Modeling Failure Rate Time Series by a Fuzzy Arithmetic-based Inference System

Tamás Jónás and Zsuzsanna Eszter Tóth
Department of Management and Corporate Economics
Budapest University of Technology and Economics
Magyar tudósok körútja 2, 1117 Budapest, Hungary

József Dombi
Department of Computer Algorithms and Artificial Intelligence
University of Szeged
Árpád tér 2, 6720 Szeged, Hungary

Abstract—In this study, a fuzzy arithmetic based inference system is introduced to model and forecast linear trends of empirical failure rate time series. Here, a simple heuristic is introduced to form the membership functions of the fuzzy rule antecedents, while each rule consequent is treated as a fuzzy number composed of a left hand side and a right hand side fuzzy set, each of which is given by a sigmoid membership function. The novelty of the proposed method lies in the application of pliant arithmetics to aggregate separately the left hand sides and the right hand sides of the individual fuzzy consequents, taking the activation levels of the corresponding antecedents into account. Here, Dombi’s conjunction operator is applied to form the fuzzy output from the aggregates of the left hand side and right hand side sigmoid functions. The introduced defuzzification method does not require any numerical integration and its speed is independent of the number of fuzzy rules. The output of the pliant arithmetic based fuzzy inference system is used to predict linear trends of failure rate time series. Next, the modeling capability of the introduced methodology is compared to that of an Adaptive Neuro-Fuzzy Inference System. Based on the results, our method may be viewed as a viable alternative modeling and prediction technique.

I. INTRODUCTION

In reliability theory, the quantity \( h(t) \Delta t \) is known as the conditional probability that a component or a product will fail in the time interval \( (t, t + \Delta t) \), given that it has survived until time \( t \). Function \( h(t) \) is called the failure rate function or hazard function that can be estimated by

\[
h(t) = \frac{N(t) - N(t + \Delta t)}{N(t) \Delta t},
\]

where \( N(t) \) is the number of components or products that have survived until time \( t \) from the number of products or components that were put into operation. If \( \Delta t = 1 \) and \( t \) is taken in discrete times then the estimated failure rate \( \omega_i \) for period \( i \) may be give by

\[
\omega_i = \frac{N(i\Delta t) - N((i + 1)\Delta t)}{N(i\Delta t)} = \frac{N(i) - N(i + 1)}{N(i)}, \quad i = 1, 2, \ldots, M,
\]

where \( N(t) \) is the number of components or products that have survived until time \( t \) from the number of products or components that were put into operation. If \( \Delta t = 1 \) and \( t \) is taken in discrete times then the estimated failure rate \( \omega_i \) for period \( i \) may be give by

\[
\omega_i = \frac{N(i\Delta t) - N((i + 1)\Delta t)}{N(i\Delta t)} = \frac{N(i) - N(i + 1)}{N(i)}, \quad i = 1, 2, \ldots, M,
\]

where \( \omega_i \) is the failure rate at time \( t_i \). The novelty of the proposed method lies in the application of pliant arithmetics to aggregate separately the left hand sides and the right hand sides of the individual fuzzy consequents, taking the activation levels of the corresponding antecedents into account. Here, Dombi’s conjunction operator is applied to form the fuzzy output from the aggregates of the left hand side and right hand side sigmoid functions. The introduced defuzzification method does not require any numerical integration and its speed is independent of the number of fuzzy rules. The output of the pliant arithmetic based fuzzy inference system is used to predict linear trends of failure rate time series. Next, the modeling capability of the introduced methodology is compared to that of an Adaptive Neuro-Fuzzy Inference System. Based on the results, our method may be viewed as a viable alternative modeling and prediction technique.

II. THE LINE SEGMENTS REPRESENTATION OF EMPirical FAILURE RATE TIME SERIES

Let \( \omega_1, \omega_2, \ldots, \omega_M \) be an empirical failure rate time series and let the window span \( u \) be a positive integer, \( u \geq 2 \). Now, we define the \( u \)-period long, normalized, left hand side and right hand side sub-time-series \( \omega_{l,1}^{(i)}, \omega_{l,2}^{(i)}, \ldots, \omega_{l,u}^{(i)} \) and \( \omega_{r,1}^{(i)}, \omega_{r,2}^{(i)}, \ldots, \omega_{r,u}^{(i)} \), respectively:

