exhibited the best performance among all the models mentioned above. Recall that the Lukasiewicz IFAM coincides with the Yager IFAM for $d = 1$.

Table 1: Set of input and output pairs used in the forecasting application

ξ	$\mathbf{x}^\xi$	$\mathbf{y}^\xi$
1	$[1.0, 0.5, 0, 0, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
2	$[0.5, 1.0, 0.5, 0, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
3	$[0.5, 1.0, 0.5, 0, 0]^T$	$[0, 0.5, 1.0, 0.5, 0]^T$
4	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
5	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0, 0.5, 1.0, 0.5, 0]^T$
6	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0, 0, 0.5, 1.0, 0.5]^T$
7	$[0, 0, 0.5, 1.0, 0.5]^T$	$[0, 0, 0.5, 1.0, 0.5]^T$
8	$[0, 0, 0.5, 1.0, 0.5]^T$	$[0, 0, 0, 0.5, 1.0]^T$
9	$[0, 0, 0, 0.5, 1.0]^T$	$[0, 0, 0, 0.5, 1.0]^T$

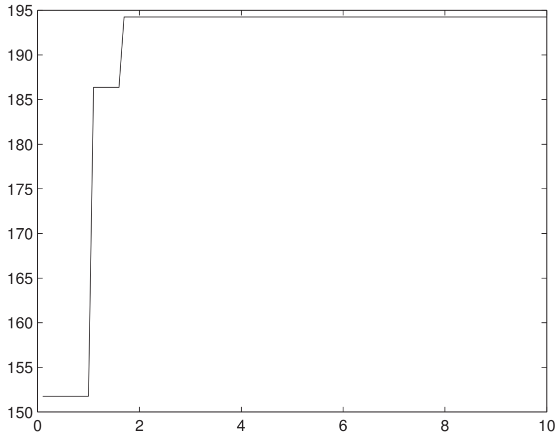

Figure 1: The distance $\left\|\mathbf{s} - \mathbf{s}^d\right\|_2$ between the actual demand of manpower between 1984 and 1995 and the demand of manpower predicted by the Yager IFAM as a function of d.

In this paper, we determine the best Yager IFAM for this prediction problem by solving the optimization problem of (15). Since the Lukasiewicz IFAM is taken into account in the optimization process, the IFAMs representing solutions of (15) are guaranteed to perform at least as well as the Lukasiewicz IFAM with respect to the MSE. In fact, the minimum mean squared error (MSE) is adopted for any value of d in the interval $(0, 1]$ as shown in Figure 1.

Table 2 presents a comparison of the FAM models and the statistical autoregressive integrated moving average (ARIMA) methods in terms of the MSE, the MPE, and the mean absolute error (MAE). Figure 2 provides for a graphical interpretation of the predictions.

The results of this experiment indicate the utility of IFAMs for prediction problems and the validity of our approach for determining the best Yager IFAM. However - as the reader may have noticed - this prediction problem does not include

ISBN: 978-989-95079-6-8

Table 2: MSE, MAE and MPE produced by the prediction models

Method	MSE ($\times 10^5$)	MAE (m^3/s)	MPE (%)
Yager IFAM ($d \in (0,1]$):	1.92	32.75	2.29
Lukasiewicz IFAM:	1.92	32.75	2.29
Kosko's FAM:	2.57	38.75	2.67
Lukasiewicz GFAM:	2.57	38.75	2.67
Gödel IFAM:	2.89	38.58	2.73
Max-min FAM/threshold:	2.89	38.58	2.73
Goguen IFAM:	3.14	42.75	2.99
ARIMA2	9.36	83.55	5.48
ARIMA1	23.26	145.25	9.79

any test data that would allow us to assess the performance of the models for years beyond the year of 1995. More realistic problems such as the one described below consist of training data and test data. In this case, the optimization should be performed on the training data and the test data should serve to validate this approach (possibly, separate validation data can be extracted from the training data and the validation data can be used to choose the best model).

