Simulation based on the Gompertz Curve Model to Assess Early Systematic and Mechanical BEEMS Failures

Takehiro Tanaka

1Toyo University, Saitama, 350-8585, Japan
*Corresponding author’s e-mail: tanaka@toyo.jp

Abstract. The purpose of this study is to introduce the simulation procedure for assessing early systematic and mechanical failures of the Building, Environment and Energy Management System (BEEMS) to prevent the newly-installed system from incurring major damage, troubles and risks. Here, the Gompertz curve model is simulated using the unit of measurement called “mean time between failures (MTBF)” to evaluate failure rates. The research focuses are placed on the types of facilities and devices, the number of failure cases, the Mean Time Between Failures (MTBF), estimations on the basis of the Weibull analysis, the Bayesian method and the Gompertz curve model. Especially, it was confirmed to be effective to use the Gompertz curve model together with the Bayesian method for data of 20 failures or less. And the Gompertz curve model would be appropriate for the estimation of a regular inspection period more precisely under several assumptions based on a limited number of data as its lower limit of probability could deliver some margin for safety.

1. Introduction
The purpose of this study is to introduce the simulation procedure for assessing early systematic and mechanical failures of the Building, Environment and Energy Management System (BEEMS) to prevent the newly-installed system from incurring major damage, troubles and risks. Here, the Gompertz curve model is simulated using the unit of measurement called “mean time between failures (MTBF)” to evaluate failure rates.

First, it uses the Weibull distribution which has taken a very important role in reliability analysis and the Weibull type cumulative hazard method According to the Weibull analysis of failure data the characteristics of reliability, that is, MTBF, failure rate and so on are estimated. Estimation of the MTBF for Random Failure is the basic subject to be studied, and much discussion of such subject is being made at present in connection with the building equipments. A close examination of the interval estimation of MTBF is indispensable in establishing the maintenance periods of the Building, Environment and Energy Management System (BEEMS), building automation system (BAS), and so on.

The second, it proposes the Gompertz curve model. Reliability is always improving. This concept is called reliability growth. In this paper, the Gompertz curve model is adopted. The Gompertz curve is able to express the total number of failures detected over an operating time in the form of a sigmoid growth curve. And the Bayesian method can be characterized in respect of using the prior information obtained from the past information jointly with the data in estimation and inspecting the parameter which is important for the estimation of failure rate.

In this study, an investigation of MTBF estimation for random failure is made using field data of building equipment operations, and a comparison of results is conducted by applying two methods of
interval estimation. The Bayesian method is used to estimate the failure rate, and then analyze it by using the data of building equipment and devices for a period of random failure. The simulation of maximum-likelihood method in comparison with the Bayesian method in estimating the failure rate is performed. And it is effective to use the Gompertz curve model. Like the Bayesian method. In this study, it touches upon the relationship of the analysis results achieved by the Bayesian approach with the assessed values of practical operations.

Especially, it evaluates the comparison in heat pump, chill unit, return fan and power facilities. These results suggest the most suitable and practical way of MTBF estimation for actual maintenance sites. So it is effective to use the Gompertz curve model. Like the Bayesian method for within 30 failure data.

2. Weibull type cumulative hazard method
Cumulative hazard method is employed to analyze the data which contain broken off data or missing data. Weibull type cumulative hazard method is adopted in the study to deal which the failure/maintenance data. The Weibull distribution is given by the following.

\[ R(t) = \exp \left( -\left( \frac{t}{\eta} \right)^{\beta} \right) = \exp \left( -\left( t - \eta \right)^{\beta} \right) \]

(1)

Where \( R(t) \) is a reliability function, \( \beta \) is the shape parameter and \( \eta \) is scale parameter, and \( \beta \) and \( \eta \) are important parameters for the Weibull analysis. The characteristics about \( \beta \) are summarized as follows.

(1) \( \beta < 1 \): Decreasing Failure Rate (DFR); Initial Failure Type
In this case, the failure rate decreases with time, and preventive maintenance is very harmful.

