Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.1815

Published online in Wiley Online Library

Clustering Empirical Failure Rate Curves for Reliability Prediction Purposes in the Case of Consumer Electronic Products

József Dombi,[a] Tamás Jónás[b] and Zsuzsanna Eszter Tóth[b*†]

In this paper, a methodology based on the combination of time series modeling and soft computational methods is presented to model and forecast bathtub-shaped failure rate data of newly marketed consumer electronics. The time-dependent functions of historical failure rates are typified by parameters of an analytic model that grabs the most important characteristics of these curves. The proposed approach is also verified by the presentation of an industrial application brought along at an electrical repair service provider company. The prediction capability of the introduced methodology is compared with moving average-based and exponential smoothing-based forecasting methods. According to the results of comparison, the presented method can be considered as a viable alternative reliability prediction technique. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: service provider; bathtub curve; empirical failure rate curve; time series; reliability prediction

1. Introduction

We present a methodology that focuses on consumer electronic products. Because of shortening life cycles and times between the design and release-to-market phases, consumer electronics are typically functionally tested rather than passed through deeper reliability tests. It results in the fact that the failure rate or hazard rate function, which is a commonly used quantity to describe electronic product life characteristics,[1–3] is bathtub-shaped with all three sections of the traditional bathtub curve. In this paper, a new way of modeling and forecasting bathtub-shaped failure rate lifetime data of newly marketed consumer electronics is addressed and discussed.

Electronic repair service provider (ERSP) companies dealing with the repair process of consumer electronic products are brought into focus in our research. This kind of electronics manufacturing service (EMS) companies possesses field return data enabling the production of empirical failure rate curves. Based on their industrial experience, the empirical failure rate curves serve as a knowledge base for reliability prediction and resource allocation purposes for repair service processes. The historical failure rate curves representing the whole life cycles of on-the-market electronic products lay the foundation of predicting the reliability of newly designed electronic products. Our paper aims to develop a model for typifying historical failure rate curves in order to predict the reliability of such products by examining the application of the time series modeling approach.

The proposed model describes the bathtub curve as three connected line segments (Figure 1); each of them refers to the corresponding phase of the traditional bathtub curve. The observed failure rates of repairable electronic products as the inputs of our model can be considered as time-dependent failure rates and hence can be treated as time series. We constructed a model that is reasonably simple and applicable for the whole product life cycle to make overall decisions. Therefore, we follow an approach that can be successfully used to discover similarities among time series of historical failure rates. The time-dependent functions of failure rates are typified by different parameters of an analytic model that grabs the most important characteristics of these curves. Based on the appropriate parameterization of the model, clustering of the different bathtub curves originating from historical field data becomes possible in a three-dimensional or four-dimensional space. Parameters of our model have semantics because of their geometric interpretation, and so, the presented approach has a cognitive aspect as well.

[a]*Institute of Informatics, University of Szeged, Szeged, Hungary*
[b]*Institute of Business Sciences, Faculty of Economics and Social Sciences, Budapest University of Technology and Economics, Budapest, Hungary*
**Correspondence to: Zsuzsanna Eszter Tóth, Institute of Business Sciences, Faculty of Economics and Social Sciences, Budapest University of Technology and Economics, 1117 Budapest, Magyar tudósok körútja 2, Building Q, Wing A, Room 309, Budapest, Hungary.*
[†]*Email: tothzs@mvt.bme.hu*

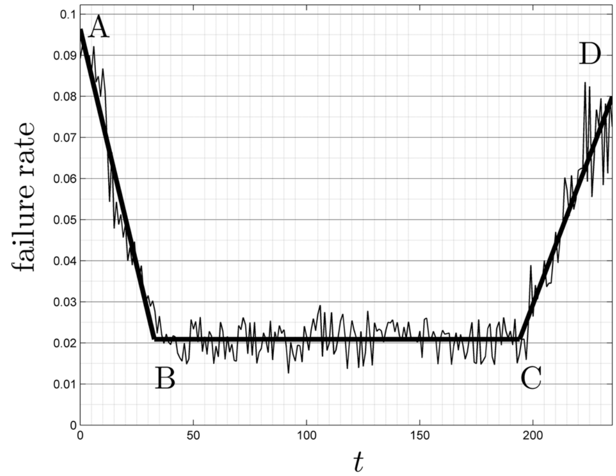

Figure 1. A typical empirical failure rate curve

The methodology we propose combines time series modeling and soft computational methods and, as such, can be considered as a hybrid one, which serves as a concrete alternative for modeling bathtub-shaped failure rate functions of newly marketed electronic products.

The information deriving from the failure rate prediction provides important inputs to decision-making in the new product development process. This kind of reliability analysis enables the manufacturer to assess and improve the dependability of the electronic product and improve the performance of next-generation products. Our model is also verified by the presentation of an industrial application brought along at an ERSP company where software was developed to apply the model. In order to evaluate the goodness of our approach as a robustness test, we also compare its results with moving average-based and exponential smoothing-based forecasts.

The remaining part of the paper is organized as follows. In Section 2, we present a brief literature review. Section 3 proposes a time series model and an analytical model for parameterization, based on which the clustering of typical failure rate curves becomes possible. In Section 4, the application of our approach is demonstrated through an industrial example. Finally, in Section 5, the paper is closed by drawing a number of key conclusions and presenting important managerial implications.

2. Literature review

New electronic products are being launched on the market at an ever-accelerated pace because of the advances in related technology and to the increasing demands and expectations of customers and users.[4] As electronic products are becoming more sophisticated, customers want to be assured of satisfactory performance in the long run. Because of these changes, dealing with reliability cannot be an afterthought for electronic manufacturers either. Reliability analysis and prediction have become an increasingly important issue for the companies competing in the electronic industry.[1,5]

The electronic sector has been undergoing a momentous transformation during the last two decades in many ways. The complexity and sophistication of electronic products and product functionality have greatly improved in the electronic sector.[1] Customers of electronic products are increasingly demanding highly customized and high-quality products at a competitive price. New products are constantly being launched, the time of which can be measured in weeks and months rather than years, which results in shortening product life cycles and delivery times. Original equipment manufacturers (OEMs) have to keep pace with the increasing requirements.[6] Because of the highly competitive nature of the electronics industry, core competencies and the need for reducing manufacturing costs have come to the front. These tendencies have resulted in the outsourcing of high-level manufacturing capabilities, design, engineering, and aftermarket services. This outsourcing trend in the electronic industry has laid a strong foundation of the subcontracting sector. Particularly, modular production has enhanced the growth of the EMS industry,[7,8] which is the sector that has been performing strongly in the last 20 years.[9]

An EMS provider is usually an independent organization without managing an own brand but having a wide range of customers managing entire product lines and whole supply chains for OEMs and offering end-to-end services.[1,10] One kind of service that EMS companies provide for OEMs is the industrial repair service as a piece of aftermarket services. The number of ERSP enterprises dealing with electronic repair services has been increasing significantly in the EMS sector.[9] These aftermarket services are developed to repair different types of electronic products, modules, and parts. In today's highly competitive electronic products market, product reliability analysis and prediction has become extremely important for OEMs.[3] ERSP companies play an important role in providing OEMs with relevant information regarding the reliability of their products and possess field failure data of on-the-market consumer electronic products, which can serve as a basis for predicting the reliability of future products.

In reliability theory, the 'bathtub shape' is widely used to describe the failure patterns of different products. This hazard-rate curve, which typically maps the failure rate versus time, has been verified by experience for many types of products.

