MODELOWANIE PLANOWYCH PRAC EKSPLOATACYJNYCH 
PRZY NIEJEDNOLITYM POJAWIANIU SIĘ DEFEKTÓW I ZMIENNYM PRAWDOPODOBIEŃSTWIE WYKRYCIA DEFEKTU

MODELING PLANNED MAINTENANCE WITH NON-HOMOGENEOUS DEFECT ARRIVALS AND VARIABLE PROBABILITY OF DEFECT IDENTIFICATION

Wenbin WANG

1. Introduction

The time based maintenance policy is still one of the effective maintenance strategies currently used in industries despite the popularity of condition based maintenance. It is particularly useful for those plant items which follows an exponential deteriorating characteristic and cannot be monitored by condition monitoring. To clarify the objective of the type of the maintenance strategy we are concerned with here, consider a plant item with a maintenance strategy of a time based maintenance every period \( t \) time units with repair of failures undertaken as they arise. The maintenance activities at a planned maintenance epoch normally consist of a general inspection of the operational state of the plant by a check list, repairs to identified defects and those reported by the operators and some other maintenance activities such as greasing, cleaning and calibrating. Such an interval \( t \) is clearly a key decision variable in such a strategy and the determination of the optimal planned maintenance interval is the subject of this paper.

There have been numerous papers contributed to this topic [7, 9], from the earlier work of Barlow and Proschan [2] to a recent review of Nicola and Dekker [8] who surveyed some developments in the area of optimal maintenance of multi-component systems since 1991. For a survey of maintenance models prior to 1991, see Cho and Parlar [5]. The work reported in this paper however follows the earlier work of Christer and Waller [3] on the use of a concept called the Delay Time (DT) to model the planned maintenance interval with respect to optimising a criterion function of interest. The delay time concept and associated modelling techniques for inspection modelling have been reported in many papers over the last 20 years, see Christer and Waller [3], Baker and Wang [1], Christer et al [4], Christer (1999), Wang and Christer [11], Wang and Jia [13], Jones et al [6] and Wang [14, 15]. The delay time, as will be discussed in length in the next section, provides a window for inspection and repair as such the...
relationship between the number of failures and planned maintenance intervals can be established.

2. The delay time concept and its inspection modelling

We are interested in the relationship between the performance of equipment and maintenance interventions, and to capture this, the conventional reliability analysis of time to first failure, or time between failures, requires enrichment. We consider a repairable item of plant. It could be, say, a component, a machine, or an integrated set of machines forming a production line, but viewed by management as a plant unit. The interaction between the maintenance concept and equipment performance may be captured using the DT concept presented below.

Let the item of plant be maintained on a breakdown basis.

The time history of breakdowns or failure events forms a random series of points. For any one of these failures, the likelihood is that had the plant been inspected at some point just prior to failure, it could have been seen that all was not well and a defect was present which, though the plant was still working, would ultimately lead to a failure. Such signals include excessive vibration, unusual noise, excessive heat, surface staining, smell, reduced output, increased quality variability etc. The first instance where the presence of a defect might reasonably be expected to be recognised by an inspection had it taken place is called the initial point of the defect, and the time to failure from this initial point is called the delay time of the defect, see Fig. 1. Had an inspection taken place in (u, u+h), the presence of a defect could have been noted and corrective actions taken prior to failure. Given that a defect arises, its delay time represents a window of opportunity for preventing a failure. Clearly, the delay time h is a characteristic of the plant concerned, the type of defect, the nature of any inspection, and perhaps the person inspecting.

Fig. 1. The delay time for a defect

To see why the DT concept is of use, consider Fig. 2 incorporating the seven defects along with the initial points and delay times associated with each failure arising under a breakdown system. Had an inspection taken place at point (A), one defect could have been identified and the seven failures could have been reduced to six. Likewise, had inspection taken place at point (A), (B) and point (C), four defects could have been identified and the seven failures could have been reduced to three. Fig. 2 demonstrated that provided it is possible to model the way defects and failures arise, that is the rate of arrival of defects λ(u), and their associated delay time h, then the DT concept can capture the relationship between inspection frequency and the number of plant failures.

Despite the extensive research in DT modelling over the past 20 years, there are still some issues remaining to be solved. In this paper we tackle two of them. Here we first clarify some confusion regarding inspection and PM. Generally speaking all actions done at a planned maintenance epoch should be called preventive maintenance contrasted with breakdown only maintenance. It is noted that at a planned maintenance epoch, not only an inspection and the removal of identified defects were carried out, some other types of maintenance activities were also carried out, such as changing or topping up lubricant oil, greasing the bearing, cleaning and calibrating etc. These types of activities are not designed for defect identification and removal, rather, if done appropriately, they can reduce the number of future defects arising, and therefore the resulting failures. However, to distinguish the three activities carried out at a planned maintenance epoch, we narrowly defined the maintenance activities done at a planned maintenance epoch excluding inspection and defect removals as Preventive Maintenance (PM). How to address the impact of PM activities gradually decrease and approaches a constant as time goes.

