

SYSTEM AND METHOD FOR ESTIMATING A FAILURE PROBABILITY OF ONE OR MORE DEVICES

TECHNICAL FIELD

[0001] Various aspects of this disclosure relate to systems and methods for estimating a failure probability of one or more devices.

BACKGROUND

[0002] Failure prediction and remaining useful life (RUL) estimation receive attention in many industries due to their close connections to the system reliability. Analyzing the failure numbers and RUL of equipment can provide engineers and automated controllers with direct insights on the health status of the equipment and hence allows making appropriate maintenance and replacement strategies for asset (i.e. equipment) management.

[0003] Statistical models have been widely adopted in the analysis of the system reliability, one of the most common which is Weibull due to its flexibility to model a variety of data sets. To build a Weibull model, traditional methods, e.g., least-square estimation or maximum likelihood estimation, have to rely on adequate failure data.

[0004] However, with the improved technology in the equipment manufacturing process and comprehensive maintenance plan in industry, equipment usually has a long life and high reliability, which results in limited failure data available to pursue Weibull modelling, making the failure prediction and RUL estimation not applicable. On the other hand, in the case where Weibull modelling is feasible it is typically employed to estimate the RUL for the entire fleet or population of assets but not individual devices. Therefore, a comprehensive process to build a statistical model for estimating failure probability of one or more devices even with limited failure data to allow failure number prediction and RUL estimation for individual devices (e.g. equipment such as vehicles or machines or their components), is desirable.

SUMMARY

[0005] Various embodiments concern a method for estimating a failure probability of one or more devices, including generating life-time information for each of a plurality of

observed devices of a class of devices to which the one or more devices belong, fitting a statistical model to the life-time information of the plurality of observed devices taking into account insufficient failure data in the life-time information by setting up a prior distribution for each parameter of the statistical model, formulating a likelihood function based on the life-time information, computing a joint posterior distribution of the parameters of the statistical model based on the prior distributions and the likelihood function according to Bayes's theorem and obtaining values of the parameters of the statistical model and estimating the failure probability of the one or more devices from the statistical model.

[0006] According to one embodiment, taking into account insufficient failure data in the life-time information comprises taking into account possible left truncation and right censoring of the life-time information.

[0007] According to one embodiment, the failure probability is estimated in terms of a probability density function for the probability of failure depending on device life-time.

[0008] According to one embodiment, the failure probability is estimated in terms of a cumulative distribution function for the probability of failure depending on device life-time.

[0009] According to one embodiment, the method includes using the failure probability to predict, for each of at least one of the devices, the probability that the device fails within a given time period.

[0010] According to one embodiment, the method includes using the failure probability to predict, for each of at least one of the devices, a remaining useful life of the device.

[0011] According to one embodiment, the method further comprises determining confidence bounds for the parameters of the statistical model.

[0012] According to one embodiment, the method further comprises determining confidence bounds for the parameters of the statistical model and determining confidence bounds for the remaining useful life from the confidence bounds for the parameters.

[0013] According to one embodiment, the method includes using the failure probability to predict a number of the devices which fail within a given time period.

[0014] According to one embodiment, the life-time information for each observed device includes an indication of the beginning of the life-time of the observed device.

[0015] According to one embodiment, the life-time information for each observed device indicates whether the life-time information of the observed device is left-truncated.

[0016] According to one embodiment, the life-time information for each observed device indicates whether the observed device has failed up to a final time.

[0017] According to one embodiment, the life-time information for each observed device that has failed includes a time of failure of the observed device.

[0018] According to one embodiment, obtaining values of the parameters of the statistical model comprises taking the expectation on the joint posterior distribution.

[0019] According to one embodiment, the likelihood function includes different terms for the following four types of life-time information: life-time information of observed devices that have failed without left-truncation, life-time information of observed devices that have failed with left-truncation, life-time information of observed devices that are still operating at the final time without left-truncation and life-time information of observed devices that are still operating at the final time with left-truncation.