Many lifetime failure data have a bathtub-shaped failure rate. Therefore, the estimation of the failure rate function is of fundamental importance in reliability analysis. According to the industrial experience, it is verified that the failure rate function of consumer electronics follows the phases of the traditional bathtub curve. The bathtub curve presents the idea that the operation of a device population can be divided into three distinct periods called infancy, useful life, and wear-out periods with each region corresponding to a specific type of failure (Figure 1).

The relevant literature regarding the bathtub curve is extremely wide-ranging. Bathtub-shaped hazard rate functions[11,12] appear frequently in reliability engineering.[13,14] A huge number of bathtub curve models are developed for modeling data that display the bathtub-shaped failure rate.[11,15,16] Hundreds of papers deal with the distributions and the parameterization describing the bathtub curve. Some new models have been proposed recently.[17–21] Lai *et al.*[11] gives an overview of bathtub-shaped failure rate distributions. There are three typical trends of methods to establish a bathtub curve model:[22] looking for an appropriate model, modifying the traditional model by complementing it with additional parameters, and transforming the data to achieve compatibility with a well-understood and convenient traditional model. Many literatures analyze the turning point of the failure rate function, which helps to determine and plan appropriate burn-in, maintenance, and repair policies and strategies.[23] Hallberg[24] introduces a model for predicting failure rate as a function of stress and time. Govil and Aggarwal[25] propose a six-parameter and five-parameter model to represent the bathtub curve. Cassanelli *et al.*[26] suggested a reliability assessment method for new products by collecting empirical data and reliability information on similar components. Lee and Lee[27] developed a Web system-based electronic components reliability prediction system based on the failure model and created a database from them. Campbell *et al.*[5] studied the reliability behavior of electronic components as a function of time.

Companies competing in the electronic industry have two opportunities when it comes to constructing a reliability prediction model for repairable electronic products. On the one hand, an appropriate probability function can be selected in advance and then applied for predicting future reliability.[28,29] Liang[28,29] argues that despite the fact that the predetermined probability function fits the failure data well, the results of prediction may not be satisfactory enough in practice. It must be also taken into consideration that there is a wide range of appropriate probability functions; so, selecting one from the many existing ones is not an easy task for reliability professionals.[28–30]

One of the most widely used applications of reliability analysis is to monitor field failure data. Considering field failure data as time series, the other viable opportunity is to construct a reliability prediction model involving time series analysis.[31] However, time series models have been successfully applied in many engineering and economic fields, and their use in the field of reliability prediction can only be dated back to the late 90s.[32–35] The aim of using time series methods is to recognize iterative patterns and nonlinear relationships.[36] Time series models are well-grounded both statistically and theoretically, and when analyzing failure rate data, no requirements and assumptions of models are required. It is also important to note that because the time series approach gives a higher weightage to more recent data, it is very suitable for repairable systems.[37]

Traditional time series methods include the exponential-smoothing method, the moving average method, the autoregressive integrated moving-average (ARIMA) model, the seasonal ARIMA models,[31] and the time series decomposition method, all of which have been used for predicting repairable system reliability.[30,38–40] Neural networks can also be employed for time series prediction.[41–44]

The presented approach is a hybrid model in the sense that we consider historical failure rates as time series to which a parametric analytical model is fit. After parameterization, we apply a fuzzy clustering method to identify typical failure rate curves. In our research, the objective was to examine the usefulness of this hybrid approach in predicting the failure rates of newly marketed electronic products. We considered the ratios of failed products per interval as a time series and then applied a time series approach to model and analyze the failure behavior. Once the model parameters were estimated, the model could be used for future prediction. We are also to prove that our method based on analyzing time series gives satisfactory results in terms of its predictive performance and hence can be a viable alternative to other widely used prediction tools such as moving average and exponential smoothing.

3. Methodology

3.1. The line segment model

Our approach is founded on the phenomenon that there are three characteristic parts recognizable in each of the studied empirical failure rate curves, and each of these parts can be represented by a simple line segment as shown in Figure 1.

The three characteristic parts of such a failure rate curve can be determined by the A, B, C, and D points. The $\overline{AB}$ line segment represents the first decreasing part of the curve, and the $\overline{BC}$ line segment corresponds to the quasi-constant middle piece of the curve, while the $\overline{CD}$ segment represents the increasing part of the bathtub curve. Having the A, B, C, and D points identified, the failure rate curve can be represented by the $\overline{AB}$, $\overline{BC}$, and $\overline{CD}$ line segments.

Let $\lambda_0, \lambda_1, \ldots, \lambda_n$ be the known weekly failure rates of an electronic product. The $\lambda_0, \lambda_1, \ldots, \lambda_n$ series can be considered as the $\lambda_{t_0}, \lambda_{t_1}, \ldots, \lambda_{t_n}$ times series, where $t_i = i$ and $i = 1, 2, \ldots, n$. Let the A, B, C, and D points be $A(t_A, \lambda_A)$, $B(t_B, \lambda_B)$, $C(t_C, \lambda_C)$, and $D(t_D, \lambda_D)$, respectively, in the time–failure rate coordinate system. As A is the first, and D is the last point of the $\overline{AB}$ and $\overline{CD}$ segments, respectively, $t_A = 0$ and $t_D = n$. The $\overline{CD}$ segment represents the constant piece of the failure rate curve; so, $\lambda_B = \lambda_C$. By this means, the A, B, C, and D

points are $A(0,\lambda_A)$, $B(t_B,\lambda_B)$, $C(t_C,\lambda_B)$, and $D(n,\lambda_D)$, where $0<t_C<t_B<n$, and $\lambda_A,\lambda_D>\lambda_B$. The segments $\overline{AB}$, $\overline{BC}$, and $\overline{CD}$ fitting best to the failure rate curve can be identified by minimizing the

$$\sum_{i=0}^{n}(f_i-\lambda_i)^2 \tag{1}$$

quantity, where

$$f(t)=\begin{cases}\dfrac{\lambda_B-\lambda_A}{t_B}t+\lambda_A, & \text{if } 0\le t<t_B\\ \lambda_B, & \text{if } t_B\le t<t_C\\ \dfrac{\lambda_D-\lambda_B}{n-t_C}(t-n)+\lambda_D, & \text{if } t_C\le t\le n\end{cases} \tag{2}$$

The unknown parameters $\lambda_A,\lambda_B,\lambda_D,t_B$, and t_C of $f(t)$ can be determined by using the generalized reduced-gradient method.[45,46]

For practical applications, it is worthwhile to calculate the root-mean-square error (RMSE) between $f(t)$ and the λ_0, λ_1, …, λ_n failure rate time series (FRTS) for each segment of $f(t)$ and link the RMSEs to $f(t)$. The σ_1, σ_2, and σ_3 RMSEs are calculated as

$$\begin{aligned}\sigma_1&=\sqrt{\frac{1}{\lfloor t_B\rfloor+1}\sum_{i=0}^{\lfloor t_B\rfloor}(f(i)-\lambda_i)^2}\\ \sigma_2&=\sqrt{\frac{1}{\lfloor t_C\rfloor-\lfloor t_B\rfloor}\sum_{i=\lfloor t_B\rfloor+1}^{\lfloor t_C\rfloor}(f(i)-\lambda_i)^2}\\ \sigma_3&=\sqrt{\frac{1}{n-\lfloor t_C\rfloor}\sum_{i=\lfloor t_C\rfloor+1}^{n}(f(i)-\lambda_i)^2}\end{aligned} \tag{3}$$

Figure 2 depicts the graph of an $f(t)$ function. Such a function is unambiguously determined by the A, B, C, and D points.

