Degradation model constructed with the aid of dynamic Bayesian networks
Anselm Lorenzoni1, Michael Kempf1 and Oliver Mannuß1

Abstract: This paper develops a generic degradation model based on Dynamic Bayesian Networks (DBN) which predicts the condition of a technical system. Besides handling bi-directional reasoning, a major benefit of this degradation model using a DBN is its ability to adequately model stochastic processes as well as Markov chains. We will assume that the behavior of the degradation can be represented as a P–F-curve (also called degradation or life curve). The model developed here is able to combine information from expert knowledge, any kind of sensor and operating data as well as information from the machine operator. Using the Bayesian approach, uncertain knowledge can be handled appropriately. Thus it is even possible to take into account the environment and stress under which the component or system is operating. Hence, it is possible to detect potential failures at an early stage and initiate appropriate remedy and repair strategies prior to catastrophic failure.

Subjects: Quality Control & Reliability; Manufacturing & Processing; Engineering Mathematics
Keywords: degradation; dynamic Bayesian networks; P–F-curve; stochastic process

1. Introduction
Machine failures often cause unexpected downtimes in manufacturing which lead to critical lapses in production and thereby cause high financial loss (Hirschmann, 2007). This financial loss may even cause small and medium enterprises to declare bankruptcy. Therefore, it is crucial to reduce if not eliminate machine failures in manufacturing using prediction tools which can estimate the durability of a component, machine, or system.
Traditional life tests are employed to approximate the durability of a machine or a component. Such life tests can, according to Lu and Meeker (1993), measure the time to failure. However, it is difficult to make a definitive statement about the durability of a specific, individual component, and accurately predict failures before they occur: failures may be caused by operating errors, construction faults, material faults, and many other reasons. Failure times do not depict the reasons for premature failures. Moreover, it is difficult to record failure times if the components or machines are new or highly reliable, as it may take a long time until one of these highly reliable components fails and as it can be expensive to generate such data (Robinson & Crowder, 2000). Therefore, it is crucial to develop new ways of predicting failures before they occur.

To develop such a model, we may consider that degradation eventually leads to a weakness that can cause failure. Therefore, if it were possible to measure degradation, this could provide more information about the reliability of production systems and the causes of their failure than time-to-failure data. Consequently, a relationship between live observations and degradation has to be found, as well as a relationship between degradation of components or systems and failure. Once both relationships are established, it is possible to estimate how badly the component or system is degraded, as well as to predict failure and the time-to-failure based on live sensor data.

This paper will develop a model of such a degradation process. This model is based on a so-called P–F-curve which is implemented as a Dynamic Bayesian Network (DBN). In so doing, it takes into account that degradation processes can be influenced bi-directionally by operating conditions, maintenance activities, and observations occurring when potential failures arise. The presented model will enhance the reliability of manufacturing equipment by providing information on the condition of the different components and systems during the lifetime of a product line. It will also indicate whether it is possible to reuse these components and systems in new product lines. Thereby, it will reduce fatal errors and significantly lower production costs.

This paper builds on an earlier paper by the author on the subject (Lorenzoni & Kempf, 2015). It significantly augments this contribution by explicating in detail the mathematical expressions as well the algorithms employed by the presented model and by including a use-case to validate the model.

2. Background

2.1. The P–F-curve

It can be assumed that the condition of a component or system is directly related to its degradation: if a component or system is half degraded, its condition is expected to be half as good as new. Hence, this paper uses the term “condition” of a component more frequently than the term “degradation”.

Most failure modes provide some sort of warning of incipient failure. In other words, there is often evidence that something is in the final stages prior to failure (Moubray, 1997). This is very helpful for creating a degradation model because it is possible to estimate the state of a component or system based on several observations such as vibration, sounds, or power consumption. Observations of failures can often be made in defined stages of the condition of a component or system. Clearly, there is a relationship between the chronological age of a component or system and its failures, but this relationship is not necessarily linear, for example in the case of a construction fault or operating error causing failure to occur earlier (or later) than expected.