(2) \( \beta = 1 \): Constant Failure Rate (CFR); Random Failure Type
In this case, the failure rate is constant with time, which means the exponential distribution, and preventive maintenance is not effective.

(3) \( \beta > 1 \): Increasing Failure Rate (IFR); Wear-out Failure Type
In this case, the failure rate increases with time, and preventive maintenance is considerably useful.

Consequently, judging from above results, it is very important and convenient to obtain the estimate from the field data through the Weibull analysis.

3. Gompertz curve model and logistic curve model
This study focuses on the phenomenon often observed in a number of devices, where a failure rate low at the beginning increases midway and then decreases again. The Gompertz curve and the logistic curve are two typical growth curves. The both are characterized by similar S-shaped curves, increasing with time \( x \) until reaching a given value \( K \). The logistic curve displays bilateral symmetry about its inflection point whereas it is characteristic of the Gompertz curve not to exhibit symmetry. Although it is not appropriate for the value of time to be negative, these curves are used by converting the time value into \( t = x + t_0 \) as \( t \geq 0 \).

3.1. Gompertz curve model
Reliability is always improving. This concept is called reliability growth. In this paper, the Gompertz curve model is adopted. The Gompertz curve is able to express the total number of failures detected over an operating time in the form of a sigmoid growth curve. In another word, it is capable of offering quality assessment criteria, by estimating the number of failures with the description of a temporal tendency of an over-all operating period, where the number of failures detected is small at the beginning, rapidly increasing in the middle of testing time and reaching a saturation level through the last stage.

The Gompertz curve is applied to failure data and its suitability has been pointed out in many cases. The Gompertz curve is expressed with the following formula.
This formula is derived from the differential equation, $\frac{dy}{dx} = Ay \times e^{-Bx}$.

Here, $\frac{dy}{dx}$ represents the growth rate of $y$. Specifically speaking, it means that the growth rate of $y$ is related to the factor which increases in proportion to the previous value of $y$ and to the factor which decreases exponentially with the time $x$. This is the characteristic feature of the Gompertz curve.

![Figure 1. The Case of Gompertz curve model](image)

Figure 1. is completed with $K=1$ and varied values of $b$ and $c$.

When $x \to \infty$: then, $e^{-cx} \to 0$, $b_0=1$, and $y \to K$

$x=0$: $e^{-cx}=1$ and $y=Kb$

$x \to -\infty$: $e^{-cx} \to \infty$ and $y \to b_\infty \to 0$ ($0 < b < 1$ is a prerequisite)

According to the graph in red with $b=0.3$ and $c=0.5$

for example, when $x_0(x=0)$, then $y_0=Kb=1 \times 0.3=0.3$.

Here, $b$ is a factor to determine the value of $y_0$, and $c$ indicates the degree of inclination.

Based on the equation $y = \frac{K}{1 + be^{-cx}}$ where $y$ comes to be equal to $aK$, $x$ is obtained from the following formula.

$x = (1/c) \times \log(\log b/\log a)$

where $b=0.3$ and $c=0.5$ are fixed, various values of $a$ lead to the following results.

$x = 1.104 4.872 6.312 9.572$

Since this curve represents the cumulative number of failures, 95% of failures is expected to be detected at $x=6.3$.

The NHPP model, which is a typical reliability growth model, assumes NHPP with the following properties in response to the detection of failures described in the counting process $\{N(t), t \geq 0\}$.

1. $N(0) = 0$
2. The counting process $\{N(t), t \geq 0\}$ contains independent increments.
3. $Pr(N(t+\Delta t) - N(t) = 1) = h(t) \Delta t + O(\Delta t)$
4. $Pr(N(t+\Delta t) - N(t) \geq 2) = O(\Delta t)$

Where, $Pr \{A\}$ expresses the probability of a phenomenon, $A$, to occur, while a function, $o(\Delta t)$, indicates that effects of the second order or higher in micro time, $\Delta t$, can be ignored, and the following Eq. (7) is obtained.
\[
\lim_{t \to \infty} \frac{O(\Delta t)}{\Delta t} = 0
\]

The item \(h(t)\) is indicating the rate of error detection, is called an NHPP intensity function. When \(t\) and \(x\) are considered to be the total operating hours executed and the designated length of testing time respectively, the degree of reliability achieved at a testing time, \(t\), can be evaluated. Although the mean time of failures (MTBF) between \((n)\) and \((n+1)\), \(E[S_{n+1} - S_n]\), does not exist, which is therefore unable to be obtained theoretically, the cumulative MTBF and instantaneous MTBF can be alternatively derived.