Let e be the closed half-line starting from point C (towards point D). Let us assume that instead of line segment $\overline{CD}$, the closed half-line e describes the third part of $f(t)$; that is, this part of $f(t)$ is given by point C and slope of e. It means that e can be given by points C and $E(t_E,\lambda_E)$, where

$$\begin{aligned}t_E&=t_B+t_C\\ \lambda_E&=f(t_E)\end{aligned} \tag{4}$$

In this way, function $g(t)$ is derived from function $f(t)$ as

$$g(t)=\begin{cases}\dfrac{\lambda_B-\lambda_A}{t_B}t+\lambda_A, & \text{if } 0\le t<t_B\\ \lambda_B, & \text{if } t_B\le t<t_C\\ \dfrac{\lambda_E-\lambda_B}{t_B}(t-t_B-t_C)+\lambda_E, & \text{if } t_C\le t\le n\end{cases} \tag{5}$$

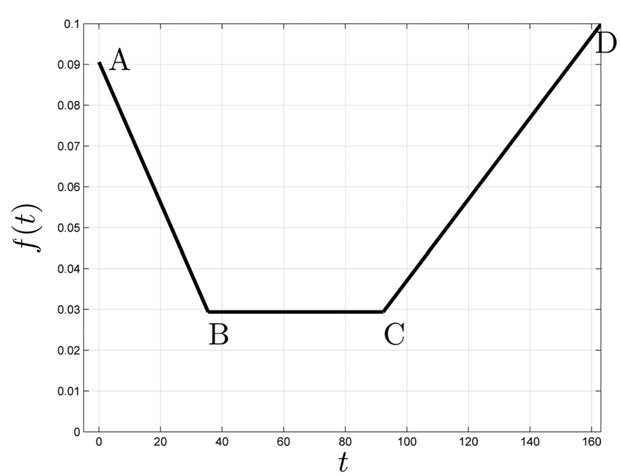

Figure 2. Graph of an $f(t)$ function

From Equation (5), function $g(t)$ is unambiguously given by parameters $\lambda_A, \lambda_B, \lambda_E, t_B$, and t_C. Function $g(t)$ can also be determined by the $h_1, \lambda_B, d, l, h_2$ parameters as shown in Figure 3, where

$$\begin{aligned} h_1 &= \lambda_A - \lambda_B \\ d &= t_B \\ l &= t_C - t_B \\ h_2 &= \lambda_E - \lambda_B \end{aligned} \tag{6}$$

Note that due to $t_E = t_B + t_C$ or $t_E = 2d + l$, $g(t)$ is given by five parameters (either $\lambda_A, \lambda_B, \lambda_E, t_B, t_C$ or $h_1, \lambda_B, d, l, h_2$), while $f(t)$ is determined by six parameters ($\lambda_A, \lambda_B, \lambda_D, t_B, t_C$ and n). By this means, $f(t)$ and $g(t)$ model the same $\lambda_0, \lambda_1, \ldots, \lambda_n$ FRTS, but $g(t)$ has less parameters. The $g(t)$ function is called the line segment model (LSM) of the $\lambda_0, \lambda_1, \ldots, \lambda_n$ FRTS (Figure 3).

3.2. The standard line segment model

Once the $\lambda_A, \lambda_B, \lambda_E, t_B$ and t_C parameters of the LSM $g(t)$ for a particular $\lambda_0, \lambda_1, \ldots, \lambda_n$ FRTS are identified, the $g(t)$ model can be standardized to the $s: [0,1] \to [0,1], x \mapsto s(x)$ function by applying the following transformation:

$$\begin{aligned} x &= \frac{t}{t_B + t_C} \\ s(x) &= \frac{g((t_B + t_C)x) - \lambda_B}{\max(\lambda_A, \lambda_E) - \lambda_B} \end{aligned} \tag{7}$$

According to this transformation, the (x, y) coordinates of the transformed A, B, C, and E points are

$$\begin{aligned} x_A &= 0; y_A = s(x_A) = \frac{\lambda_A - \lambda_B}{\max(\lambda_A, \lambda_E) - \lambda_B} \\ x_B &= \frac{t_B}{t_B + t_C}; y_B = s(x_B) = 0 \\ x_C &= \frac{t_C}{t_B + t_C}; y_C = s(x_C) = 0 \\ x_E &= 1; y_E = s(x_E) = \frac{\lambda_E - \lambda_B}{\max(\lambda_A, \lambda_E) - \lambda_B} \end{aligned} \tag{8}$$

So, the $s(x)$ function, which we call the standard LSM (SLSM) of the $\lambda_0, \lambda_1, \ldots, \lambda_n$ FRTS, can be written as

$$s(x) = \begin{cases} \frac{-y_A}{x_B} x + y_A, & \text{if } 0 \le x < x_B \\ 0, & \text{if } x_B \le x < x_C \\ \frac{y_E}{1 - x_C}(x - 1) + y_E, & \text{if } x_C \le x \le 1 \end{cases} \tag{9}$$

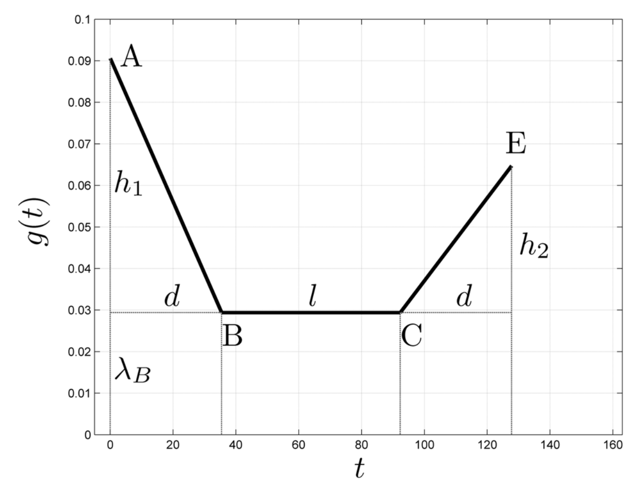

Figure 3. Graph of function $g(t)$ derived from function $f(t)$

As the LSM in Equation (5) can be described by the $h_1, \lambda_B, d, l, h_2$ parameters, the SLSM in Equation (9) is determined by y_1, x_d, x_l and y_2, where

$$\begin{aligned} y_1 &= y_A \\ x_d &= x_B \\ x_l &= x_C - x_B \\ y_2 &= y_E \end{aligned} \tag{10}$$

and as $2x_d + x_l = 1$,

$$x_d = \frac{1 - x_l}{2} \tag{11}$$

It means that $s(x)$ is given by the y_1, y_2 height parameters and the x_l middle-length parameter as shown in Figure 4. Note that due to Equation (11), the middle segment of $s(x)$ (with a length of x_l) is centered to the $(0.5, 0)$ point. In what follows, let $s_{y_1,y_2,x_l}(x)$ denote the SLSM with the y_1, y_2 and x_l parameters.