[0020] According to one embodiment, in the likelihood function life-time information of observed devices with truncation is weighed higher than life-time information of observed devices without truncation.

[0021] According to one embodiment, in the likelihood function life-time information of each observed device with truncation is weighed the higher the higher the left-truncation time of the life-time information is.

[0022] According to one embodiment, in the likelihood function life-time information of each observed device with truncation is weighed by the inverse of the probability, given by the statistical model, that the device has not failed during the left-truncation time.

[0023] According to one embodiment, the statistical model is a Weibull distribution.

[0024] According to one embodiment, a method for controlling one or more devices is provide including estimating failure probability for the one or more devices according to the method of any one of the embodiments described above and controlling the one or more devices according to the modelled failure behaviour.

[0025] According to one embodiment, a computer program element is provided including program instructions, which, when executed by one or more processors, cause the one or more processors to perform the method of any one of the embodiments described above.

[0026] According to one embodiment, a computer-readable medium is provided including program instructions, which, when executed by one or more processors, cause the one or more processors to perform the method of any one of the embodiments described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0027] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

- FIG. 1 illustrates statistical failure modelling.
- FIG. 2 shows a diagram with the cumulative distribution function of the two-parameter Weibull distribution for three different combinations of its parameters.
- FIG. 3 shows a flow diagram illustrating a process of failure prediction and remaining useful life (RUL) estimation for individual devices according to an embodiment.
- FIG. 4 shows an example of a dataset used in the statistical modelling according to an embodiment.
- FIG. 5 shows a flow diagram illustrating a workflow of a Bayesian Weibull modelling procedure.
- FIG. 6 shows a diagram illustrating an example of a Weibull model with an illustration of the RUL computation.
- FIG. 7 shows a flow diagram illustrating a method for estimating a failure probability of one or more devices.
- FIG. 8 shows a data processing system according to an embodiment.

DETAILED DESCRIPTION

[0028] The following detailed description refers to the accompanying drawings that show, by way of illustration, specific details and embodiments in which the disclosure may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the disclosure. Other embodiments may be utilized and structural, and logical changes may be made without departing from the scope of the disclosure. The various embodiments are not necessarily mutually exclusive, as some embodiments can be combined with one or more other embodiments to form new embodiments.

[0029] Embodiments described in the context of one of the devices or methods are analogously valid for the other devices or methods. Similarly, embodiments described in the context of a device are analogously valid for a vehicle or a method, and vice-versa.

[0030] Features that are described in the context of an embodiment may correspondingly be applicable to the same or similar features in the other embodiments. Features that are described in the context of an embodiment may correspondingly be applicable to the other embodiments, even if not explicitly described in these other embodiments. Furthermore, additions and/or combinations and/or alternatives as described for a feature in the context of an embodiment may correspondingly be applicable to the same or similar feature in the other embodiments.

[0031] In the context of various embodiments, the articles “a”, “an” and “the” as used with regard to a feature or element include a reference to one or more of the features or elements.

[0032] As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

[0033] In the following, embodiments will be described in detail.

[0034] FIG. 1 illustrates statistical (failure) modelling.

[0035] Data 102 is gathered which holds information about the time of failures of devices of equipment 101 (e.g. a fleet of vehicles, a set of machines etc.). The data 102 may for example indicate a time at which devices (e.g. vehicle components) have failed. It is therefore also referred to as failure data. A statistical (failure) model 103 is generated from the failure data 102. So, the statistical model 103 may for example, model a failure probability per device depending on the life-time (e.g. time since start of operation) of the device. The statistical (failure) model 103 can for example predict the total number of failures of a fleet of devices (e.g. vehicles, machine etc.) for a given period and/or estimate the remaining useful life (RUL) for the devices.

[0036] The statistical model 103 may include a set of probability distributions P_{θ} , $\theta \in \Theta$, which depend on a set of observations $S = \{t_1, t_2, \dots\}$ (which are reflected in the data 102).