4.2 *Prediction of Average Monthly Streamflow of a Hydroelectric Plant*

Furnas is a large hydroelectric plant that is located in southeastern Brazil. Magalhaẽs et al. as well as Sussner and Valle have previously discussed the problem of forecasting the average monthly streamflow [26, 15]. The seasonality of the monthly streamflow suggests the use of 12 different models, one for each month of the year. The training data correspond to the years $1931-1990$ and the test data correspond to the years $1991-1998$.

Let s^ξ, where $\xi = 1, \ldots, p$, be samples of a seasonal streamflow time series. The goal is to estimate the value of s^γ from a subsequence of $(s^1, s^2, \ldots, s^{\gamma-1})$. Here, we employ subsequences that correspond to a vector of the form

$$\mathbf{v}^\gamma = (s_{\gamma-h}, \ldots, s_{\gamma-1})^T, \qquad (18)$$

where $h \in \{1, 2, \ldots, \gamma-1\}$. In this experiment, we employed a fixed number of three antecedents in our IFAM models. For example, only the values of January, February, and March were taken into account to predict the streamflow of April.

The uncertainty that is inherent in hydrological data suggests the use of fuzzy sets to model the streamflow samples. For $\xi < \gamma$, a fuzzification of $\mathbf{v}^\xi$ and s^ξ using Gaussian membership functions yields fuzzy sets $\mathbf{x}^\xi : \mathcal{U} \to [0,1]$ and $\mathbf{y}^\xi : \mathcal{V} \to [0,1]$, respectively, where $\mathcal{U}$ and $\mathcal{V}$ represent finite universes of discourse. A subset S of the resulting input-output pairs $(\mathbf{x}^\xi, \mathbf{y}^\xi)$, where $\xi = 1, \ldots, p$, is implicitly stored in an IFAM model (we only construct the parts of the weight matrix that are actually used in the recall phase). We employed the *subtractive clustering method* to determine the set S [27]. The centers and the widths of the Gaussian-type membership functions of the input patterns $\mathbf{x}^\xi$ and the output patterns $\mathbf{y}^\xi$ in S were estimated by means of the MATLAB function `subclust`. Here, we used a constant width of $r = 0.5$.

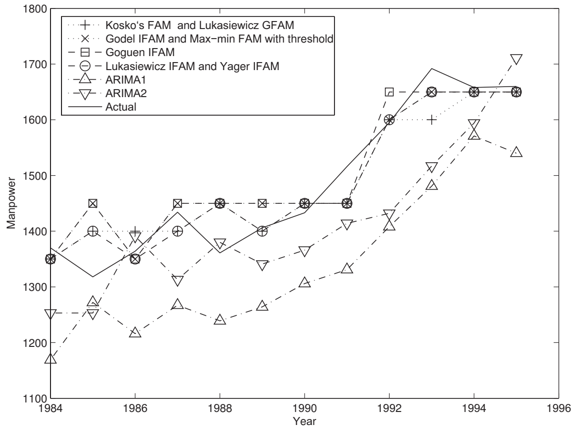

Figure 2: Comparison in forecasting manpower. The continuous line represents the actual manpower. The dashed line marked by '○' corresponds to the Lukasiewicz IFAM model, i.e., the Yager IFAM for $d = 1$ and the dashed line marked by '□' corresponds to the Goguen IFAM. The dotted line marked by '×' refers to Kosko's FAM model as well as the Lukasiewicz Generalized FAM, and the dotted line marked by '+' refers to the max-min FAM with threshold and the Gódel IFAM. The lines marked by '△' and '▽' represent ARIMA1 and ARIMA2.

Upon presentation of the input pattern $\mathbf{x}^\gamma$, the IFAM with parameter d yields the corresponding output pattern $\mathbf{y}^{\gamma,d}$. For computational reasons, $\mathbf{x}^\gamma$ is modeled as a discrete Dirac-δ (impulse) function. A defuzzification of $\mathbf{y}^{\gamma,d}$ using the centroid method yields $s^{\gamma,d}$.