3.3. Clustering the SLSMs

Let

$$\begin{aligned} &\lambda_{1,0}, \lambda_{1,1}, \ldots, \lambda_{1,n_1} \\ &\lambda_{2,0}, \lambda_{2,1}, \ldots, \lambda_{2,n_2} \\ &\vdots \\ &\lambda_{m,0}, \lambda_{m,1}, \ldots, \lambda_{m,n_m} \end{aligned} \tag{12}$$

be FRTSs that we wish to cluster. The SLSM model for each FRTS can be created as discussed before.
Let

$$s_{y_{1_i}, y_{2_i}, x_{li}}(x) \tag{13}$$

be the SLSM for the $\lambda_{i,0}, \lambda_{i,1}, \ldots, \lambda_{i,n_i}$ FRTS, where $i = 1, \ldots, m$. These SLSMs can be clustered into the $C_1, \ldots, C_N$ clusters ($N \le m$) based on their y_{1i}, y_{2i}, x_{li} parameters in the three-dimensional space using the fuzzy C-means clustering method.[47,48] Let N_j denote the number of SLSMs in cluster C_j, and let I_j be the set of indexes of SLSMs that belong to cluster C_j, that is,

$$I_j = \left\{ k | s_{y_{1_k}, y_{2_k}, x_{l_k}}(x) \in C_j, k \in \{1, \ldots, m\} \right\} \tag{14}$$

where $j = 1, \ldots, N$. For cluster C_j, the parameters of the cluster characteristic SLSM

$$s_{y_{1C_j}, y_{2C_j}, x_{lC_j}}(x) \tag{15}$$

are defined as the

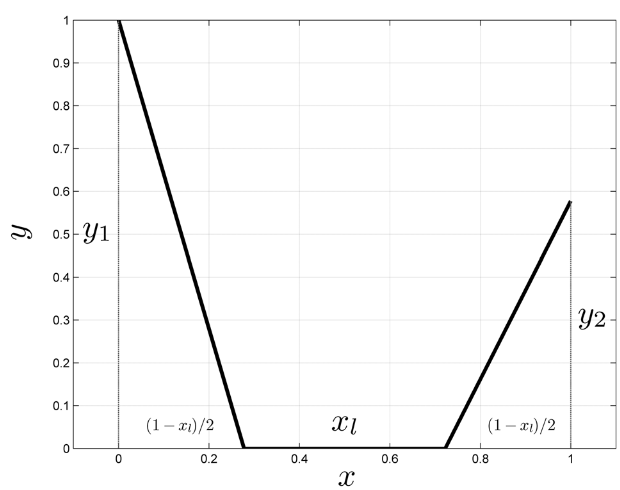

Figure 4. Graph of $s(x)$

$$
\begin{aligned}
y_{1_{C_j}} &= \frac{1}{N_j} \sum_{r \in I_j} y_{1r} \\
y_{2_{C_j}} &= \frac{1}{N_j} \sum_{r \in I_j} y_{2r} \\
x_{I_{C_j}} &= \frac{1}{N_j} \sum_{r \in I_j} x_{Ir}
\end{aligned}
\tag{16}
$$

averages, where $j = 1, \ldots, N$; that is, the cluster characteristic SLSMs are the cluster centroids.

The FRTSs in Equation (12) are all complete ones; that is, they represent FRTSs of end-of-life consumer electronic products that are not repaired anymore. In other words, each studied FRTS represents the entire failure rate history of a product type. Thus, the cluster characteristic SLSMs represent the typical SLSMs of the FRTSs.

3.4. Fitting LSMs to a fractional failure rate time series

A fractional FRTS is not complete; that is, it will be continued in the future, and its new values are becoming known week by week. Now, we discuss how the LSM, which fits to a fractional FRTS the best, can be generated from the characteristic SLSMs. Let $\lambda_{F,0}, \ldots, \lambda_{F,M}$ be a fractional FRTS. For each $s_{y_{1_{C_j}}, y_{2_{C_j}}, x_{I_{C_j}}}(x)$ characteristic SLSM, we are looking for the $\alpha_{C_j} > 0, \beta_{C_j} \geq 0$ and $\gamma_{C_j} > 0$ parameters so that

$$
\begin{aligned}
& g_{C_j} : \left[0, \alpha_{C_j}\right] \rightarrow [0, 1] \\
& \alpha_{C_j} \geq M \\
& g_{C_j}(t) = \gamma_{C_j} s_{y_{1_{C_j}}, y_{2_{C_j}}, x_{I_{C_j}}} \left(\frac{t}{\alpha_{C_j}} \right) + \beta_{C_j} \\
& \sum_{i=0}^{M} \left(g_{C_j}(i) - \lambda_{F,i} \right)^2 \rightarrow \min
\end{aligned}
\tag{17}
$$

Solution for each fitting problem in Equation (17) ($j = 1, \ldots, N$) can be found by applying the generalized reduced-gradient method.[45,46] Let $F(t)$ be the best-fitting function derived from the SLMSs based on Equation (17), that is,

$$
\begin{aligned}
& F(t) = g_{C_k}(t) \Big|_{\sum_{i=0}^{M} \left(g_{C_k}(i) - \lambda_{F,i} \right)^2 = \min\limits_{j=1,\ldots,N} \sum_{i=0}^{M} \left(g_{C_j}(i) - \lambda_{F,i} \right)^2} \\
& k = 1, \ldots, N
\end{aligned}
\tag{18}
$$

3.4.1. Predicting failure rates. Let us assume that $F(t)$ has the α, β, and γ parameters. As $\alpha \geq M$, the $F(i)$ values for $\lfloor \alpha \rfloor \geq i > M$ can be taken as predictions of the unknown $\lambda_{F,M+1}, \ldots, \lambda_{F,\lfloor \alpha \rfloor}$ future values; that is, the $\lambda_{F,M+1}, \ldots, \lambda_{F,\lfloor \alpha \rfloor}$ series can be taken as a possible continuation of the $\lambda_{F,0}, \ldots, \lambda_{F,M}$ fractional FRTS. Figure 5 illustrates how the LSM fitting the best to the $\lambda_{F,0}, \ldots, \lambda_{F,M}$ fractional FRTS can be used as a predictor.

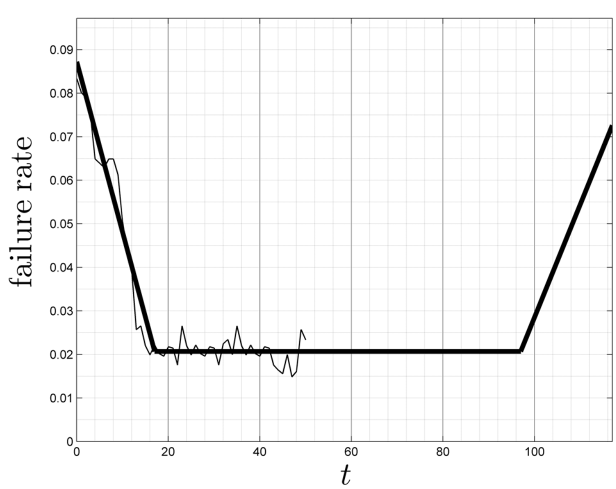

Figure 5. Predicting failure rates

3.5. *Discovering similarities among the FRTSs*

Similarities among the studied FRTSs can be characterized by using methods that are suitable to measure either the similarity or the distance among them. Having the appropriate similarity or distance metric identified, the FRTSs can be clustered based upon it. A good overview of the well-known time series clustering techniques is available in a study by Liao.[49]

In our case, the bathtub shape of the graphs of the studied FRTSs can be taken as their common property. The LSM of an FRTS is a representation of it with the h_1, λ_B, d, l, and h_2 parameters that describe the shape of the FRTS. Following this approach, the FRTSs can be clustered using the h_1, λ_B, d, l, and h_2 parameters, or if the length of FRTSs need to be considered, their n parameters can be taken into account as well. Furthermore, the σ_1, σ_2 and σ_3 RMSEs, which are measures of goodness-of-model fit, can also be considered as model parameters and used in clustering of FRTSs a multidimensional space.