[0037] An example of a probability distribution which may be fitted to information contained in the failure data 102 as part of the statistical model 103 is a two-parameter Weibull distribution which has the probability density function and the cumulative distribution function (cdf)

$$f(t; \beta, \eta) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} \text{ and } F(t; \beta, \eta) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta}. \quad (1)$$

[0038] A controller 104 (e.g. a data processing device such as a server computer controlling the equipment, in particular in response to user commands) and/or human operator (e.g. engineer) may then use the statistical model 103 for controlling the equipment group, e.g. deactivate a device (which has a high risk of failure, e.g. above a tolerable threshold), i.e. take the device out of service, perform maintenance of the device, replace the device (e.g. a machine component), change operating parameters of the device (to reduce risk of failure or impact of failure), supplement backup devices etc.

[0039] FIG. 2 shows a diagram 200 with the cumulative distribution function of the two-parameter Weibull distribution for three different combinations of the (two) parameters β and η . The variable t for example refers to the time, e.g. a time since starting operation of a device of the equipment group 101 and $F(t)$ for example is the probability that the device has failed since the beginning of its operation until time t .

[0040] FIG. 3 shows a flow diagram 300 illustrating a process of failure prediction and remaining useful life (RUL) estimation for individual devices according to an embodiment.

[0041] The process includes processing by a data pre-processing module 301, a model development module 302 (which e.g. develops the statistical model 102), a failure prediction module 303 and a RUL estimation model 304.

[0042] In the data pre-processing module 301, key information required for the statistical modelling by the model development module 302 is prepared, e.g. from the failure data 102. The key information includes life-time information for each of a plurality of devices (in particular the information whether a device has failed until a current time, also referred to as final time in the failure data 102, and the time the device has been operating before failure).

[0043] FIG. 4 shows an example of a dataset 400 (i.e. the prepared key information) used in the subsequent statistical modelling. Each row corresponds to one observed device (of a class of devices of interest) and holds the life-time information for the respective device. The information in all columns can for example be obtained from a database maintained by an organization or company through some simple calculations (e.g. applied to the stored failure data 102)).

[0044] In the example of FIG. 4, each row corresponds to a respective device and the 'status' flag indicates whether the data of the respective line is failure data (i.e. data for a device which has failed, value 1) or an operating data (i.e. data for a device which is still operational, value 0). The column 'time' refers to the life-time of the respective device determined by subtracting the time of the beginning of the life-time (e.g. manufacturing time or time of putting the device in operation, second column) from the failure time (third column) for an observed device that has failed or from the current time for an observed device that is still operating. The 'trunc' flag indicates whether the data for the respective device is left-truncated (value 0) or not (value 1).

[0045] If the manufacturing time is before the time of starting the data record (e.g. start of recording the failure data, year 2000 in this example), the data is left-truncated; otherwise, it is not left-truncated. The 'time_lt' represents the left-truncation time (by subtracting the respective manufacturing time from the time of starting the data record) if the data is left-truncated, and zero otherwise.

[0046] With a pre-prepared dataset as shown in FIG. 4, the next step, performed by the model development module 302, is to build a statistical model. It should be noted that the failure data may be limited or there may be no failures at all contained in the failure data in the extreme case. Hence, standard Weibull modelling may not be applicable. Therefore, as remedy for such a case, Bayesian Weibull modelling is used in the model development module 302, i.e. a Weibull model whose probability density function (pdf) and cumulative distribution function (cdf) are as given by equation (1) is generated using a Bayesian approach (i.e. Bayesian inference), where the shape parameter β and the scale parameter η are regarded as random variables.

[0047] FIG. 5 shows a flow diagram 500 illustrating a workflow of a Bayesian Weibull modelling procedure (performed by the model development module 302). To facilitate the Bayesian Weibull modelling, prior information about the Weibull parameters is needed, which is specified in terms of some prior distribution 501 such as a normal, lognormal, beta or uniform distribution.