3.2. Logistic curve model

The logistic curve model is expressed with the following formula.

\[
y = \frac{K}{1 + be^{-cx}}
\]

Fig. 2 shows the graph drawn on the basis of the logistic equation. Because living things are born to their parents, the growth rate of population \(dy/dt\) is supposed to be proportional to the previous population size \(y\).

![Logistic curve model](image)

**Figure 2.** The Case of Logistic curve model

Thus, \(dy/dt = ry\) holds true. In reality, however, the growth does not infinitely continue.

Getting closer to its limit \(K\), the growth rate decreases. As this assumption is taken into account, \(dy/dt = ry(K-y)\) is obtained.

The logistic equation is derived from this relation.

From the viewpoint of biological growth, \(e \geq 0\), and \(y = 0\) when \(t = t_0\). But to make it generalized, this prerequisite is ignored in the logistic equation. Hence, \(x = t + t_0\) and \(t_0 = -\infty\). This mathematically indicates the birth of living things in times long past.

According to the logistic equation,

\[
\begin{align*}
\text{when } x & \to -\infty: \text{ then, } e^{-cx} \to 0, \text{ denominator} \to 1 \text{ and } y \to K \\
\text{and } x & = 0: e^{-cx} = 1, \text{ denominator} = 1+b \text{ and } y = K/(1+b) \\
\text{and } x & \to \infty: e^{-cx} \to 0 \text{ and } y \to 0
\end{align*}
\]

Based on the differential equation, \(dy/dt\) (growth rate) is maximized at \(y^* = K/2\). From the logistic equation, \(x\) at that point is found to be \(x^* = (\log b - \log 2)/c\).

\[
y = \frac{K}{1 + be^{-cx}}
\]

Based on the equation \(y = aK\) where \(y\) comes to be equal to \(aK\), \(x\) is obtained from the following formula.
\[
\frac{\log b - \log((1/\alpha)-1)}{x} = \frac{x}{c}
\]

where, \( \alpha = \log((1/\alpha)-1) \)

| \( \alpha \) | \( \log((1/\alpha)-1) \) |
|-------------|------------------|
| 0.5         | 0                |
| 0.9         | -2.197           |
| 0.95        | -2.944           |
| 0.99        | -4.595           |

The graphs are drawn with \( K=1 \) and varied values of \( b \) and \( c \). According to the graph in red with \( b=3 \) and \( c=1 \), when \( x=0 \), then \( y_0=1/(1+3)=2.5 \). The growth rate is maximized at \( x^*=\frac{\log 3 - \log 2}{1}=1.098-0.693=0.405 \) and \( y^*=1/2=0.5 \). It is also understood that \( y_0 \) is determined by \( b \) and the inclination varies depending on \( c \).

4. Theoretical background of assessment in random failure

The concept of random failure differs depending on respective building facility system or devices as there is no definite criterion at present. However, any failure which occurs during the period after elapse of the usual one-year guarantee period and before start of abrasion failure may be considered as the random failure.

The conceptual maintenance curve is well known as "bathtub" curve. But in reality the function conditions are changing upon the regular inspection conditions during the period of random failure. It is attempted in this study to estimate the interval between the upper limit value and the lower limit value of MTBF in order to maintain the wholesome functions required for building equipments. Then it is examined and compared with application of the field data of actual building equipments by two methods of interval estimation. Moreover, it refers to the relationship between the analysis respect by the Bayesian approach and the assessed values.