Having the SLSM generated for each FRTS, the opportunity to use the SLMSs for clustering the FRTSs presents itself. Recall that the SLSM of an FRTS has the y_1, y_2, and x_l parameters, but it does not have any information about the position of the FRTS graph on the vertical axis. In other words, because of the construction of the LSMs, if there is a constant difference between corresponding values of two FRTSs, their LSMs are identical. On the other hand, if the λ_B parameters of the LSMs are considered besides the y_1, y_2, and x_l SLSM parameters, the vertical positions of the FRTS graphs are captured through the λ_B parameters. By this means, the y_1, y_2, x_l, and λ_B parameters allow us to cluster the FRTSs in a four-dimensional space. If the λ_B parameters are close to each other, we recommend normalizing them to the $[0, 1]$ interval and using the normalized values for clustering. The advantage of this approach, which we demonstrate in our case study, lies in the fact that the FRTSs can be clustered in a low-dimensional space.

4. Industrial application

The methods described in the preceding text were applied to 30 real-life FRTS of a consumer electronic product family with the purpose of studying their similarities and typifying them. The SLSM of each FRTS was generated according to Section 3. After that, the SLSMs were clustered – using the fuzzy C-means method – into seven clusters, based on their y_1, y_2, and x_l parameters. Figure 6 shows the clustering results including the cluster centroids that are marked by bigger bold symbols.

Figure 7 depicts the SLSMs (thin lines) and the cluster characteristic SLSM (thick lines) belonging to each cluster formed.

We can conclude from these images that the SLSMs within a cluster are similar, while the cluster characteristic SLSMs reflect the dissimilarities between the clusters well.

4.1. *Using SLSMs to predict failure rates*

The FRTS of a product type, which had not been involved in establishing the SLSMs, was selected to demonstrate how the cluster characteristic SLSMs can be used to predict the FRTS by applying the method discussed in Section 3.4.1. This FRTS represents the empirical failure rates of a consumer product type during 235 consecutive weeks; that is, the time series consists of 235 data. The SLSM-based prediction was compared with two widely applied forecasting methods: the moving average and exponential

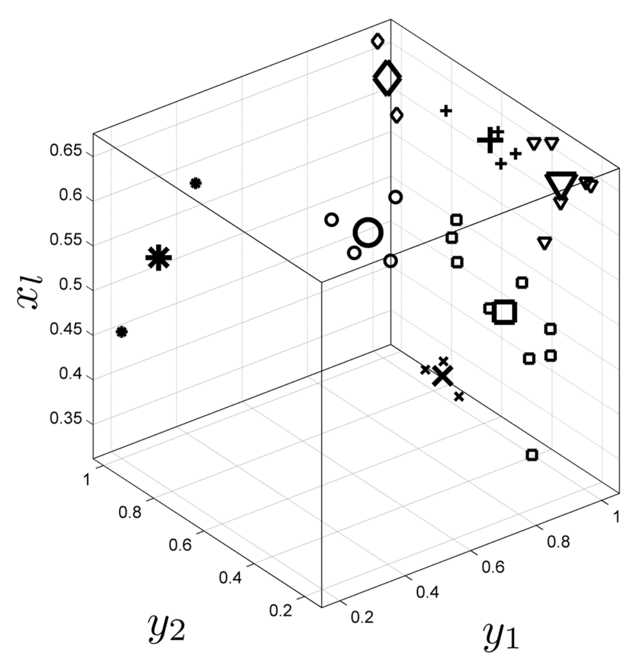

Figure 6. Clustered standard line segment models

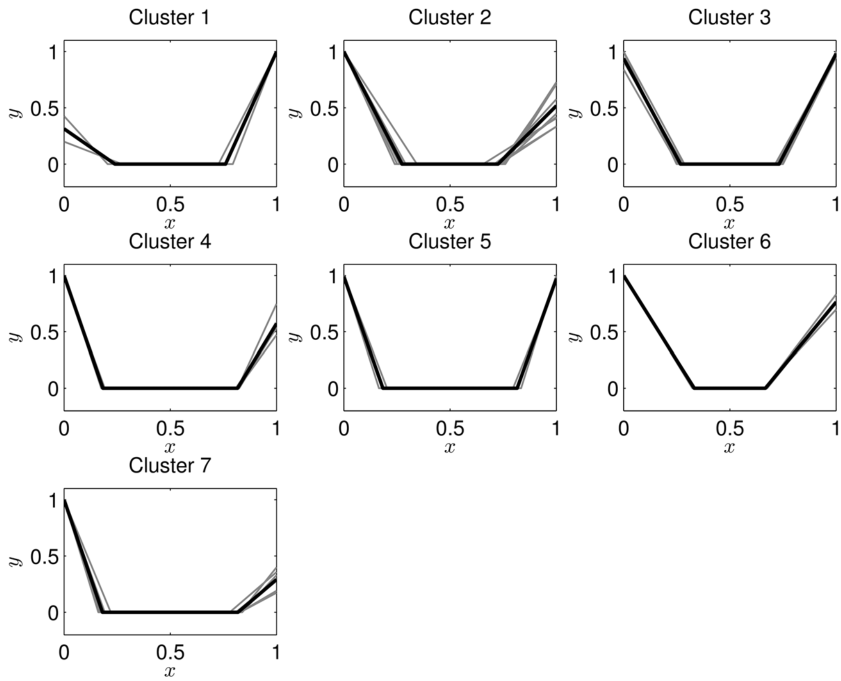

Figure 7. Standard line segment models and cluster characteristic standard line segment models

smoothing. The moving average was applied with a span of five, and default weight for the exponential smoothing was computed by fitting an ARIMA (0, 1, 1) model to the data, and back-casting was used to calculate the initial smoothed value. The mean squared error (MSE) of predicted failure rate was calculated for each forecast period to characterize the goodness of the applied forecasting methods. In order to evaluate the methods both in the short run and in the long run, we forecasted failure rates for 5, 10, 30, and 60 weeks (periods) based on the first 10-week-long, 30-week-long, 80-week-long, and 160-week-long known fractional FRTS, respectively. Results of these ex post predictions are summarized in Table I.

Predictions based on known FRTS fractions of first 10, 30, 80, and 160 weeks were also computed for the entire time range of FRTS, that is, for 225, 205, 105, and 85 weeks, respectively. Figure 8 depicts graph of the entire FRTS and the predictions founded on its different known fractions for its entire time range. In Figure 8, the solid gray lines, the dashed black lines, and the dashed gray lines indicate the SLSM-based, the moving average, and the exponential-smoothing forecasts, respectively.

Table II shows the MSE values of predicted failure rates calculated for 225, 205, 105, and 85 weeks of forecasts.

There are couples of notable properties of the SLSM-based predictions that can be seen in our example.

The best-fitting cluster characteristic SLSM changes as the known fraction of the studied FRTS develops; that is, predictions change as more information is presented to the prediction model.