As before, we generated the vectors $\mathbf{s} = (s^1, \ldots, s^p)^T$ and $\mathbf{s}^d = (s^{1,d}, \ldots, s^{p,d})^T$. We employed the MATLAB function FMINBND to solve the optimization problem given by (15). This optimization process resulted in the parameter $d = 4.3568$. Figure 3 depicts $\left\|\mathbf{s} - \mathbf{s}^d\right\|_2$ for $d \in (0, 10]$. For the training data concerning the years $1931-1990$, the MSE of the Yager IFAM with $d = 4.3568$ is 12087.3 whereas the MSE for the Lukasiewicz IFAM, i.e., the Yager IFAM with $d = 1$, is 12387.4.

Table 3 provides the MSEs, the MPEs, and the MAEs produced by some IFAMs and other models during the testing phase. The values of the periodic auto-regressive moving average model (PARMA) [21], the fuzzy neural network NFN [28], and the predictive fuzzy clustering method FPM-PRP were drawn from the literature [26].

Figure 4 shows the forecasted streamflows estimated by the prediction model based on the FMAM for the Furnas reservoir from 1991 to 1998. Table 3 compares the errors that were generated by the IFAMs and several other models [26]. In contrast to the IFAM models, the MLP, NFN, and FPM-PRP models were initialized by optimizing the number of the parameters for each monthly prediction. For example, the MLP considers 4 antecedents to predict the streamflow of January and 3 antecedents to predict the streamflow for February. Moreover, the FPM-PRP model also takes advantage of slope information which requires some additional "fine tuning". Nevertheless, the Yager IFAM resulting from the min-

ISBN: 978-989-95079-6-8

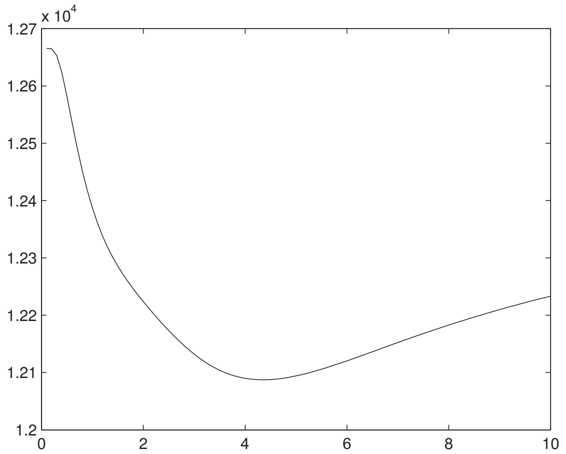

Figure 3: The distance $\left\|\mathbf{s}-\mathbf{s}^d\right\|_2$ between the actual streamflows given by the training data and the streamflows predicted by the Yager IFAM as a function of d.

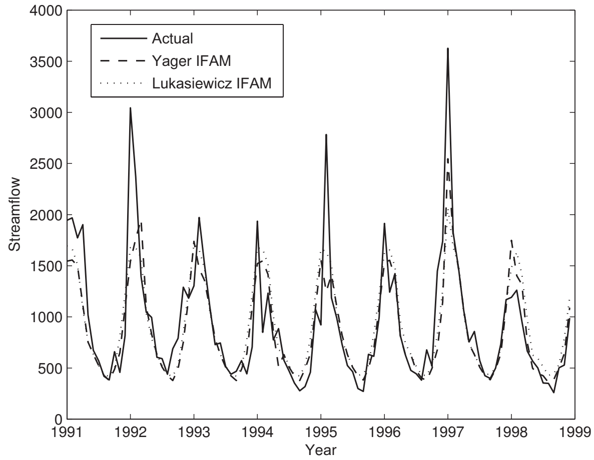

Figure 4: The streamflow prediction for the Furnas reservoir from 1991 to 1998. The continuous line corresponds to the actual values and the dashed line corresponds to the prediction of the Yager IFAM with $d = 4.3568$. The dotted line refers to the values predicted by the Lukasiewcz IFAM.

Table 3: Mean square, mean absolute, and mean relative percentage errors produced by the prediction models.