To perform the interval estimation of MTBF in exponential distribution, it is necessary to estimation the interval from the assessed values. In other words, there exists a certain value in the interval of the upper and lower limits of MTBF, and the interval extinction is to estimate the change of the value between those limit. The probability for existence of MTBF within that interval is called the assessed value.

In most cases, the maintenance data of building facilities are the closing data. In fact, there are many cases in conducting sample analysis of the time. The closing is generally divided into two categories; one is fixed time closing and the other is fixed number closing. \( c^2 \) (chi-square) distribution is used for interval estimation. In the cases of fixed time closing and fixed number closing, \( c^2 \) distribution, assuming the assessed values as \( (1-a) \), multipliers to the value calculated from MTBF=\( T/r \) (where \( T \) is total operation time, and \( r \) is total number of failure) by \( c^2 \) (2r, a), \( 2r \), in \( c^2 \) (2r, a) denotes the degree of freedom of \( c^2 \) distribution and \( a \), the probability \( ( \text{ratio of risk} ) \) that goes out of the interval.

The Bayesian method can be characterized in respect of using the prior information obtained from the past information jointly with the data in estimation and inspecting the parameter which is important for the estimation of failure rate.

The different between the conventional method and the Bayesian method can be summarized as follows.

1. With the use of prior information, the Bayesian method will work with a small number of data in obtaining the same accuracy compared with the case of conventional sampling distribution method. And it can alleviate the problem of decreasing number of the data, when subdividing the item-classifications of the field data at the site.

2. The Bayesian method provides wider scope of applications being capable of using the past information reflected as the prior information despite the population distribution is changed due to the improvement of system, facilities or devices. Changed due to the improvement of system, facilities or
devices. Here, the expansion of the Bayesian method for random failure (period) is shown. In this study, the probability density function of the population distribution is considered as the exponential one. That is, given with $f(t|\alpha) = \exp(-\alpha t) (t \geq 0, \alpha > 0)$, and when evaluating ex post facto distribution $g(\alpha|T)$ of $L$ by the Bayesian theorem after data $T=\{t_1,\ldots,t_n\}$ is observed, it becomes, in the above equation, ex post facto distribution $g(\alpha|T)$ becomes gamma distribution of the parameter $(\alpha|T), (\beta + n)$. By using the equation, one can estimate failure rate by the Bayesian method.

In the study, the random failure period is assumed. A maximum of twenty-two random numbers are generated, a total of 30 piece observatory data are input, and five closing data corresponding to three, five, ten, fifteen, 20, 25 and 30 are set per year, respectively. Furthermore, the failure rate or ex post factor distribution. Finally, the calculated results are examined and with the result of the conventional maximum-likelihood method to obtain the following discussions.

The difference between the result by the maximum-likelihood method and the Bayesian method becomes small with the growth of the number of data but when the number of the available data is less than 20, the Bayesian method has to be used. As for the MTBF, the difference between the results by the two methods is not large at all. Therefore for newly completed building facility system and devices. The Bayesian method should be applied.

5. Simulation based on random numbers in the period of random failure

By way of experiment, a hypothetical building facility (e.g. a multi air-conditioning system) that had not been maintained for 36 months was simulated in the following manners, using 500 pieces of data with 50 pieces of random numbers being generated ten times. Here, the Weibull Type Cumulative Hazard Method and the Gompertz curve model were employed as the authors have already reached the conclusion that simulations on the basis of the Bayesian Method are effective when the number of data is 20 or less.

1) Random numbers of the Weibull Type Cumulative Hazard Method and the Gompertz curve model were assumed to be normally distributed. Therefore the lower limit of probability was set at 15.9% which was the 3σ value of normal distribution.

2) Random numbers of the Weibull Type Cumulative Hazard Method and the Gompertz curve model were assumed to be normally distributed. Therefore the upper limit of probability was set at 84.1% which was the 3σ value of normal distribution.

Fifty pieces of random numbers were put together into one group. Simulation results of Groups 1 through 10 are shown in Table 1, where numbers are expressed in months.