When only the first 10 data of FRTS are known (left-upper graph in Figure 8), the moving average and exponential-smoothing forecasts differ very much from the real FRTS values. This is due to the fact that the failure rate curve is declining in this phase, and these two methods give constant forecasts that are computed based on the first 10 values of a decreasing series of

Table I. Results of ex post predictions for different time ranges based on different known FRTS fractions

			Fitting parameters of best SLSM			MSE of forecast (for the predicted time range)		
Best-fitting SLSM	Length of fractional FRTS	Number of forecast periods	Alpha	Beta	Gamma	SLSM	Moving average	Exponential smoothing
5	10	5	227.42	0.0116	0.0857	0.000086324	0.000327629	0.000264068
4	30	10	227.13	0.0069	0.0869	0.000020014	0.000107852	0.000051725
7	80	30	210.33	0.0207	0.0757	0.000015417	0.000016251	0.000024988
5	150	60	236.86	0.0198	0.0661	0.000016623	0.000046311	0.000050558

First column in Table I contains the index of best-fitting cluster characteristic SLSM for each prediction. Columns labeled with 'alpha', 'beta', and 'gamma' contain the α, β, and γ fitting parameters (described in Section 3.4.1), respectively, for the best-fitting cluster characteristic SLSMs. For example, cluster characteristic SLSM number 5 results the best fitting based on the fractional failure rate time series that contains failure rates of the first 10 weeks.
FRTS: failure rate time series; SLSM: standard line segment model; FRTS: failure rate time series; MSE; mean squared error.

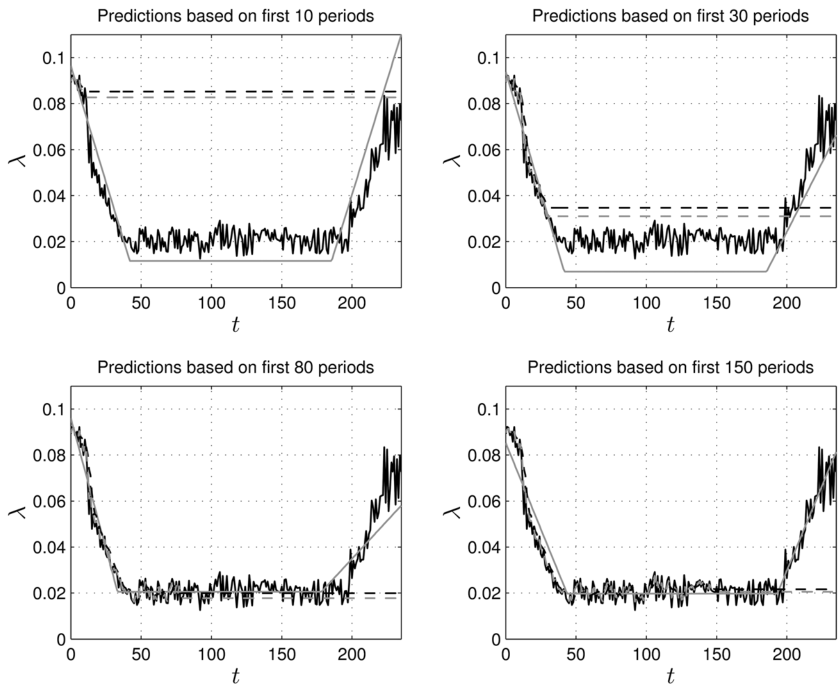

Figure 8. Ex post predictions based on different known factions of failure rate time series

Table II. Results of ex post predictions for entire time range of FRTS based on its different known fractions

			Fitting parameters of best SLSM			MSE of forecast (for the predicted time range)		
Best-fitting SLSM	Length of fractional FRTS	Number of forecast periods	Alpha	Beta	Gamma	SLSM	Moving average	Exponential smoothing
5	10	225	227.42	0.0116	0.0857	0.000152829	0.003419144	0.003017330
4	30	205	227.13	0.0069	0.0869	0.000161862	0.000390271	0.000210055
7	80	155	210.33	0.0207	0.0757	0.000047985	0.000381051	0.000265106
5	150	85	236.86	0.0198	0.0661	0.000035476	0.000361879	0.000226632

FRTS: failure rate time series; SLSM: standard line segment model; FRTS: failure rate time series; MSE; mean squared error.

approximately 40 values. One advantage of the SLSM-based prediction is that it gives a long-term forecast and so provides an indication on turning points of the failure rate curve. Certainly, goodness of this indication is becoming better as more data of the curve are becoming known. In Tables I and II, the MSE values for predictions based on the first 10 and first 30 known data, that is, for periods before the first turning point of FRTS, indicate that the SLSM-based prediction is considerably better than the other two.

As the failure rate curve approaches its second, quasi-constant phase, the moving average and exponential smoothing deliver similar results to the SLSM-based forecast both in the short run and in the long run.

The MSE values calculated for the predicted failure rates (Tables I and II) indicate that the SLSM-based predictions give better results in each case that was studied in our example. From a practical point of view, we also need to mention that in the quasi-constant phase of the failure rate curve, the differences among the three methods in terms of their goodness are small, and these differences are more dominant in the declining and rising phases of the curve, especially in the long run.

4.2. Grouping the original empirical failure rate curves

If we wish to discover similarities among FRTSs considering not just their shape, which can be grasped through their SLSM parameters, but their positions along the vertical axis as well, we can use the approach described in the succeeding text.

The FRTSs were clustered in the four-dimensional space based on their SLSMs and LSMs considering the y_1, y_2, x_I, and the normalized λ_B parameters. As the result of the fuzzy C-means clustering applied, 10 clusters of FRTSs were generated that are shown in the images cluster 1 to cluster 10 (Figure 9)