[0048] For example, the model development module 302 provides two commonly used prior distributions with one being uninformative and the other being informative: the uniform distribution and the lognormal distribution, whose pdfs are given by

$$\text{uniform: } f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{otherwise} \end{cases},$$

$$\text{lognormal: } f(x) = \begin{cases} \frac{2}{\text{erf}(\frac{\ln b - \mu}{\sigma\sqrt{2}}) - \text{erf}(\frac{\ln a - \mu}{\sigma\sqrt{2}})} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} & x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

where a, b are lower and upper bounds of x for both distributions, and μ, σ are the mean and standard deviation in the lognormal distribution. Let $\pi(\beta)$ and $\pi(\eta)$ denote the probability density functions of the prior distributions for the shape parameter and scale parameter, respectively.

[0049] If there are past information about the parameters available, such as historical model parameters $(\{\beta_1, \beta_2, \dots\}, \{\eta_1, \eta_2, \dots\})$ of the same type of equipment (e.g. device group or class), then the informative normal distribution can be used as a prior 501, whose hyperparameters μ, σ can be obtained by fitting it to the historical model parameters. If there is no historical information available, then the uninformative uniform distribution can be used as a prior 501.

[0050] As β and η are always positive, the lower bound a (respectively, upper bound b) can be set reasonably small (respectively, large) to avoid bias in general. Specifically, for the shape parameter β , the bounds a, b can be set based on the failure stage (e.g. degradation phase) of the equipment if it can be identified, i.e., for example $0 < \beta < 1$ for the infant stage, and $\beta > 1$ for the wear-out stage. The trade-off is that the obtained confidence bounds could be conservative.

[0051] To obtain a tighter confidence bounds, it is preferable to set tighter bounds a, b for the prior distributions of the parameters for both β and η . To do so, it is possible to look into various analysis of similar fleets or equipment type whose parameters are estimated by relatively sufficient dataset. For example, for an electrical equipment with real data, the shape parameter for different brands ranges from 2.48 to 2.66. Then, it is possible to obtain a reasonable and practical lower bound a around 2.2 and upper bound b around 2.8 for the

shape parameter such that the middle 80% of the prior distribution for the shape parameter lies within 2.2 to 2.8.

[0052] The likelihood function 502 of the dataset 503, denoted by $L(\beta, \eta|\text{data})$, is obtained by considering four types of data. Let $\{t_i\}_{i=1}^{n_f}$ denote the life-time (i.e. life-time until failure) for the failure data without left-truncation, $\{t_j\}_{j=1}^{n_s}$ denote the life-time (i.e. life time up to the current time) for the operating data without left-truncation, $\{t_n, \tau_n\}_{n=1}^{n_{f,t}}$ denote the life-time and left-truncation time for the failure data with left-truncation, and $\{t_m, \tau_m\}_{m=1}^{n_{s,t}}$ denote the life-time and left-truncation time for operating data with left-truncation. Then, the likelihood function of the data is given by

$$L(\beta, \eta|\text{data}) = \prod_{i=1}^{n_f} f(t_i; \beta, \eta) \times \prod_{j=1}^{n_s} (1 - F(t_j; \beta, \eta)) \\ \times \prod_{n=1}^{n_{f,t}} \frac{f(t_n; \beta, \eta)}{1 - F(\tau_n; \beta, \eta)} \times \prod_{m=1}^{n_{s,t}} \frac{1 - F(t_m; \beta, \eta)}{1 - F(\tau_m; \beta, \eta)},$$

where $f(t; \beta, \eta)$ and $F(t; \beta, \eta)$ are the pdf and the cdf of the Weibull distribution according to (1), respectively.