Initially, the condition of a component will remain good for a certain period of time (performance is without any changes). However, at a certain point, where an observation indicating a potential failure emerges as likely, the condition of the component as well as its performance decrease precipitously. Typical degradation behavior follows the characteristics of a P–F-curve as developed by Moubray (1997), also called the “life curve” (Sugier & Anders, 2010) or degradation curve.
(Nelson, 1998). Figure 1 depicts such a P–F-curve. It “shows how a failure starts, deteriorates to the point at which it can be detected (e.g. vibrations could be detected) until it reaches the point of functional failure” (Moubray, 1997).

It is very useful to integrate this curve into degradation models. At the beginning of the operating time, the degradation of a component or system is minor, but there comes a certain point where the degradation of the component or system starts and “degradation accelerates in the final stages of most failures” (Moubray, 1997).

2.2. Markov property
Andrei Andreevich Markov (1856–1922) was a Russian mathematician. He is well known for his work on number theory, analysis and probability theory. A primary subject of his work is known as Markov chains and processes (Basharin, Langville, & Naumov, 2004).

A Markov process is a special stochastic process. Markov developed this concept in 1907 for discrete time processes with finite state spaces (Yin & Zhang, 2013). In a classical discrete Markov chain, each random variable \((x_1, x_2, \ldots)\) is dedicated to time. The state of the variables at any given time is defined. That means that variable \(x\) at time \(m\) has the state \(r\) (in formula \(x_m(t = m) = r\)). The state of the node at time point \(m\) is dependent on the state of the variable at time point \((m-1)\) but is absolutely independent of any former time point \((m - y, y = 2, 3, 4, \ldots)\) (Upton, Graham, & Cook, 2008). As Jensen and Nielsen succinctly put the so-called “Markov property”: “The past has no impact on the future given the present” (Jensen & Nielsen, 2007).

2.3. Bayesian networks
In the following, the basic tenets of Bayesian Networks (BN) will be explained. According to Heckermann, BN are graphical models which are able to visualize probabilistic relationships among a set of variables (Heckermann, 1995). BN are directed acyclic graphs (DAG): nodes represent the variables in the model. Each variable (and thus each node) can have different conditions (called states) in the net. These states can be numbered values but may also be defined as interval values or binary values, and they can be distributed with freely selectable texts. The nodes in the BN can be connected by omnidirectional arcs; it is not possible to connect the nodes with arcs in a circular way (hence, the term “directed acyclic graph”). To distribute the probabilities behind the nodes and the arcs, so-called conditional probability tables (CPD) are used (Weidl, Madsen, & Israelson, 2005).

Figure 2 shows a very simple example of a BN. This BN was generated using a software called HUGIN EXPERT. The graphical user interface (GUI) only shows nodes and arcs. Every arc between two nodes indicates a direct probabilistic dependence between them. By implication, the lack of an arc between two nodes indicates a conditional independence (McNaught & Chan, 2011).
Jensen and Nielsen (2007) summarize Bayesian Networks as follows:

- "A set of variables and a set of directed edges [arcs] between variables"
- "Each variable has a finite set of mutually exclusive states"
- "The variable together with the directed edges form a DAG [...]"
- "To each variable $B$ with parents $A_1, \ldots, A_n$, a conditional probability table (CPT) $P(B|A_1, \ldots, A_n)$ is attached"

### 2.4. Dynamic Bayesian Networks

Bayesian Networks are useful when the state is static and time is irrelevant. However, to extend a Bayesian Network into a time dimension, a DBN can be used (Hulst, 2006). The structure of the network does not change dynamically but one can model a dynamic system with it. "A DBN is a directed acyclic graphical model of a stochastic process. It consists of time-slices (or time-steps), with each time-slice containing its own variables" (Hulst, 2006). Figure 3 shows an example of a DBN.

The dependency of the individual time-slices of a DBN obeys the so-called Markov property (see above): "the future is conditionally independent of the past given the present" (Kjærulff, 1995).
In the example illustrated by Figure 3, “Node A” and “Node C” are time-variant variables which obey the Markov property in each time-slice. Jensen and Nielsen summarize time-variant variables as follows: “A dynamic Bayesian network is first-order Markovian when the variables at time step \( t + 1 \) are d-separated from the variables at time step \( t - 1 \) given the variables at time step \( t \)” (Jensen & Nielsen, 2007).