\[
\omega_{l,p}^{(i)} = \begin{cases} 
\frac{\omega_{(i-1)u+p} - \omega_{i}}{s_i}, & \text{if } s_i > 0 \\
\frac{\omega_{(i-1)u+p} - \omega_{i}}{s_i}, & \text{if } s_i = 0 
\end{cases},
\]

\[
\omega_{r,p}^{(i)} = \begin{cases} 
\omega_{i+u+p} - \omega_{i}, & \text{if } s_i > 0 \\
\omega_{i+u+p} - \omega_{i}, & \text{if } s_i = 0 
\end{cases},
\]

where \( i = 1, 2, \ldots, \lfloor M/u \rfloor - 1 \), \( p = 1, 2, \ldots, u \). \( \overline{\omega}_i \) and \( s_i \) in (3) are the average and standard deviation of
\[ \omega_{i,1}, \omega_{i,2}, \ldots, \omega_{i,n}, \] respectively. Here, we fit trend lines both to the left hand side and right hand side normalized sub-time-series. Let \( \alpha^{(l)}_i \) and \( \beta^{(l)}_i \) denote the slope and intersection of the line fitted to the time series \( \omega_{i,1}, \omega_{i,2}, \ldots, \omega_{i,2} \), respectively. Similarly, let \( \alpha^{(r)}_i \) and \( \beta^{(r)}_i \) denote the slope and intersection of the line fitted to the time series \( \omega^{(r)}_{i,1}, \omega^{(r)}_{i,2}, \ldots, \omega^{(r)}_{i,n} \), respectively. If \( s_i = 0 \), then \( \alpha^{(l)}_i = \alpha^{(r)}_i = 0, \beta^{(l)}_i = \omega^{(l)}_{i,1}, \beta^{(r)}_i = \omega^{(r)}_{i,1} \). If \( s_i > 0 \), then \( \alpha^{(l)}_i, \beta^{(l)}_i, \alpha^{(r)}_i \) and \( \beta^{(r)}_i \) are identified so that
\[
\begin{align*}
\sum_{t=1}^{n} \left( \alpha^{(l)}_i t + \beta^{(l)}_i - \omega^{(l)}_{i,t} \right)^2 &\rightarrow \min, \\
\sum_{t=1}^{n} \left( \alpha^{(r)}_i t + \beta^{(r)}_i - \omega^{(r)}_{i,t} \right)^2 &\rightarrow \min.
\end{align*}
\]
We call the series of vector pairs
\[
\left( (\alpha^{(l)}_i, \beta^{(l)}_i, s_i), (\alpha^{(r)}_i, \beta^{(r)}_i) \right)
\]
the normalized line segments model of the failure rate time series \( \omega_{1,1}, \omega_{1,2}, \ldots, \omega_{M} \). Figure 1 shows how the normalized line segments model, after its denormalization, can represent a failure rate time series. Now, we will introduce a fuzzy arithmetic-based inference system that can learn the relations
\[
\begin{align*}
(\alpha^{(l)}_i, \beta^{(l)}_i, s_i) &\Rightarrow \alpha^{(r)}_i, \\
(\alpha^{(r)}_i, \beta^{(r)}_i, s_i) &\Rightarrow \beta^{(r)}_i.
\end{align*}
\]

III. CONSTRUCTION OF A FUZZY INFERENCE SYSTEM

From the line segment model of the failure rate time series \( \omega_{1,1}, \omega_{1,2}, \ldots, \omega_{M} \) we have the input vectors
\[
x_i = (x_{i,1}, x_{i,2}, x_{i,3}) = (\alpha^{(l)}_i, \beta^{(l)}_i, s_i)
\]
and the outputs
\[
y^{(l)}_i = \alpha^{(r)}_i, \\
y^{(r)}_i = \beta^{(r)}_i.
\]

\[ i = 1, 2, \ldots, d, \] where \( d = [M/u] - 1 \) is the number of input vectors. We will construct a separate fuzzy inference system for each output. For a simpler notation during the fuzzy modeling, we will use \( y_i \) to denote the scalar output that belongs to input vector \( x_i \). We will introduce our modeling in a generic way, thus we will use \( n \) to denote the number of dimensions of the input vectors.