[0053] With the specified prior distributions on the shape and scale parameter shape parameter β and the scale parameter η and the likelihood function 502 of the dataset 503, the joint posterior distribution 504 of β and η can be obtained by

$$\pi(\beta, \eta|\text{data}) = \frac{\pi(\beta)\pi(\eta)L(\beta, \eta|\text{data})}{\int_{\eta} \int_{\beta} \pi(\beta)\pi(\eta)L(\beta, \eta|\text{data})d\beta d\eta}. \quad (2)$$

[0054] The estimates 505 of the shape and scale parameter $\hat{\beta}, \hat{\eta}$ can be obtained by taking the expectation with respect to each of them, given by

$$\hat{\beta} = \mathbf{E}[\beta|\text{data}] = \int_{\eta} \int_{\beta} \beta \cdot \pi(\beta, \eta|\text{data})d\beta d\eta, \\ \hat{\eta} = \mathbf{E}[\eta|\text{data}] = \int_{\eta} \int_{\beta} \eta \cdot \pi(\beta, \eta|\text{data})d\beta d\eta, \quad (3)$$

where $\pi(\beta, \eta|\text{data})$ is obtained from (2). Now, the Bayesian Weibull modelling is completed.

[0055] For the two-sided confidence bounds of scale and shape parameters, denote the upper and lower bounds of scale parameter by η_u and η_l , the upper and lower bounds of shape parameter by β_u and β_l corresponding to confidence level, CL , then η_u and η_l can be obtained by solving:

$$\frac{1 - CL}{2} = \int_{\beta} \int_0^{\eta_l} \pi(\beta, \eta|\text{data}) d\eta d\beta, \frac{1 + CL}{2} = \int_{\beta} \int_0^{\eta_u} \pi(\beta, \eta|\text{data}) d\eta d\beta.$$

β_u and β_l can be obtained by solving:

$$\frac{1 - CL}{2} = \int_0^{\beta_l} \int_{\eta} \pi(\beta, \eta|\text{data}) d\eta d\beta, \frac{1 + CL}{2} = \int_0^{\beta_u} \int_{\eta} \pi(\beta, \eta|\text{data}) d\eta d\beta.$$

[0056] For the two-sided confidence bounds of Bayesian Weibull cdf, (i.e., probability of failure), denote the upper and lower bounds of probability of failure at time t by F_u and F_l corresponding to confidence level, CL , then they can be obtained by solving:

$$\begin{aligned} \frac{1 - CL}{2} &= \int_{\beta} \int_0^{te^{-\frac{\ln(-\ln(1-F_u))}{\beta}}} \pi(\beta, \eta|\text{data}) d\eta d\beta, \\ \frac{1 + CL}{2} &= \int_{\beta} \int_0^{te^{-\frac{\ln(-\ln(1-F_l))}{\beta}}} \pi(\beta, \eta|\text{data}) d\eta d\beta. \end{aligned}$$

[0057] The failure prediction module 303 uses the developed Bayesian Weibull to predict the total number of failures in the fleet of assets for a given period. For example, the failure numbers over a time duration $T \geq 0$ should be predicted. Then, the conditional probability that the equipment of age t_i will fail during this period, denoted by $P(t_i, T)$, can be obtained by

$$P(t_i, T) = \frac{F(t_i + T; \hat{\beta}, \hat{\eta}) - F(t_i; \hat{\beta}, \hat{\eta})}{1 - F(t_i; \hat{\beta}, \hat{\eta})},$$

where $\hat{\beta}, \hat{\eta}$ are the estimated Bayesian Weibull parameters obtained from (3). Let there be a number n_i of (still operating) devices at age t_i , then the total number of failures among

all these devices at age t_i over a time duration T , denoted by $n_F(t_i, T)$, can be obtained by

$$n_F(t_i, T) = P(t_i, T) \times n_i.$$

[0058] Finally, failure prediction module 301 can obtain the total number of failures among all the equipment of all ages over a time duration T , denoted by $N_F(T)$, can be by