Some DBN-nodes are however also time-invariant or stationary. In the example given, this is the case for the nodes B and Node D: there is no arc between the current and the next time-slices and they are conditionally independent of their values in earlier time-slices.

3. The degradation model

3.1. Reversed exponential function as a Markovian Polygon

The basic concept of the degradation model in this paper is based on the work of Straub who describes the process of deterioration depending on time, a set of time-variant and time-invariant parameters (root causes) as well as observations (Straub, 2009). The set of time-variant parameters is represented by the P–F-curve, the time-invariant parameters are given as influencing factors of degradation and, finally, the observations made are given as available sensor data as well as from human inspections. In the following, the integration of the parameters or observations in a DBN will be explained.

The first task will be to develop a curve which has the same curve progression as a P–F-curve. The reversed exponential curve is very similar to the P–F-curve, which can be expressed as follows:

\[
eC(t) = \begin{cases} 
101 - e^{\ln(101)} \frac{t}{D}, & t < D \\
0, & t \geq D 
\end{cases}
\]

The problem with this function is that time is used, but it is not possible to use a continuous time variable in DBN. As Moubray (1997) notes: “Many failure modes are not age-related, most of them give some sort of warning that they are in process of occurring or are about to occur”. For example, if there is an observation or parameter which indicates that a failure will probably occur, the degradation process has started—and if the degradation process has already started, the condition of a component or system often deteriorates rapidly. Thus, the condition of a component or system clearly depends on its condition one time step before. Therefore, a stochastic process obeying the Markov property has to be sought which exhibits the same trend as a reversed exponential curve. A way to accomplish this is to model a polygon function from the reversed exponential function.

A polygon function consists of polygonal lines. Owing to memory capacity and computational speed, five polygonal lines are chosen here. In other words, for every 20% change in the condition of
a component or system, the polygonal line changes. It is advantageous to implement linear polygonal lines in DBN. Figure 4 illustrates the “Markovian polygon”:

The formula of the expected condition \( eC \) curve (P–F-curve) is:

\[
eC(s) = \begin{cases} 
\frac{eC(s - 1) - 20 \cdot \frac{c}{G(s)}}{G(s)}, & eC(s) > 0 \\
0, & eC(s) \geq 0
\end{cases}
\]

(2)

With Gradient:

\[
G(s) = \begin{cases} 
0.65968, & eC(s - 1) > 80 \\
0.14497, & eC(s - 1) > 60 \\
0.0861, & eC(s - 1) > 40 \\
0.06144, & eC(s - 1) > 20 \\
0.04781, & eC(s - 1) > 0
\end{cases}
\]

(3)

And for the time-slice length \( \frac{t_{\text{slice}}}{D} \) and Durability \( D \):

\[
\frac{t_{\text{slice}}}{D} = c
\]

(4)

3.2. Implementation of a dynamic environment and maintenance activities

Influences by operational conditions cannot be excluded—operational issues and the environment might well be expected to influence wear over time. If the component or system is not being used, the "Degradation level" \( Dl \) becomes 0 (no stress). If the component or system operates under normal circumstances, \( Dl = 1 \) (normal stress). If the environment is suboptimal for the component or system (for example, very dirty, hot, or high humidity) or the operating conditions exceed those for which the component or system was designed, \( Dl \) could be 2 (high stress) or in very unusual circumstances, \( Dl = 3 \) (very high stress). To implement the Degradation level into the formula (Figure 5):

\[
eC(s) = \begin{cases} 
\frac{eC(s - 1) - 20 \cdot \frac{c}{G(s)} \cdot Dl(s)}{G(s)}, & eC(s) > 0 \\
0, & eC(s) \geq 0
\end{cases}
\]

(5)

Maintenance will also influence the degradation process. According to the European Standard, maintenance is defined as “combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function.” There are three different kinds of maintenance which have to be distinguished: (1) no maintenance action (2) maintenance like re-lubrication or re-tensioning; (3) replacement of components of the system. To implement the maintenance activity “replacement” in the model, the formula has to be changed to:

\[
eC(s) = \begin{cases} 
100, & M = “\text{replace}” \\
\frac{eC(s - 1) - 20 \cdot \frac{c}{G(s)} \cdot Dl(s)}{G(s)}, & eC(s) > 0 \\
0, & eC(s) \leq 0
\end{cases}
\]