A. Membership Functions of the Fuzzy Rule Antecedents

Each component of the input vector \( x \) is considered being a linguistic variable. Let \( k_j \) denote the number of linguistic values (levels) for the \( j \)th dimension of \( x \), and let \( A_{j,h} \) be a fuzzy set that represents the \( h \)th linguistic value in the \( j \)th dimension of the input space \( (k_j \geq 3) \). Furthermore, let \( \mu_{A_{i,h}} \) denote the membership function of fuzzy set \( A_{j,h} \); that is, \( \mu_{A_{j,h}}(x) \) is the membership value of variable \( x \) in the fuzzy set \( A_{i,h} \). For \( j = 1, 2, \ldots, n; h = 1, 2, \ldots, k_j \), we construct the membership function \( \mu_{A_{i,h}}(x) \) as follows. Let \( q^{(j)}_i \) be the \( t/(k_j - 1) \) quantile of set \( X_j \), where \( X_j = \{ x_{i,j} \} \)
\[
i = 1, 2, \ldots, d,\ t = 0, 1, 2, \ldots, k_j - 1,
\]
\[
q^{(j)}_0 = \min_{i \in \{1, 2, \ldots, n\}} (x_{i,j}), \\
q^{(j)}_{k_j - 1} = \max_{i \in \{1, 2, \ldots, n\}} (x_{i,j}).
\]

Regarding the membership function \( \mu_{A_{i,h}}(x) \), we differentiate among three cases. If \( h = 1 \) (case a), we construct the membership function for the lowest linguistic level, if \( h = k_j \) (case c), we define the membership function of the fuzzy set that represents the highest linguistic level, and if \( 1 < h < k_j \) (case b), we give the membership functions for all the other linguistic levels. In our constructions, we use membership functions that are derived from the sigmoid function.

Definition 1: The sigmoid function \( \sigma^{(\lambda)}_a \) with parameter \( a \) and \( \lambda \) is given by
\[
\sigma^{(\lambda)}_a(x) = \frac{1}{1 + e^{-\lambda(x-a)}},
\]
where \( x, a, \lambda \in \mathbb{R} \) and \( \lambda \) is nonzero. The main properties of the sigmoid function \( \sigma^{(\lambda)}_a(x) \) are as follows. The function is monotone increasing if \( \lambda \) is positive, and it is monotone decreasing if \( \lambda \) is negative.

\[
\lim_{x \to +\infty} \sigma^{(\lambda)}_a(x) = \begin{cases} 1, & \text{if } \lambda > 0 \\ 0, & \text{if } \lambda < 0 \end{cases},
\]
\[
\lim_{x \to -\infty} \sigma^{(\lambda)}_a(x) = \begin{cases} 1, & \text{if } \lambda < 0 \\ 0, & \text{if } \lambda > 0 \end{cases}.
\]

\( \sigma^{(\lambda)}_a(a) = 0.5 \), the slope of \( \sigma^{(\lambda)}_a(y) \) in the \( (a, 0.5) \) point is determined by \( \lambda \) as
\[
\frac{d\sigma^{(\lambda)}_a(x)}{dx} \bigg|_{x=a} = \frac{\lambda}{4}.
\]

Case a \((h = 1)\). If \( h = 1 \), then \( \mu_{A_{i,h}}(x) = \mu_{A_{j,1}}(x) \) is given by
\[
\mu_{A_{i,1}}(x) = \frac{1}{1 + e^{-\lambda^{(j)}(x-a^{(j)})}}.
\]
where
\[
a^{(j,r)}_1 = \frac{q_0^{(j)} + q_1^{(j)}}{2}, \tag{15}
\]
\[
\lambda_1^{(j)} = \frac{1}{1 - \varepsilon} \cdot \frac{2}{q^{(j)}_1 - q^{(j)}_0}.
\tag{17}
\]

As \(1/(1 - \varepsilon) - 1 < 1\), \(\lambda_1^{(j)}\) is negative, \(\mu_{A^{(j)}_1}(x)\) is a decreasing function. Note that construction of \(\mu_{A^{(j)}_1}(x)\) also ensures that \(\mu_{A^{(j)}_1}(q^{(j)}_0) \approx 0\).

**Case b (\(h \in \{2, \ldots, k_j - 1\}\)).** Following Dombi’s Plant Inequality Model [11], we will use the following definitions, where the indexes \(l\) and \(r\) stand for “left” and “right”, respectively, and will be used to denote left hand side and right hand side components of fuzzy numbers.