$$N_F(T) = \sum_{t_i} n_F(t_i, T).$$

[0059] The RUL estimation module 304 uses the built Weibull model to estimate the RUL for each individual component or equipment. To facilitate the estimation, a certain level of unreliability is specified, called unreliability threshold (denoted by F_{critical}), which can be interpreted as the highest probability of failure that can be tolerated for the equipment. For example, $F_{\text{critical}} = 1\%$ implies that at most 1% of the total fleet failure can be tolerated. Hence, for different types of equipment, F_{critical} could be set to different values. From the cdf of the Weibull model, the time which corresponds to the probability of failure F_{critical} is denoted as the reliability age gap, denoted by t_{critical} with upper and lower bounds $t_{\text{critical,upper}}$ and $t_{\text{critical,lower}}$. With the reliability age gap, RUL estimation module 304 obtains the RUL for the equipment of age t_i , denoted by $\text{RUL}(t_i)$, by

$$\text{RUL}(t_i) = t_{\text{critical}} - t_i. \quad (4)$$

[0060] The confidence bounds for the remaining useful life, RUL for the equipment of age t_i , denoted by $\text{RUL}_{\text{upper}}(t_i)$ and $\text{RUL}_{\text{lower}}(t_i)$, by

$$\text{RUL}_{\text{upper}}(t_i) = t_{\text{critical,upper}} - t_i, \text{RUL}_{\text{lower}}(t_i) = t_{\text{critical,lower}} - t_i$$

[0061] The controller 104, for example, may use the results of the failure prediction module 301, the RUL estimation module 304 or both to control the equipment (in particular for example one or more devices for which RUL was determined).

[0062] FIG. 6 shows a diagram 600 illustrating an example of the cdf of the Weibull model (determined by the Bayesian approach as described above) with an illustration of the RUL computation as well as the cdf according to the confidence bounds and the resulting RUL bounds. For different equipment with different ages, the same computation (4) applies,

and hence the RUL estimation module 304 can obtain the RUL for each individual device is obtained. Moreover, it can be seen from (4) that a higher unreliability threshold F_{critical} results in a larger t_{critical} , and a younger equipment with smaller t_i leads to a larger RUL as expected.

[0063] In summary, according to various embodiments, a method is provided as illustrated in FIG. 7.

[0064] FIG. 7 shows a flow diagram illustrating a method for estimating a failure probability of one or more devices.

[0065] In 701, life-time information is generated for each of a plurality of observed devices of a class of devices to which the one or more devices belong.

[0066] In 702, a statistical model (e.g. a 2- or 3-parameter Weibull distribution but other statistical models may also be used) is fit to the life-time information of the plurality of observed devices taking into account insufficient failure data in the life-time information. This is done by setting up a prior distribution for each parameter of the statistical model, formulating a likelihood function based on the life-time information, computing a joint posterior distribution of the parameters of the statistical model based on the prior distributions and the likelihood function according to Bayes's theorem and obtaining values of the parameters of the statistical model.

[0067] In 703, the failure probability of the one or more devices is estimated from the statistical model.

[0068] According to various embodiments, in other words, a process is provided that allows statistical failure modelling in the case of limited failure data, in particular statistical failure modelling for a dataset (e.g. extracted from failure data 102) with left-truncation and/or right-censoring. According to various embodiments, the process establishes a statistical (failure) model that can predict the total number of failures of a fleet of equipment for a given period and/or can estimate the remaining useful life (RUL) of each individual device (of the equipment, e.g. a machine or vehicle component). The process thus can help to improve and optimize the maintenance and replacement strategies for asset management with accurate failure prediction and RUL estimation of each individual device.

[0069] So, according to various embodiments, a process for predicting the total number of failures for a given period and estimating the remaining useful life (RUL) for each

individual device (piece of equipment or component) with limited failure data is provided that, according to one embodiment, builds a statistical model in the case of limited failure data, predicts the total number of failures, and estimates the RUL for each individual component or equipment is provided. The process is for example performed by a data pre-processing module, a model development module, a failure prediction module and a RUL estimation module (performing corresponding operations (or steps) of the process). Specifically, in the data pre-processing module, data is processed and key information is retrieved to be prepared for the model development (i.e. building). The model development module is responsible for building a statistical model based on the data prepared from the data pre-processing module and some prior knowledge of the model parameters. With the developed model, the failure prediction module computes the total number of failures in the fleet of assets for a given period, and the RUL estimation module will calculate the RUL for each individual device.