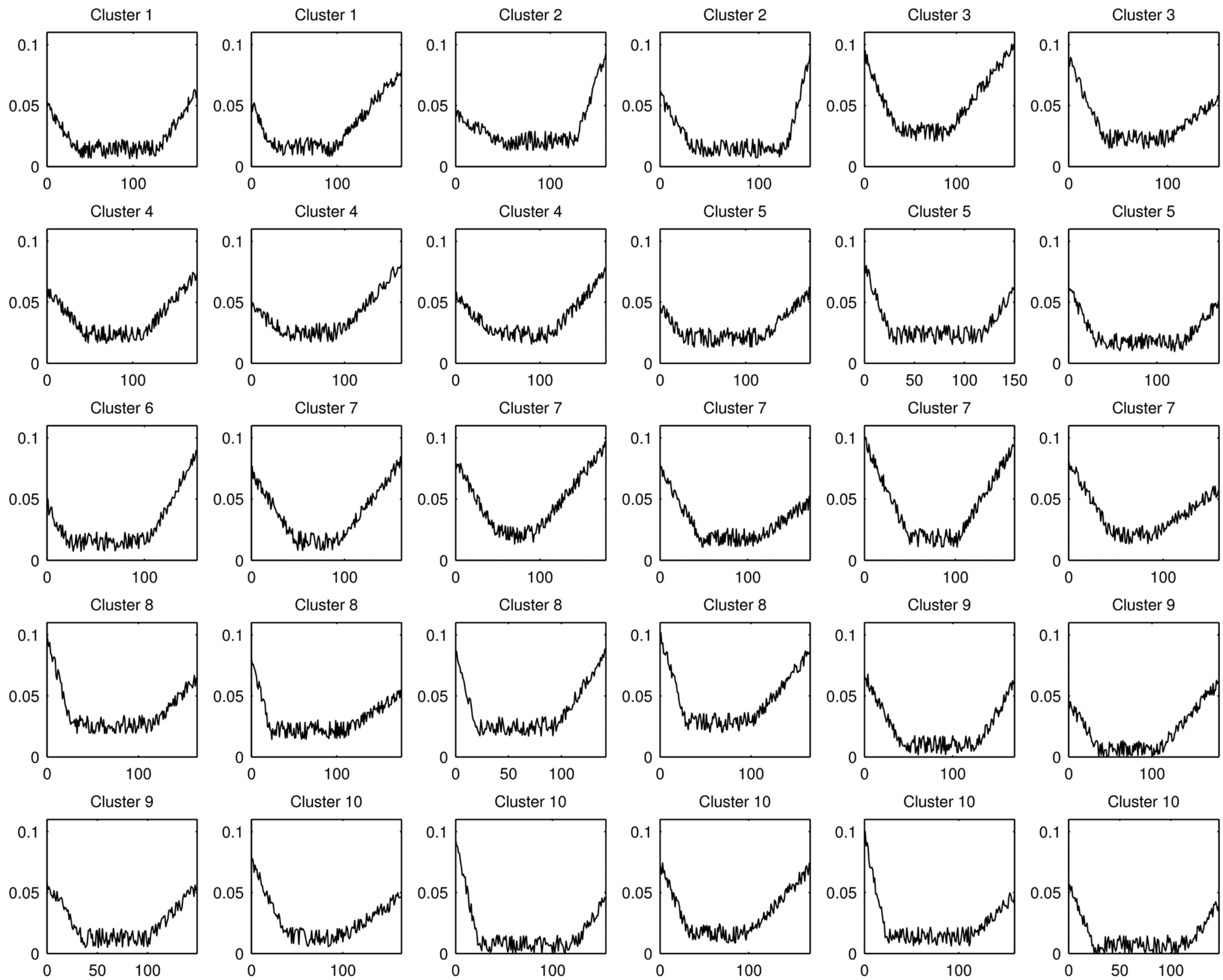

Figure 9. Clustered failure rate time series

The clusters formed and shown in the images cluster 1 to cluster 10 highlight that clustering the FRTS based on the SLSMs parameters, considering the λ_B parameter of each SLSM too, results in a grouping of FRTSs that reflects the similarities of FRTS within the groups and dissimilarities between the groups well.

5. Conclusions, managerial implications, and further research plans

In this paper, we presented a hybrid technique for modeling and forecasting bathtub-shaped failure rate data of newly marketed consumer electronics by studying repair services. Empirical failure rate curves describing the whole life cycles of on-the-market electronic products considered as time series were applied as inputs for typifying FRTSs, based on which future reliability of new-generation products can be predicted. Similarities among historical FRTSs are characterized by SLSM model parameters. These parameters describe an FRTS as three connected line segments, referring to the corresponding sections of the traditional bathtub curve. The SLSMs of FRTSs are clustered based on the model parameters as well as the FRTSs using the same parameters supplemented with the normalized λ_B parameter.

In spite of the fact that SLSM with its y_1, y_2, and x_l parameters is a simple representation of an FRTS, it can be used to capture the main properties of FRTS of consumer electronic products. The model parameters of an FRTS carry information about the shape of the failure rate curve, and so, clustering the SLSMs results in a typical SLSM that can be applied for predicting the future values of the failure rate curves. In this sense, the cluster characteristic SLSMs represent the knowledge of failure rate curves gained from historical data. From a managerial perspective, discovering similarities among empirical failure rate curves generates value both for the repair service providers and for the original manufacturers. While ERSP companies can use it to predict resource needs for particular repair services, the latter ones can conclude on typical reliability characteristics of their products.

As the parameters of the model have a geometric interpretation as well, the proposed model has a cognitive aspect.

The case study presented in this paper justifies that the SLSMs along with the λ_B parameters can be used to discover similarities of the studied failure rate curves. The introduced method is founded on a model that is mathematically simple but suites well the needs of industrial applications.

A comparison was made to moving average and exponential smoothing, and it was observed that the results obtained from our modeling are encouraging, and it has the potential to be a suitable alternative predicting technique.

The model's flexibility is twofold. On the one hand, in the case of a newly marketed electronic product, we can obtain further information about the actual failure rates in the course of time. From time to time, the typical failure rate curve that fits the best to the actual curve may change. This kind of fitting by moving forward in the product life cycle can change and can be revised according to new information. On the other hand, by producing more and more products, the database of curves can be complemented, and historical failure rate curves can be clustered again, resulting in the re-identification of typical curves as time goes on.

We should emphasize that this approach is founded on the phenomenon that the empirical failure rate curves, for which the method is applied, are bathtub-shaped ones. We do not know yet if the same method would be appropriate for other product types as well. One of our future research plans is to answer this question.

In the presented approach, we focused solely on properties of the failure rate curves in order to capture similarities among them. We have not taken into consideration the functional and technological properties of the studied products; however, one can say that products with similar functional and technological parameters presumably have similar failure rate curves. Starting from this thought, we plan to study whether products that are functionally and technologically similar also show similarities in their SLSM parameters. For this purpose, our future research plan is to study how effectively the product SLSMs (or LSMs) could be typified according to clustering criteria learnt from clustering them based on their functional and technological attributes using self-organizing feature maps.[50]

Acknowledgements

The authors are grateful to all anonymous referees whose comments and suggestions have significantly improved our original version of this paper. This study was partially supported by TÁMOP-4.2.2.C-11/1/KONV-2012-0005, Well-being in the Information Society.

References

1. Goel A, Graves RJ. Electronic system reliability: collating prediction models. *IEEE Transactions on Device and Materials Reliability* 2006; **6**:258–265. doi:10.1109/TDMR.2006.876570.
2. Economou M. The merits and limitations of reliability predictions. Reliability and Maintainability, 2004 Annual Symposium - RAMS, 26-29 Jan, 2004, 352–357.
3. Campbell DS, Hayes JA, Jones JA, Schwarzenberger AP. Reliability behaviour of electronic components as a function of time. *Quality and Reliability Engineering International* 1992; **8**:161–166. doi:10.1002/qre.4680080303.
4. Kita M. How the EMS (electronics manufacturing service) business model can help Japanese corporations revolutionize their factories? *JBIC Review* 2001; **4**:1–24.
5. Harms JW. Revision of MIL-HDBK-217, Reliability prediction of electronic equipment, Reliability and Maintainability Symposium (RAMS), 2010 Proceedings - Annual, San Jose, CA, 25–28. Jan, 2010
6. Zhai E, Shi Y, Gregory M. The growth and capability development of electronics manufacturing service (EMS) companies. *International Journal of Production Economics* 2007; **107**:1–19.
7. Barnes E, Dai J, Deng S, Down D, Goh M, Lau HC, Sharafali M. Electronics Manufacturing Service Industry. National University of Singapore: Georgia Institute of Technology, 2000.
8. Salvador F, Forza C, Rungtusanatham M. Modularity, product variety, production volume, and component sourcing: theorizing beyond generic prescriptions. *Journal of Operations Management* 2002; **20**:549–575. doi:10.1016/S0272-6963(02)00027-X.
9. Sherman R. The Worldwide Electronics Manufacturing Services Market. Electronic Trend Publications, New Venture Research Corp: Nevada, CA, 2013.
10. Leung A, Yip G. The global OEM: the transformation of Asian supplier companies. In: The Future of Multinational Companies Birkinshaw et al. (eds) Chichester: Wiley UK, 2003, 86–99.
11. Lai CD, Xie M, Murthy DNP. Bathtub shaped failure rate life distributions. *Handbook of Statistics* 2001; **20**:69–104.
12. Lai CD, Xie M. Stochastic Ageing and Dependence for Reliability. Springer: New York, 2006.
13. Block HW, Savits TH. Burn-in. *Statistical Science* 1997; **12**:1–19.
14. Block HW, Jong YK, Savits TH. Bathtub functions and burn-in. *Probability in the Engineering Sciences* 1999; **13**:497–507.
15. Rajarshi S, Rajarshi MB. Bathtub distributions: a review. *Communications in Statistics, Part A-Theory and Methods* 1998; **17**:2597–2621.
16. Pham H, Lai CD. On recent generalizations of the Weibull distribution. *IEEE Transactions on Reliability* 2007; **56**:454–458.
17. Silva RB, Barreto-Souza W, Cordeiro GM. A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Computational Statistics and Data Analysis* 2010; **54**:935–944.
18. Paranaíba PF, Ortega EMM, Cordeiro GM, Pescim RR. The beta Burr XII distribution with application to lifetime data. *Computational Statistics and Data Analysis* 2011; **55**:1118–1136.
19. Liu J, Wang Y. On Crevecoeur's bathtub-shaped failure rate model. *Computational Statistics and Data Analysis* 2013; **57**:645–660.
20. Cancho VG, Louzada-Neto F, Barriga GDC. The Poisson-exponential lifetime distribution. *Computational Statistics and Data Analysis* 2011; **55**:677–686.
21. Hemmati F, Khorram E, Rezakhah S. A new three-parameter ageing distribution. *Journal of Statistical Planning and Inference* 2011; **141**:2266–2275.
22. Mudholkar GS, Asubonteng KO, Hutson AD. Transformation of the bathtub failure rate data in reliability for using Weibull-model analysis. *Statistical Methodology* 2009; **6**:622–633.