**Definition 2:** The truth of inequality \(a_h^{(j,l)} < x < a_h^{(j,r)}\) is given by the sigmoid function
\[
\{a_h^{(j,l)} < x < a_h^{(j,r)}\} = \sigma^{(\lambda_h^{(j,l)})}_{a_h^{(j,l)}}(x) = \frac{1}{1 + e^{-\lambda_h^{(j,l)}(x-a_h^{(j,l)})}},
\tag{18}
\]
where \(a_h^{(j,l)}, \lambda_h^{(j,l)} \in \mathbb{R}, \lambda_h^{(j,l)} > 0\).

**Definition 3:** The truth of inequality \(x < a_h^{(j,r)}\) is given by the sigmoid function
\[
\{a_h^{(j,r)} > x\} = \sigma^{(\lambda_h^{(j,r)})}_{a_h^{(j,r)}}(x) = \frac{1}{1 + e^{\lambda_h^{(j,r)}(x-a_h^{(j,r)})}},
\tag{19}
\]
where \(a_h^{(j,r)} \in \mathbb{R}\).

Now, we will use the Dombi disjunction operator [13] to implement intersection of two fuzzy sets.

**Definition 4:** (Dombi’s t-norm) Let \(A_1\) and \(A_2\) be fuzzy sets defined over the crisp universe \(X\) given by the membership functions \(\mu_{A_1}(x)\) and \(\mu_{A_2}(x)\), respectively. The Dombi intersection of \(A_1\) and \(A_2\) is given by the membership function \(\mu_{A_1 \cap A_2}(x)\):
\[
\mu_{A_1 \cap A_2}(x) = \mu_{A_1}(x) *_{(D)} \mu_{A_2}(x) =
\frac{1}{1 + \left(\frac{1-\mu_{A_1}(x)}{\mu_{A_1}(x)}\right)^{1/\alpha} + \left(\frac{1-\mu_{A_2}(x)}{\mu_{A_2}(x)}\right)^{1/\alpha}},
\tag{20}
\]
where \(\alpha \in \mathbb{R}, \alpha > 0\), and \(*_{(D)}\) denotes the Dombi intersection operator.

The truth of inequality \(a_h^{(j,l)} < x < a_h^{(j,r)}\) is defined as the intersection of fuzzy sets given by (18) and (19) using Dombi’s t-norm in (20), which hereafter is always applied with \(\alpha = 1\).

**Definition 5:** The truth of inequality \(a_h^{(j,l)} < x < a_h^{(j,r)}\) is given by
\[
\{a_h^{(j,l)} < x < a_h^{(j,r)}\} = \sigma^{(\lambda_h^{(j,l)})}_{a_h^{(j,l)}}(x) *_{(D)} \sigma^{(\lambda_h^{(j,r)})}_{a_h^{(j,r)}}(x),
\tag{21}
\]
and
\[
d_h^{(j)} = \min \{q_{h-1}^{(j)} - q_{h-2}^{(j)}, q_{h}^{(j)} - q_{h-1}^{(j)}\}.
\tag{24}
\]

Setting \(\lambda_h^{(j)}\) so that
\[
\frac{1}{1 + e^{-\lambda_h^{(j)}(x-a_h^{(j,l)})}} = 1 - \varepsilon,
\tag{25}
\]
ensures that \(\mu_{A_h^{(j)}}(x)\) is close to zero in the locus \(q_{h-1}^{(j)}\) and \(q_{h+1}^{(j)}\), while it is close to one in the locus \(q_h^{(j)}\). From (25)
\[
\lambda_h^{(j)} = -\ln \left(\frac{1}{1 - \varepsilon} - 1\right) \frac{2}{d_h^{(j)}},
\tag{26}
\]
h \(\in \{2, \ldots, k_j - 1\}\). Owing to its construction, the membership function \(\mu_{A_h^{(j)}}(x)\) is symmetric and consists of a left hand side and a right hand side sigmoid function \(h \in \{2, \ldots, k_j - 1\}\).