[0070] According to various embodiments, methods adopted in the process to build a statistical model in the case of limited failure data, predict the total number of failures, and estimate the RUL for each individual component or equipment are provided. These methods for example include using a Bayesian Weibull model for statistical modelling, wherein the model parameters are estimated based on the composition of the prior distributions of the model parameters and the likelihood function constructed by the existing data. The failure number prediction is computed based on the conditional probability of failure obtained from the developed Bayesian Weibull distribution functions. The RUL estimation of each individual component or equipment is derived based on the cumulative distribution function of the developed Bayesian Weibull model and the age of each component or equipment.

[0071] So, according to various embodiments, a process for predicting the total number of failures for a given period and estimating the remaining useful life (RUL) for each individual component or equipment with limited failure data is provided, specifically including (performing the functions of):

- a data pre-processing module to prepare the key information required for the statistical modelling from the database maintained by the organization/company;
- a model development module to build a statistical model that can handle limited failure data in the dataset with left-truncation and right-censoring;

- a failure prediction module to predict the total number of failures in the fleet of assets for a given period; and
- a remaining useful life (RUL) estimation module to estimate the RUL of each individual component or equipment.

[0072] The data pre-processing may further include a collection of the key information including manufacturing time, time of starting the data record (if any), current time and failure time (if any) and procedures to compute the required information including time, left-truncation time, flags of status (failure data or operating data), flags of truncation (with or without truncation).

[0073] The model development module may further include

- a module for the prior distribution which can take uniform or lognormal distribution based on the prior knowledge for the Bayesian Weibull modelling;
- a module to construct likelihood function based on the data prepared in the data pre-processing module with the consideration of four types of data:
 - failure data without left-truncation,
 - operating data without left-truncation,
 - failure data with left-truncation and
 - operating data with left-truncation;
- a module to compute joint posterior distribution from the specified prior distributions on the parameters and the obtained likelihood function of the dataset based on Bayes' theorem; and
- a module to estimate the Bayesian Weibull parameters by taking the expectation with respect to each of the parameters.

[0074] The failure prediction module may further include (performing)

- a method to compute the conditional probability of failure for the equipment of a particular age over a period;
- a method to compute the total number of failures of all equipment of a particular age over a period; and
- a method to compute the total number of failures of all equipment of all ages (total fleet of assets) over a period.

[0075] The RUL estimation module may further include a concept of the unreliability threshold specified by required reliability level to facilitate the RUL estimation; and a method

to estimate the RUL for each component or equipment based on the cdf of the obtained Bayesian Weibull model and the specified unreliability threshold.

[0076] The method of FIG. 7 is for example carried out by a data processing system (e.g. a computer or multiple computers) as illustrated in FIG. 8.

[0077] FIG. 8 shows a data processing system 800 according to an embodiment.

[0078] The data processing system 800 includes a communication interface 801 (e.g. configured to receive failure data, e.g. via user input and/or from sensors). The server computer 800 further includes a processing unit 802 and a memory 803. The memory 803 may be used by the processing unit 802 to store, for example, data to be processed, such as the failure data and the derived life-time information. The data processing system is configured to perform the method of FIG. 7.

[0079] The methods described herein may be performed and the various processing or computation units and the devices and computing entities described herein (in particular the various modules described above) may be implemented by one or more circuits. In an embodiment, a "circuit" may be understood as any kind of a logic implementing entity, which may be hardware, software, firmware, or any combination thereof. Thus, in an embodiment, a "circuit" may be a hard-wired logic circuit or a programmable logic circuit such as a programmable processor, e.g. a microprocessor. A "circuit" may also be software being implemented or executed by a processor, e.g. any kind of computer program, e.g. a computer program using a virtual machine code. Any other kind of implementation of the respective functions which are described herein may also be understood as a "circuit" in accordance with an alternative embodiment.