23. Bebbington M, Lai CD, Zitikis R. Estimating the turning point of a bathtub-shaped failure rate distribution. *Journal of Statistical Planning and Inference* 2008; **138**:1157–1166.
24. Hallberg Ö. Failure-rate as a function of time due to log-normal life distribution(s) of weak parts. *Microelectronics and Reliability* 1977; **16**:155–158.
25. Govil KK, Aggarwal KK. Modelling early, chance and wearout failures by a single failure rate equation. *Reliability Engineering* 1982; **3**:339–343.
26. Casanelli G, Mura G, Cesaretti F, Vanzi M, Fantini F. Reliability predictions in electronic industrial applications. *Microelectronics Reliability* 2005; **45**:1321–1326.
27. Lee SW, Lee HK. Reliability prediction system based on the failure rate model of electronic components. *Journal of Mechanical Science and Technology* 2008; **22**:957–964.
28. Liang YH. Combining neural networks and genetic algorithms for predicting the reliability of repairable systems. *International Journal of Quality and Reliability Management* 2008; **25**:201–210.
29. Liang YH. Analyzing and forecasting the reliability for repairable systems using the time series decomposition method. *International Journal of Quality and Reliability Management* 2011; **28**:317–327.
30. Tong LI, Liang YH. Forecasting field failure data for repairable systems using neural networks and SARIMA model. *International Journal of Quality and Reliability Management* 2005; **22**:410–420.
31. Ascher H, Feingold H. Repairable System Reliability. Marcel Dekker: New York, NY, 1984.
32. Chatterjee S, Misra RB, Alam SS. Prediction of software reliability using an autoregressive process. *International Journal of System Science* 1997; **28**:211–216.
33. Burtschy B, Albeanu G, Boros DN, Popentiu F, Nicola V. Improving software reliability forecasting. *Microelectronics and Reliability* 1997; **37**:901–907.
34. Xie M, Hong GY, Wohlin C. A study of exponential smoothing technique in software reliability growth prediction. *Quality and Reliability Engineering International* 1997; **13**:347–353.
35. Billinton R, Chen H, Ghajar R. Time series models for reliability evaluation of power systems including wind energy. *Microelectronics and Reliability* 1996; **36**:1253–1261.
36. Box GEP, Jenkins GM. Time Series Analysis Forecasting and Control. Holden-Day: San Francisco, CA, 1997.
37. Xie M, Ho SL. Analysis of repairable system failure data using time series models. *Journal of Quality in Maintenance Engineering* 1996; **5**:50–61.
38. Liang YH, Tong LI. Analysing and forecasting the field failure data for repairable systems. *Journal of Quality* 2001; **8**:75–96.
39. Healy J. A simple procedure for reliability of repairable systems. Proceedings of the 1997 Annual Reliability and Maintainability Symposium, 1997, 171–175.
40. Wang WL, Hemminger TL. A moving average non-homogeneous Poisson process reliability growth model to account for software with repair and system structures. *IEEE Transactions on Reliability* 2007; **56**:411–21.
41. Chen KY. Forecasting systems reliability based on support vector regression with genetic algorithms. *Reliability Engineering & System Safety* 2007; **92**:423–32.
42. Xue K, Xie M, Tang LC, Ho SL. Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing* 2003; **2**:255–268.
43. Ho SL, Xie M, Goh TN. A comparative study of neural network and Box-Jenkins ARIMA modelling in time series prediction. *Computer & Industrial Engineering* 2002; **42**:371–375.
44. Al-Garni AZ, Jamal A. Artificial neural network application of modelling failure rate for Boeing 737 tires. *Quality & Reliability Engineering International* 2011; **27**:209–219.
45. Dunn JC, Harshbarger S. Conditional gradient algorithms with open loop step size rules. *Journal of Mathematical Analysis and Applications* 1978; **62**:432–444.
46. Jaggi M. Revisiting Frank-Wolfe: projection-free sparse convex optimization. *Journal of Machine Learning Research: Workshop and Conference Proceedings* 2013; **28**:427–435.
47. Bezdek JC. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press: New York, 1981.
48. Chiu S. Fuzzy model identification based on cluster estimation. *Journal of Intelligent & Fuzzy Systems* 1994; **2**:267–278. doi:10.3233/IFS-1994-2306.
49. Liao TW. Clustering of time series data — a survey. *Pattern Recognition* 2005; **38**:1857–1874.
50. Kohonen T. Self Organizing Maps. Springer Series in Information Sciences. Springer: Berlin, 2001.

Authors' biographies

József Dombi (66 years old) is full professor at University of Szeged. His academic degrees are as follows: university doctor's degree (1977, summa cum laude), Candidate of Mathematical Sciences (1994, CSc, title of dissertation: Fuzzy sets' structure from the aspect of multi-criteria decision aid), Doctor of Hungarian Academy of Sciences (2011). His editorial memberships are as follows: editorial board member of the Foundations of Computing and Decision Sciences magazine until 1997, editorial board member of the *International Journal of Systems Science*, and editorial board member of the *International Journal of Advanced Intelligence Paradigms*. His research interests are as follows: computational intelligence, theory of fuzzy sets, multi-criteria decision-making, genetic and evolutional algorithms, operation research, and visualization. His teaching interests are as follows: information visualization, which has an enormously growing interest in Hungary and abroad, the theory of artificial intelligence, fuzzy systems, multi-criteria decision-making, and intelligent systems.

Tamás Jónás (42 years old) is Research and Development Manager at Flextronics International Ltd, Budapest, and Associate Professor at the Department of Management and Corporate Economics, Budapest University of Technology and Economics (BUTE), where he teaches reliability theory and quantitative methods. His research interest includes quantitative modeling of business processes. He earned his Master of Computer Science degree at the University of Szeged and his MBA degree at BUTE. He earned his PhD in Business Economics in 2011.

Zsuzsanna Eszter Tóth (35 years old) is Associate Professor at the Department of Management and Corporate Economics, BUTE. Her researches concentrate on the application of reliability management tools. She teaches applied quantitative methods and quality and reliability management in undergraduate and graduate levels. She earned her PhD in Business Economics in 2008 at BUTE.