**Case c (\(h = k_j\)).** If \(h = k_j\), then \(\mu_{A_h^{(j)}}(x) = \mu_{A_{j,k}}(x)\) is given by
\[
\mu_{A_{j,k}}(x) = \frac{1}{1 + e^{-\lambda_{k_j}^{(j)}(x-a_{k_j}^{(j,l)})}},
\tag{27}
\]
where
\[
a_{k_j}^{(j,l)} = \frac{q_{k_j}^{(j)} + q_{k_j-1}^{(j)}}{2}.
\tag{28}
\]
\(\lambda_{k_j}^{(j)}\) is identified based on the requirement \(\mu_{A_{j,k}}(q_{k_j}^{(j)}) = 1 - \varepsilon\). From this:
\[
\lambda_{k_j}^{(j)} = -\ln \left(\frac{1}{1 - \varepsilon} - 1\right) \frac{2}{q_{k_j}^{(j)} - q_{k_j-1}^{(j)} - q_{k_j-2}^{(j)}}.
\tag{29}
\]

Figure 2 shows an example of the membership functions of rule antecedents that were formed based on the heuristics introduced above.
functions of the fuzzy numbers

B. Fuzzy Rules

Segmentation of a failure rate time series typically results in a few tens of input-output pairs \((x_i, y_i), i = 1, 2, \ldots, d\). It allows us to build the fuzzy rule base according to the one rule per data method [9]. The fuzzy rule for the \(i\)th input pair is

If \(x_1 \) is \(A_{1,i}^{(i)}\) and \(\ldots\) and \(x_n \) is \(A_{n,i}^{(i)}\), then the output is \(B_i\),

where \((x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\) is an input vector and the fuzzy set \(A_{j,i}^{(i)}\) is given as follows:

\[
I_{j,i}^{(i)} = \{ p : \mu_{A_{j,i}}(x_j) = \max_{l=1, \ldots, k_j} \{ \mu_{A_{l,i}}(x_j) \} \}
\]

\[
k = \min_{m \in I_{j,i}^{(i)}} (m); \quad A_{j,i}^{(i)} = A_{j,k}.
\]

\(B_1, B_2, \ldots, B_m\) are fuzzy sets defined over set \(Y\) which is the domain of crisp system outputs. We will treat each consequent \(B_i\) as a fuzzy number that we will discuss later. In order to make the fuzzy inference system complete, the consequent of each rule needs to be identified; that is, the membership function of each fuzzy number \(B_i\) needs to be supplied \((i = 1, 2, \ldots, d)\).

C. Membership Functions of the Fuzzy Rule Consequents

Let the \(Y\) domain of the crisp outputs of our FIS be

\[
Y = [y_L - \Delta, y_H + \Delta],
\]

where \(y_L = \min_{i=1,\ldots,d}(y_i)\), \(y_H = \max_{i=1,\ldots,d}(y_i)\), \(\Delta = c(y_H - y_L)\), \(c > 0\), \(c \in \mathbb{R}\). In our implementation, \(c = 0.1\). We will use sigmoid functions to compose the membership functions of the fuzzy numbers \(B_1, B_2, \ldots, B_m\). The membership function \(\mu_{B_i^{(l)}}(y)\) is given by Dombi intersection of the left hand side \(B_i^{(l)}\) and the right hand side \(B_i^{(r)}\) fuzzy sets. The \(\mu_{B_i^{(l)}}(y)\) and \(\mu_{B_i^{(r)}}(y)\) membership functions of the fuzzy sets \(B_i^{(l)}\) and \(B_i^{(r)}\) are given by the following sigmoid functions:

\[
\mu_{B_i^{(l)}}(y) = \frac{1}{1 + e^{-\lambda_i^{(l)}(y-a_i^{(l)})}},
\]

\[
\mu_{B_i^{(r)}}(y) = \frac{1}{1 + e^{-\lambda_i^{(r)}(y-a_i^{(r)})}},
\]

where \(y \in \mathbb{Y}, \lambda_i^{(l)}, \lambda_i^{(r)}, a_i^{(l)}, a_i^{(r)} \in \mathbb{R}, \lambda_i^{(l)}, \lambda_i^{(r)} > 0\). Using (20), \(\mu_{B_i}(y)\) is given by

\[
\mu_{B_i^{(l)}}(y) * (D) \mu_{B_i^{(r)}}(y) = \frac{1}{1 + e^{-\lambda_i^{(l)}(y-a_i^{(l)})} + e^{\lambda_i^{(r)}(y-a_i^{(r)})}}.
\]

We use the following heuristics to construct the membership functions. Here,

\[
\tilde{y}_i = \frac{\sum_{v=1}^{d} \mu_i(x_v) y_v}{\sum_{v=1}^{d} \mu_i(x_v)},
\]

where

\[
\mu_i(x_v) = \min_{j=1,\ldots,n} \left( \mu_{A_{i,j}}(x_v) \right).
\]

\(\mu_i(x_v)\) measures how much the \(i\)th input vector \(x_v = (x_{v,1}, \ldots, x_{v,n})\) activates the \(i\)th fuzzy rule \((i = 1, 2, \ldots, d)\).