Table 1. The time of failure occurrence after the system is left unmaintained for 60 months (Weibull)

| Group | Lower limit | 50% point | Upper limit |
|-------|-------------|------------|-------------|
| Group1 | 29M 0D | 30M 2D | 31M 3D |
| Group2 | 29M 2D | 29M28D | 30M23D |
| Group3 | 29M 0D | 29M29D | 30M27D |
| Group4 | 29M 4D | 29M29D | 30M24D |
| Group5 | 28M29D | 29M27D | 30M24D |
| Group6 | 28M26D | 29M29D | 31M 3D |
| Group7 | 29M 3D | 29M29D | 30M26D |
| Group8 | 28M29D | 30M 1D | 31M 2D |
| Group9 | 29M 1D | 30M 2D | 31M 4D |
| Group10 | 28M28D | 29M25D | 30M22D |
| Mean | 29M 0D | 29M29D | 30M28D |

※Explanatory notes: Months: M, Days: D

The result shown in Table 1 indicate: supposing that a certain construction facility is not or unable to be maintained for 60 months, the first failure is expected to occur in 29 months at the lower limit of probability, 29 months and 29 days at the 50% point, and 30 months and 28 days at the upper limit of probability. That is to say, the facility must be inspected at least within 29 months.
Table 2. The time of failure occurrence after the system is left unmaintained for 60 months (Gompertz)

| Group | Lower limit | 50% point | Upper limit |
|-------|-------------|-----------|-------------|
| Group1 | 28M 7D      | 30M11D    | 31M 4D      |
| Group2 | 29M 0D      | 30M 6D    | 30M24D      |
| Group3 | 28M27D      | 30M 7D    | 30M28D      |
| Group4 | 29M 2D      | 30M 6D    | 30M24D      |
| Group5 | 28M27D      | 30M 5D    | 30M24D      |
| Group6 | 28M23D      | 30M 9D    | 31M 3D      |
| Group7 | 29M 1D      | 30M 7D    | 30M26D      |
| Group8 | 28M24D      | 30M11D    | 31M 4D      |
| Group9 | 28M28D      | 30M11D    | 31M 4D      |
| Group10| 28M26D      | 30M 3D    | 30M22D      |
| Mean   | 28M28D      | 30M 7D    | 30M28D      |

※Explanatory notes: Months: M, Days: D

The result shown in Table 2 indicates: supposing that a certain construction facility is not or unable to be maintained for 60 months, the first failure is expected to occur in 28 months and 28 days at the lower limit of probability, 30 months and 7 days at the 50% point, and 30 months and 28 days at the upper limit of probability. That is to say, the facility must be inspected at least within 28 months.

Ten data groups of which each consisted of 50-piece data were simulated in this experiment. Although no significant differences were obtained, we considered that the Gompertz curve model would be appropriate for the estimation of a regular inspection period more precisely from the limited number of data as its lower limit of probability could deliver some margin for safety.

6. Application to building equipment--mainly estimate the comparison between many facilities--

The data used in this study are the results of an analysis of four systems whose total operating hours could be clearly determined among the air-conditioning devices and the electrical installations of a certain building. Table 3 summarizes the systems chosen, names of equipment and devices, number of cases, and MTBF.

In this study, the MTBF estimation for random failure is investigated using the field data of building equipment operations and the comparison of the results is conducted by applying two methods for the interval estimation. It analyzed with the data acquired in thirty facilities for the office building on the Tokyo Japan. The author because all those facilities are completed in Oct. of 1995 and their age is the 18 years in Sep. of 2013. Especially it estimates the comparison between heat pump system and other building system. These results suggest the most suitable and practical way of MTBF estimation for actual maintenance sites.

Estimation of MTBF of the period of random failure is examination and compared by applying the interval estimation methods of fixed time closing and fixed number closing, with the use of actual building facilities. In this paper, it shows interval estimation methods of fixed number closing. The assessed values are estimated on both sides at interval of 60%, 80%, 90% and 95%. Those results are shown in Table 4 & Table 5.