[0080] While the disclosure has been particularly shown and described with reference to specific embodiments, it should be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the invention as defined by the appended claims. The scope of the invention is thus indicated by the appended claims and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced.

CLAIMS

1. A method for estimating a failure probability of one or more devices, comprising:
generating life-time information for each of a plurality of observed devices of a class of devices to which the one or more devices belong;
fitting a statistical model to the life-time information of the plurality of observed devices taking into account insufficient failure data in the life-time information by
 setting up a prior distribution for each parameter of the statistical model;
 formulating a likelihood function based on the life-time information;
 computing a joint posterior distribution of the parameters of the statistical model based on the prior distributions and the likelihood function according to Bayes's theorem;
 obtaining values of the parameters of the statistical model;
estimating the failure probability of the one or more devices from the statistical model;
using the failure probability to predict, for each of at least one of the one or more devices, a remaining useful life of the device;
determining confidence bounds for the parameters of the statistical model; and
determining confidence bounds for the remaining useful life from the confidence bounds for the parameters.
2. The method of claim 1, wherein taking into account insufficient failure data in the life-time information comprises taking into account possible left truncation and right censoring of the life-time information.
3. The method of any claim 1 or 2, wherein the failure probability is estimated in terms of a probability density function for the probability of failure depending on device life-time.
4. The method of any one of claims 1 to 3, wherein the failure probability is estimated in terms of a cumulative distribution function for the probability of failure depending on device life-time.

5. The method of any one of claims 1 to 4, comprising using the failure probability to predict, for each of at least one of the devices, the probability that the device fails within a given time period.
6. The method of any one of claims 1 to 5, comprising using the failure probability to predict a number of the devices which fail within a given time period.
7. The method of any one of claims 1 to 6, wherein the life-time information for each observed device includes an indication of the beginning of the life-time of the observed device.
8. The method of any one of claims 1 to 7, wherein the life-time information for each observed device indicates whether the life-time information of the observed device is left-truncated.
9. The method of any one of claims 1 to 8, wherein the life-time information for each observed device indicates whether the observed device has failed up to a final time.
10. The method of claim 9, wherein the plurality of observed devices include at least one observed device that has failed and at least one observed device that is still operating at the final time.
11. The method of claim 9 or 10, wherein the life-time information for each observed device that has failed includes a time of failure of the observed device.
12. The method of any one of claims 1 to 11, wherein obtaining values of the parameters of the statistical model comprises taking the expectation on the joint posterior distribution.
13. The method of claim 12, wherein the likelihood function comprises different terms for the following four types of life-time information: life-time information of observed

devices that have failed without left-truncation, life-time information of observed devices that have failed with left-truncation, life-time information of observed devices that are still operating at the final time without left-truncation and life-time information of observed devices that are still operating at the final time with left-truncation.

14. The method of claim any one of claims 1 to 13, wherein in the likelihood function life-time information of observed devices with truncation is weighed higher than life-time information of observed devices without truncation.
15. The method of any one of claims 1 to 14, wherein in the likelihood function life-time information of each observed device with truncation is weighed the higher the higher the left-truncation time of the life-time information is.
16. The method of any one of claims 1 to 15, wherein in the likelihood function life-time information of each observed device with truncation is weighed by the inverse of the probability, given by the statistical model, that the device has not failed during the left-truncation time.
17. The method of any one of claims 1 to 16, wherein the statistical model is a Weibull distribution.
18. Method for controlling one or more devices comprising estimating failure probability for the one or more devices according to any one of claims 1 to 17 and controlling the one or more devices according to the modelled failure behaviour.
19. A data processing system comprising a communication interface, a memory and a processing unit configured to perform the method of any one of claims 1 to 18.
20. A computer program element comprising program instructions, which, when executed by one or more processors, cause the one or more processors to perform the method of claims 1 to 18.

21. A computer-readable medium comprising program instructions, which, when executed by one or more processors, cause the one or more processors to perform the method of any one of claims 1 to 18.