Let the vector \(y^* = (y_1^*, \ldots, y_d^*)\) contain the \(\tilde{y}_i\) values in non-decreasing order; that is, \(y_1^* \leq y_2^* \leq \ldots \leq y_d^*\) and \(y_i^* \in \{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_d\}\). The parameters \(a_i^{(l)}, a_i^{(r)}, \lambda_i^{(l)}, \lambda_i^{(r)}\) of \(\mu_{B_i^{(l)}}(y)\) and \(\mu_{B_i^{(r)}}(y)\) may be set as

\[
a_i^{(l)} = y_i^* - d_i^{(l)} / 2; \quad a_i^{(r)} = y_i^* + d_i^{(r)} / 2
\]

\[
\lambda_i^{(l)} = -\ln \left( \frac{1}{1 - \varepsilon} - 1 \right) \frac{2}{d_i^{(l)}},
\]

\[
\lambda_i^{(r)} = -\ln \left( \frac{1}{1 - \varepsilon} - 1 \right) \frac{2}{d_i^{(r)}},
\]

where

\[
d_i^{(l)} = y_i^* - y_{i-1}^* \quad \text{for } i = 2, 3, \ldots, d
\]

\[
d_1^{(l)} = d_2^{(l)}
\]

\[
d_i^{(r)} = y_{i+1}^* - y_i^* \quad \text{for } i = 1, 2, \ldots, d - 1
\]

\[
d_d^{(r)} = d_{d-1}^{(r)}
\]

D. Aggregation of Fuzzy Outputs

As we remarked above, we will treat the \(B_i\) consequent of each fuzzy rule as a fuzzy number, namely, as a so-called Pl¨a¨an number that is given by two sigmoid functions. We will aggregate the Pl¨a¨an numbers using the Pl¨a¨an arithmetics based on the next proposition [11].

Proposition 1: If \(A_1, A_2, \ldots, A_m\) are fuzzy sets with the membership functions \(\sigma_k^{(l)}(y), \sigma_k^{(r)}(y), \ldots, \sigma_k^{(l)}(y)\), respectively, \(\text{sgn}(\lambda_1) = \text{sgn}(\lambda_2) = \cdots = \text{sgn}(\lambda_m)\), and the fuzzy set \(A\) is given by the linear combination

\[
A = \sum_{i=1}^{d} w_i A_i,
\]

Fig. 2. Membership functions of rule antecedents
where $\sum_{i=1}^{d} w_i = 1$, then $A$ is also sigmoid-shaped with the membership function $\sigma_a^{(l)}(y)$, where

$$a = \sum_{i=1}^{d} w_i a_i; \quad 1 = \sum_{i=1}^{d} \frac{w_i}{\lambda_i}. \quad (40)$$

**Proof.** See [11].

Similarly to (36), let the quantity $\mu_i(x)$ measure how much the $n$-dimensional input vector $x$ activates the $i$th fuzzy rule. In other words, $\mu_i(x)$ may be interpreted as the level of applicability of rule $i$ to input $x$. Based on this, the normalized weight of rule $i$, $w_i$, for input $x$ is

$$w_i = \frac{\mu_i(x)}{\sum_{k=1}^{d} \mu_k(x)} \quad \text{if} \quad \sum_{k=1}^{d} \mu_k(x) \neq 0. \quad (41)$$

Notice that if $\sum_{k=1}^{m} \mu_k(x) = 0$, we infer nothing. Proposition 1 allows us to aggregate the left hand sides and right hand sides of $B_1, B_2, \ldots, B_m$ separately using the $w_i, w_2, \ldots, w_m$ weights and derive the parameters of the left hand side $\mu_i^{(l)}(y)$ and the right hand side $\mu_i^{(r)}(y)$ of the fuzzy output $B$ in the following way:

$$a^{(l)} = \sum_{i=1}^{d} w_i a_i^{(l)}; \quad a^{(r)} = \sum_{i=1}^{d} w_i a_i^{(r)}$$

$$\frac{1}{\lambda^{(l)}} = \sum_{i=1}^{d} \frac{w_i}{\lambda_i^{(l)}}, \quad \frac{1}{\lambda^{(r)}} = \sum_{i=1}^{d} \frac{w_i}{\lambda_i^{(r)}}. \quad (42)$$