Table 3. Names of Facilities and Devices, Number of Failure Cases and MTBF

| Names of Equipment And Devices | Number of failures | MTBF   |
|--------------------------------|--------------------|--------|
| Heat pump                      | 489                | 632.40 |
| Chiller unit                   | 132                | 1431.00 |
| Return fan                     | 102                | 1920.73 |
| Power facilities               | 508                | 620.84 |
Table 4. Relation between Assessed Values and Upper Limit for Fixed Number

| Names of Equipment And devices | Assessed Values |
|-------------------------------|----------------|
|                               | 60% | 80% | 90% | 95% |
| Heat pump                     | 7.9x10^6 (minutes) | 7.1x10^6 | 6.4x10^6 | 5.2x10^6 |
| Chiller unit                  | 19.2x10^6 | 15.9x10^6 | 12.3x10^6 | 11.2x10^6 |
| Return fan                    | 27.3x10^6 | 21.8x10^6 | 17.8x10^6 | 14.3x10^6 |
| Power facilities              | 7.6x10^6 | 6.9x10^6 | 5.8x10^6 | 5.4x10^6 |

Table 5. Relation between Assessed Values and Lower Limit for Fixed Number

| Names of Equipment And devices | Assessed Values |
|-------------------------------|----------------|
|                               | 60% | 80% | 90% | 95% |
| Heat pump                     | 6.5x10^6 (minutes) | 6.2x10^6 | 5.8x10^6 | 4.5x10^6 |
| Chiller unit                  | 15.7x10^6 | 14.4x10^6 | 13.5x10^6 | 10.2x10^6 |
| Return fan                    | 20.5x10^6 | 18.6x10^6 | 14.3x10^6 | 12.6x10^6 |
| Power facilities              | 6.4x10^6 | 5.8x10^6 | 4.7x10^6 | 3.9x10^6 |

7. Summary and future issues

In this study, we examined the simulation of construction facility systems on the reliability growth model. The failure rate differs depending on systems and devices. The simulation result of our study proved that we were able to estimate regular inspection periods for commonly used construction facility systems and devices, based on the Bayesian method together with the Gompertz curve model, simply with a limited number of data. It was figured out that the lower limit of probability could deliver some margin for safety. In estimating an inspection period by using a small volume of data, the Bayesian method is considered to be appropriate for cases with less than 20 pieces of data, whereas the Gompertz curve model produces precise simulation results where the number of data is between 20 and 50. A further attempt was made to compare the Weibull cumulative hazard function and the Gompertz curve model by simulating more than 50 pieces of aggregated data on them. As a result, no big difference was found between them. That is to say, the Weibull cumulative hazard function which has been conventionally used is still applicable to the simulation of systems and devices with more than 50 pieces of data.

References

[1] Nelson. W, 1982, Applied Life Data Analysis, New York, John Wiley & Sons, Inc.
[2] T. Tanaka, 1999, Investigation on Operational Reliability of Absorption Refrigerating System, Proc 20th IIR, CD-ROM.
[3] T. Tanaka, and S. Fujii, 2006, Simulation of Operational, Reliability in Facilities System by using Bayesian Approach, proc. BUEE2006, CD-ROM.
[4] T. Tanaka, 2005, Investigation on Operational Reliability of Heat Pump System, Proc. 8th International Energy Agency, HEAT PUMP CONFERENCE 2005, CD-ROM.
[5] T. Tanaka, 1995, Reliability of Refrigerating and Air-conditioning System, Proc. 19th IIR, Vol. IIIa, pp 507-514.
[6] S. Hosoyamada, T. Tanaka, 2015, Simulation of Construction Facility Systems on the Reliability Growth Model, Proceeding of the Cold Climate HVAC 2015, China (Dalian), CD-ROM.
[7] S. Hosoyamada, T. Tanaka, 2014, Simulation of Reliability Growth for Facility System using the Random Number, Proceeding of the ICEERB 2014, China (Zhengzhou), CD-ROM.
[8] H. Kogure, http://www.kogures.com/hitoshi/webtext/stat-seicho-kyokusen/