Method	MSE ($\times 10^5$)	MAE (m^3/s)	MPE (%)
FPM-PRP	1.20	200	18
Yager IFAM ($d = 4.3568$)	1.28	216	21
Lukasiewicz IFAM	1.27	229	24
PARMA	1.85	280	28
MLP	1.82	271	30
NFN	1.73	234	20

imization performed in (15) and the Lukasiewicz IFAM produced very satisfactory predictions that are visualized in Figure 4. Note that in the testing phase the MSE produced by the Yager IFAM with $d = 4.3568$ is slightly higher than the one produced by the Lukasiewicz IFAM although the Yager IFAM with $d = 4.3568$ outperforms the Lukasiewicz IFAM in terms of MAE and MPE. This fact may be due to overfitting to the training data. With some fine tuning of the parameters, it is possible to experimentally determine a Lukasiewicz IFAM (and thus a Yager IFAM) that outperforms the FPM-PRP with respect to the three error measures MSE, MAE, and MPE [15]. In our opinion, however, the fine tuning of the parameters should preferably not be performed experimentally but should be part of the optimization process.

5 Concluding Remarks

This paper represents the first attempt of tackling the problem of selecting the best IFAM for a given application. To this end, we introduced parameterized IFAM models, specifically the Yager class of IFAMs, and we formulated the problem of determining the Yager IFAM that yields the best fit to the training data as an optimization problem.

Recall that an IFAM is uniquely determined by a continuous t-norm, such as a Yager t-norm T^d with parameter d for $0 < d < \infty$, which gives rise to an R-implication. Although other continuous parameterized t-norms could have been chosen [29], the Yager t-norms T^d have the advantage that T^d approaches T_W, the pointwise smallest t-norm, for $d \to 0$ and T^d approaches T_M, the pointwise largest t-norm, for $d \to \infty$. Moreover, T^1 equals the Lukasiewicz t-norm which implies that the Lukasiewicz IFAM, i.e., the IFAM that exhibited the best performance in our previous experiments, is taken into account in the optimization process.

We applied our approach to two prediction problems from the literature. In the first problem concerning the prediction of the manpower requirement in steel manufacturing industry in West Bengal, India, only training data are available. The minimization of the MSE for the Yager IFAMs resulted in an infinite number of Yager IFAMs corresponding to $d \in (0, 1]$ that exhibit the same errors on the training data as the Lukasiewicz IFAM ($d = 1$). For the streamflow prediction problem in a major Brazilian hydroelectric plant, we dispose of training data and test data. The optimization process generated a Yager IFAM that outperforms the Lukasiewicz IFAM with respect to the MSE on the training data but exhibits a slightly higher MSE on the test data. This result may be due to overfitting and in fact we know that better Yager IFAMs exist for this problem. Nevertheless, the particular Yager IFAM derived from the optimization process exhibits very satisfactory results (which are better than those of the Lukasiewicz IFAM with respect to the MAE and MPE error measures), especially in view of the fact that we used a fixed number of antecedents and other parameters in conjunction with the Yager IFAM.

In the near future, we intend to incorporate the selection of additional parameters into the optimization process and we intend to use validation and/or regularization techniques to avoid overfitting. Furthermore, we plan to investigate other classes of parameterized IFAMs that are based on other types of parameterized t-norms.

ISBN: 978-989-95079-6-8

References

[1] P. Sussner and M. Valle. Implicative fuzzy associative memories. *IEEE Transactions on Fuzzy Systems*, 14(6):793–807, 2006.

[2] M. Valle and P. Sussner. A general framework for fuzzy morphological associative memories. *Fuzzy Sets and Systems*, 159(7):747–768, 2008.

[3] M. Nachtegael and E. Kerre. Connections between binary, grayscale and fuzzy mathematical morphologies. *Fuzzy Sets and Systems*, 124(1):73–85, Nov. 2001.

[4] P. Sussner and M. Valle. Classification of fuzzy mathematical morphologies based on concepts of inclusion measure and duality. *Journal of Mathematical Imaging and Vision*, 32(2):139–159, 2008.