The membership function $\mu_i(y)$ of the fuzzy output $B$ is computed according to (20):

$$\mu_B(y) = \frac{1}{1 + e^{-\lambda^{(l)}(y-a^{(l)})} + e^{-\lambda^{(r)}(y-a^{(r)})}}. \quad (43)$$

Here, it is worth mentioning that there is no implication applied in our inference method, the fuzzy output $B$ is inferred via fuzzy arithmetic operations.

**E. Defuzzification.**

In order to use our FIS for modeling and forecasting, a defuzzification method, which transforms the fuzzy output $B$ into a crisp $\hat{y}$, needs to be identified. Here, we will use a maximum defuzzification method; that is, we will represent the fuzzy output $B$, which is given by $\mu_B(y)$ in (43), by the crisp $\hat{y}$ for which

$$\mu_B(\hat{y}) = \max_{y \in X} \mu_B(y). \quad (44)$$

It can be seen that the unique solution of (44) is

$$\hat{y} = \frac{\lambda^{(l)}}{\lambda^{(l)} + \lambda^{(r)}} a^{(l)} + \frac{\lambda^{(r)}}{\lambda^{(l)} + \lambda^{(r)}} a^{(r)} + \frac{1}{\lambda^{(l)} + \lambda^{(r)}} \left( \ln(\lambda^{(l)}) - \ln(\lambda^{(r)}) \right). \quad (45)$$

$$\quad + \frac{1}{\lambda^{(l)} + \lambda^{(r)}} \left( \ln(\lambda^{(l)}) - \ln(\lambda^{(r)}) \right). \quad (46)$$

In this form $\hat{y}$ may be interpreted as the weighted average of $a^{(l)}$ and $a^{(r)}$ in (45) corrected with the term in (46). The introduced defuzzification method has two notable advantages. On the one hand, it does not require any numerical integration to generate the crisp output, on the other hand, it runs in a constant time as its speed does not depend on the number of fuzzy rules.

We call the introduced inference system Pliant Arithmetic-based Fuzzy Inference System (PAFIS).

**IV. SIMULATION AND FORECASTING.**

Let $PAFIS_\alpha$ and $PAFIS_\beta$ be two Pliant Arithmetic-based Fuzzy Inference systems that were trained with the training pairs

$$\begin{pmatrix} \alpha_i^{(l)}, \beta_i^{(l)}, s_i, \alpha_i^{(r)}, \beta_i^{(r)} \end{pmatrix}, \quad (47)$$

respectively, $i = 1, 2, \ldots, d$. The meaning of $\alpha_i^{(l)}, \beta_i^{(l)}, s_i, \alpha_i^{(r)}$ and $\beta_i^{(r)}$ are the same as described in section II. Let $\omega_1, \omega_2, \ldots, \omega_n$ be a $u$-period long sub-time-series in a failure rate time series. $PAFIS_\alpha$ and $PAFIS_\beta$ are able to provide with the parameters $\alpha_o$ and $\beta_o$ of linear trend of the $u$-period long continuation of the input time series $\omega_1, \omega_2, \ldots, \omega_n$.

Firstly, we identify the slope parameter $\alpha$ and the intersection parameter $\beta$ of the line that fits the best to the normalized input time series. The normalization is done by using the average $\bar{y}$ and the standard deviation $s$ of the input time series as discussed in section II. Once the input vector $x = (x_1, x_2, x_3) = (\alpha, \beta, s)$ has been identified, we can use it as an input to $PAFIS_\alpha$ and $PAFIS_\beta$. Let the outputs of the systems be $\alpha_o = PAFIS_\alpha(x)$ and $\beta_o = PAFIS_\beta(x)$. Having $\alpha_o$ and $\beta_o$, the linear trend prediction $\hat{\omega}_{u+t}$ for the continuation of failure rate time series $\omega_1, \omega_2, \ldots, \omega_u$ may be given as follows:

$$\hat{\omega}_{u+t} = \begin{cases} \alpha_o t + \beta_o + \bar{y}, & \text{if } s = 0 \\ (\alpha_o t + \beta_o) s + \bar{y}, & \text{if } s > 0. \end{cases} \quad (48)$$