(6)

Figure 5. P–F-curve modelled in HUGIN 8.1.
After Maintenance actions like re-lubrication or re-tensioning, the degradation process will be slower than before maintenance. While the condition of the component is not improved, the degradation process is delayed. To implement a degradation delay in this network, a new node called “Degradation Delay” \( DD \) is introduced \( 0.1 \leq DD(s) \leq 1 \). The variable \( x \) is defined according to expert knowledge and specifies the extent to which the degradation process is delayed.

What happens to the “Degradation Delay” when the maintenance activity “replace” is carried out? When a new component is installed, this new component has not enjoyed the same maintenance activity as the previous component. Therefore, there has not been any maintenance activity “reset” regarding this component. To express the “Degradation Delay” \( DD \) mathematically Figure 6:

\[
DD(s) = \begin{cases} 
1, & M = "replace" \\
DD(s - 1), & M = "none" \\
DD(s - 1) - 0.3, & M = "reset" \\
0.1, & DD(s) \leq 0.1
\end{cases}
\]  

(7)

\[
eC(s) = \begin{cases} 
100, & M = "replace" \\
eC(s - 1) - 20 \cdot \frac{\gamma}{\sigma(s)} \cdot Dl(s) \cdot DD(s), & eC(s) > 0 \\
0, & eC(s) \leq 0
\end{cases}
\]  

(8)

3.3. Implementation of the condition states and the observations

After modeling the P–F-curve as an expected Condition curve in HUGIN 8.1, two problems arise: (1) the number of states in the node Expected Condition curve is very high (50) and thus very difficult to interpret; (2) a degree of uncertainty exists as both the curve and the durability are estimated and as sensor data and observations may be inaccurate. To resolve these problems, a categorization of states and a normal distribution will be integrated into the model via an extra node.

To simplify the conceptualization of the continuous degradation of a component or system, a so-called damage-index is introduced. This paper refers to the examples of IEEE Trondheim PowerTech (2011) and McNaught and Zagorecki (2009). Both use five states to define the condition of a component or system. These states are as follows: with a new replacement component, there will be no indication for degradation (state 1). The component is as good as new. After some time, the condition of the component will decline to the next state, state 2. Now there will be some indications of degradation: for example, vibration or particles in the oil could be observed. The next state, state 3, is accompanied by serious degradation, as indicated in this example by noise or stronger vibrations. Progression to state 4 is then very rapid and the condition of the component becomes critical. High temperatures and smoke might now, for example, be observed. If the condition of the component reaches state 5, the probability of a failure is very high.
To implement this in the mathematical model, we need a new variable which gives notice of the Condition \([C]\) of the component or system. The five states of this variable are \([0-20, ..., 80-100]\). However, HUGIN 8.1 requires a normal distribution to be used for the border values \(-\infty\) and \(\infty\). Therefore, the states are \([-\infty - 20, ..., 80 - \infty]\). The formula for the Condition is now:

\[
P(C(s)) = \text{Normal}(\epsilon C(s), \nu)
\]  \hspace{1cm} (9)

The variance \([\nu]\) depends on the knowledge of the experts assessing the system Figure 7.

4. Case study

A simple case study is used here to test the degradation model developed. This case study of the degradation of bearings is based on a case described by Hirschmann (2007) in his doctoral thesis. The generic degradation model is modified accordingly, and the results of the case study are entered into the model in order to show how close the model might reflect reality. Hirschmann (2007) describes an experiment testing bearings. This experiment is a kind of accelerated life test. The results are illustrated by Figure 8:
After three hours and ten minutes, the middle bearing is worn out and failure occurs. 25 min beforehand, the temperature rises very quickly. The motor current increases after approximately three hours. In our model, the durability is set to \( D = 20 \) h and the time-slice length \( t_{\text{slice}} = 5 \) min. As the results show, two “Observations” can be made: Temperature and Motor current. The Motor current is divided into two states \([\text{true}; \text{false}]\) (true = Motor current high; false = Motor current normal) The Temperature is divided into three States \([0–30; 30–40; 40–\infty]\). The CPTs appear as follows:

For the simulation, it has to be determined at which points of time the temperature and the current change so that new states are entered. These points of time can be estimated based on Figure 8. The following results for the points of time are then suggested:

The Degradation model for this experiment then looks like this (Figure 9):

![Degradation model of the bearing.](image)

### Table 1. CPTs of the nodes “Current high” and “Temperature”

| State | Current high = false | Current high = true |
|-------|---------------------|---------------------|
| State 5 [−\infty–20] | 5 | 95 |
| State 4 [20–40] | 50 | 50 |
| State 3 [40–60] | 100 | 0 |
| State 2 [60–80] | 100 | 0 |
| State 1 [80–\infty] | 100 | 0 |

### Table 2. Points of time of the observation “Current high” and “Temperature”

| State | 0–30 | 30–40 | 40–\infty | State | False | True |
|-------|------|-------|-----------|-------|-------|------|
| Time [h] | 0–2.40 | 2.40–2.55 | 2.55–3.10 | Time [h] | 0–3.00 | 3.00–3.10 |
| Time (min) | 0–165 | 165–180 | 180–190 | Time (min) | 0–185 | 185–190 |
| Time-slice number | 0–33 | 33–35 | 35–38 | Time-slice number | 0–37 | 37–38 |
4.1. Results

As stated above, the failure occurs after approximately three hours and ten minutes, i.e. after 190 min. In the model this is time-slice 38:

\[
\text{time-slice-number} = \frac{\text{time}}{\text{t}_{\text{slicen}}} = \frac{190 \text{ min}}{5 \text{ min}} = 38
\]  

(10)

The condition of the bearing in the absence of any observations is reflected in the degradation model as follows Figure 10:

According to this figure, there is a 23% probability that the condition of the bearing is in state 2, but up to 74% that it is in state 1. When the observations shown in Tables 1 and 2 are included, the following results emerge (Figure 11):

When including the observations, the condition in time-slice 37 (185 min after starting the test) is very bad. It is expected that a failure will occur with a 95.94% probability. Predicting time-slice 38 (190 min after starting the test), the probability that a failure occurs is over 97%. As stated above, failure did in fact occur in time-slice 38. This demonstrates the close agreement of the degradation model developed here with reality. In addition, using this case study, the importance of bi-directional reasoning in DBNs is demonstrated. The results including observations are completely different from the results not including any observations: without observations, the condition of the bearing is estimated to be as good as brand new with a probability of approximately 74%, but including observations the bearing is expected to fail with a probability of over 97%.
5. Conclusion

This paper develops a degradation model based on DBN. The model is based on a P–F-curve which is implemented in such a DBN. The P–F-curve describes typical degradation processes, is modeled based on a reversed exponential function and includes the Markovian property. The characteristics of the P–F-curve can be influenced by maintenance activities, operating conditions as well as by observations made while the component or machine is working. The maintenance activities as well as the operating conditions are described as time-invariant parts which influence the durability of the machine or component. The observations, on the other hand, are represented as time-variant parameters. They are caused by potential failures. Therefore, this predictive model can consider both influencing factors as operating conditions and maintenance activities and observations indicating potential failures (symptoms of failures). Moreover, a so-called damage-index simplifies the complex results of the model, thereby making them more accessible. It integrates both the five condition states of the component or machine and an uncertainty factor (caused by sensor inaccuracy and the estimation of probability values). As shown in chapter 3.3 in a case study of the degradation of bearings, reliability data or so-called life time data which is to date traditionally employed to estimate the durability of components in the case at hand only roughly indicates how long the components regarded here (bearings) will be working. By contrast, the model presented can, by including all life data it receives from sensors, machine operators and the machine control, accurately predict the condition of the components regarded. Moreover, the model has several more strengths. Firstly, it can predict the condition of the system including all life data it got before from sensors, machine operators, and the machine control. Secondly, the model is able to learn while it is running.