[5] P. Sussner and M. Valle. Fuzzy associative memories and their relationship to mathematical morphology. In W. Pedrycz, A. Skowron, and V. Kreinovich, editors, *Handbook of Granular Computing*, chapter 33. John Wiley and Sons, Inc., New York, 2008.

[6] B. Kosko. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, N.J., 1992.

[7] F. Junbo, J. Fan, and S. Yan. A learning rule for fuzzy associative memories. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, volume 7, pages 4273 – 4277, June 1994.

[8] P. Liu. The fuzzy associative memory of max-min fuzzy neural networks with threshold. *Fuzzy Sets and Systems*, 107:147–157, 1999.

[9] R. Belohlávek. Fuzzy logical bidirectional associative memory. *Information Sciences*, 128(1-2):91–103, 2000.

[10] W. Bandler and L. Kohout. Fuzzy relational products and fuzzy implication operators. In *International Workshop on Fuzzy Reasoning Theory and Applications*, Queen Mary College, University of London, London, 1978.

[11] L. Běhounek and M. Daňková. Relational compositions in fuzzy class theory. *Fuzzy Sets and Systems*, 160(8):1005–1036, April 2009.

[12] J. Drewniak. Fuzzy relation equations and inequalities. *Fuzzy Sets and Systems*, 14(3):237–247, December 1984.

[13] M. Valle, P. Sussner, and F. Gomide. Introduction to implicative fuzzy associative memories. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pages 925 – 931, Budapest, Hungary, July 2004.

[14] H. Heijmans. *Morphological Image Operators*. Academic Press, New York, NY, 1994.

[15] P. Sussner and M. Valle. Morphological and certain fuzzy morphological associative memories for classification and prediction. In V. Kaburlassos and G. Ritter, editors, *Computational Intelligence Based on Lattice Theory*, volume 67, pages 149 – 173. Springer Verlag, Heidelberg, Germany, 2007.

[16] R. Yager. On a general class of fuzzy connectives. *Fuzzy Sets and Systems*, 4(3):235–242, 1980.

[17] G. Klir and T. Folger. *Fuzzy Sets, Uncertainty and Information*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[18] C. Lin and C. Lee. *Neural Fuzzy Systems*. Prentice Hall, Upper Saddle River, NJ, 1996.

[19] J. Choudhury, B. Sarkar, and S. Mukherjee. Forecasting of engineering manpower through fuzzy associative memory neural network with arima: a comparative study. *Neurocomputing*, 47:241–257, 2002.

[20] B. Kosko. *Fuzzy Engineering*. Prentice Hall, Upper Saddle River, N.J., 1992.

[21] J. Box, G. Jenkins, and C. Reinsel. *Time Series Analysis, Forecasting, and Control*. Holden-Day, Oakland, CA, third edition, 1994.

[22] T. Mills. *Time Series Techniques for Economists*. Cambridge University Press, Cambridge, United Kingdom, 1990.

[23] F. Chung and T. Lee. On fuzzy associative memory with multiple-rule storage capacity. *IEEE Transactions on Fuzzy Systems*, 4(3):375–384, August 1996.

[24] G. Ritter, P. Sussner, and J. Diaz de Leon. Morphological associative memories. *IEEE Transactions on Neural Networks*, 9(2):281–293, 1998.

[25] P. Sussner and M. Valle. Grayscale morphological associative memories. *IEEE Transactions on Neural Networks*, 17(3):559–570, May 2006.

[26] M. Magalhães, R. Ballini, R. Gonçalves, and F. Gomide. Predictive fuzzy clustering model for natural streamflow forecasting. In *Proceedings of the IEEE International Conference on Fuzzy Systems*, pages 390–394, Budapest, Hungary, 2004.

[27] S. Chiu. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, 1994.

[28] M. Figueiredo and F. Gomide. Adaptive neuro-fuzzy modeling. *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, 3:1567–1572, 1997.

[29] E. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, 2001.

ISBN: 978-989-95079-6-8