Based on empirical observations, it may be assumed that electronic components of the same commodity have similar failure rate curves; that is, their empirical failure rate time series are similar, too. For example, we may have the empirical failure rate times series of 100 end-of-life laptop motherboard types. Here, let us assume that we have the empirical failure rate time series of $N$ electronic components that belong to the same commodity. In this case, we can built a $PAFIS_{i,\alpha}$ and $PAFIS_{i,\beta}$ system for each of the failure rate time series, $i = 1, 2, \ldots, N$. Furthermore, let $\mu_{\max_{i,\alpha}}(x)$ be the maximum of the normalized activation levels of the rules in $PAFIS_{i,\alpha}$ for the input $x$; that is,

$$\mu_{\max_{i,\alpha}}(x) = \max_{m=1,\ldots,d_i} \frac{\mu_m^{(i,\alpha)}(x)}{\sum_{k=1}^{d_i} \mu_k^{(i,\alpha)}(x)}. \quad (49)$$

where $\mu_k^{(i,\alpha)}(x)$ is the activation level of the $k$th rule in $PAFIS_{i,\alpha}$ for the input $x$, $d_i$ is the number of rules in $PAFIS_{i,\alpha}$. The quantity $\mu_{\max_{i,\alpha}}(x)$ for $PAFIS_{i,\beta}$ can be calculated similarly to (49). Having the outputs $\alpha_{t,\alpha}$ and $\beta_{t,\alpha}$ of $PAFIS_{i,\alpha}$ and $PAFIS_{i,\beta}$, respectively, for the input $x,$
the \( i \)th linear trend prediction \( \hat{\omega}_{i,u+t} \) for the continuation of failure rate time series \( \omega_1, \omega_2, \ldots, \omega_n \) may be given as follows:

\[
\hat{\omega}_{i,u+t} = \begin{cases} 
\alpha_i + \beta_i + \omega, & \text{if } s = 0 \\
(\alpha_i + \beta_i + \omega_i) s + \omega, & \text{if } s > 0
\end{cases},
\]

(50)

The linear trend prediction \( \hat{\omega}_{u+t} \) for the continuation of failure rate time series \( \omega_1, \omega_2, \ldots, \omega_n \) may be calculated by using each of the individual linear trend predictions \( \hat{\omega}_{i,u+t} \):

\[
\hat{\omega}_{u+t} = \sum_{i=1}^{N} \frac{\mu_{max,i,\alpha}(x) \mu_{max,i,\beta}(x)}{\mu_{max,i,\alpha}(x) \mu_{max,i,\beta}(x)} \hat{\omega}_{i,u+t}.
\]

(51)

V. EMPIRICAL RESULTS AND CONCLUSIONS

Performance of the Pliant Arithmetic-based Fuzzy Inference System was tested on 38 real life failure rate time series. These time series represent the weekly failure rates of 38 different motherboard types, typically over 150 to 210 weeks. 30 of these failure rate time series were used for training purposes and 8 for validation. For the training data set, 6 different normalized line segment models were created for each time series with 5, 10, 15, 20, 25 and 30 training records and for each case a PAFIS and ANFIS system was tested on 38 real life failure rate time series. These failure rate time series were compared to those of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) programmed in MATLAB R2016a by using its ANFIS tool. The simulation results are summarized in Table I. The MSE results of simulations on the training data set are shown in Table II.

Next, each PAFIS and ANFIS that had been build from the training data was tested on the validation data set; that is, on data that had not been used for building the fuzzy inference systems. The average of MSE values of the individual simulations were computed for each PAFIS and ANFIS system. The simulation results on the validation data are summarized in Table II. We can see from Table I that on the training data, the ANFIS simulation performs better than the PAFIS simulation, but at the same time PAFIS gives nearly the same goodness when the number of fuzzy rules is low (e.g. line 1 and 2 two in Table I); that is, when the modeling can be done by a relatively small number of fuzzy rules, the PAFIS method is a good alternative of the much more complex ANFIS method. From Table II, we can see that ANFIS and PAFIS perform similarly on the validation data set.

It is worth mentioning that the PAFIS method is much simpler than the ANFIS method, moreover, based on the empirical results, application of the Pliant Arithmetic-based Fuzzy Inference System for failure rate forecasting may be viewed as a viable alternative modeling and prediction technique.

REFERENCES