Another contribution made in this paper is the relaxation of the assumption of a constant defect identification probability. Many papers in DT studies have used this assumption to describe the quality of inspections, but always assumed it is constant no matter the defect is in its earlier stage or at the advanced stage just before the failure. In this paper however we allow the probability of defect identification be a function of the delay time so that the technician will have a different degree of easiness for identifying the defect if it is there. Obviously, the probability of defect identification should increase towards the end of the delay time.

To put the above situation into a framework of modelling, the following are the characteristics of modelling a piece of production plant using the DT concept.

- Many failures can be characterized by a two-stage failure process, that is, from new to the initial point of a defect, and from this point to failure if the defect was not attended to.
- The initial points of defects are random and as such can be modelled by a stochastic process along the time axis.
- The rate of the arrival of defects is an increasing function of the time since the last PM/inspection but should tail off toward a constant if the PM interval is long. This implies that the impact of PM activities gradually decrease and approaches a constant as time goes.
- The time interval (delay time) between the initial point of the defect and failure is uncertain and can be modelled by a probability distribution function.
- By inspections at discrete points, one can detect if the defect has initiated with a certain probability, and then maintenance decisions can be taken to avoid failure.
- The defect identification probability is an increasing function of the delay time with an upper bound. This means that one has a less probability to detect the defect when the defect is in its earlier stage than that in the later stage.

Fig. 2. Defect, delay and failure process incorporating three inspection at A, B and C, where ○ --- initial points; ● --- failure points
• Failures can be observed immediately and need to be rectified through corrective maintenance (CM) actions at the time of the failure.
• All actions (inspection, defect removals at inspection, PM and CM actions) may cost money and result in downtime.
• The problem under study is to decide on the optimal maintenance intervals that can be periodic or non-periodic.

One needs to build models to determine the optimal maintenance intervals based on some objective function (or performance measure) – either downtime or cost or reliability.

3. Assumptions and notation

Assumptions
1) A multi-component complex engineering system with many components.
2) All failures follow a two-stage failure process as defined by the well known delay time concept, Christer and Waller [3] that is, from new to the initial point of identification of a hidden defect, then from this initial point to an eventual failure caused by the defect if not attended to.
3) If an inspection service is performed during the second stage before the failure, the defect may be rectified by either a repair or replacement of the same or upgraded part.
4) An inspection is not guaranteed to identify the defect. It has a probability \( r \) for successfully identification. However, \( r \) is also a function of the delay time in that \( r \) increases towards the end of the delay time, which is commonly expected as the defect propagates it become more obviously to be identified.
5) The arrival of defects follows an NHPP with the rate of the arrival of defects being a function of the time since the last PM. This assumption allows us to model the impact of PM type of activities such as greasing, cleaning and adjusting at a planned maintenance epoch.
6) The NHPP process over each PM interval is the same if the interval is constant.
7) Once a defect arrived, it follows a delay time \( h \) before failure with a pdf., The pdfs of the delay times of all defects are not identical.
8) The interval between inspections is constant.

Notation
\( u \) the initial point of a random defect, 
\( t_i \) the time of the \( i \)th planned maintenance since new, 
\( \Delta = t_i – t_i-1 \) the planned maintenance interval, assumed to be constant for now,
\( D(\Delta) \) expected downtime per unit time,
\( C(\Delta) \) expected cost per unit time,
\( \lambda(u; t_i) = A – Be^{-C(u-t_i)} \) rate of the arrival of defects, 
\( r(h) = a – be^{-ah} \) the probability of defect identification, 
\( p(h) \) the delay time pdf. of defects, 
\( F_0(h) \) the delay time cdf. of defects, 
\( D_i \) the mean downtime per failure, 
\( D_i \) the mean downtime per maintenance/inspection, 
\( D_i \) the mean downtime per defect removal, 
\( C_i \) the mean cost per failure, 
\( C_i \) the mean cost per maintenance/inspection, 
\( C_i \) the mean cost per defect removal, 
\( \nu(t) \) the rate of the arrival of failures, \( t \leq t_i \).
\( E[N(t_i, t)] \) the expected number of failures over \( [t_i, t] \), 
\( E[N(u, t)] \) the expected number of the defects identified at \( t \).

The formulations of \( \lambda(u; t_i) = A – Be^{-C(u-t_i)} \) and \( r(h) = a – be^{-ah} \) need some explanation. Both formulas ensure an increasing function and approach constants \( A \) and \( a \) when \( u \) and \( h \) are very large. This is what we needed since we assume that the rate of the arrival of defects is smaller just after the PM and then gradually increases towards a constant long after the PM. This model delates the impact of PM on the rate of the arrival of defects. The same argument also applies to \( r(h) = a – be^{-ah} \) where \( r(h) \) should increase towards the end of the delay time to reflect the degree of easiness for identifying the defect, but constrained under \( a \).

4. Model formulation

From assumptions 6 and 8, it can be shown that when \( i \) is large, \( E[N(t_i, t)] \approx E[N(t_i, t)] \) and \( E[N(t_i, t)] \approx E[N(t_i, t)] \), so the long term expected measures of \( D(\Delta) \) and \( C(\Delta) \) are

\[
D(\Delta) = \frac{D(N_d(\Delta)) + D_N + D_N E[N_d(\Delta)]}{T + D + D_E[N_d(\Delta)]} \tag{1}
\]

and

\[
C(\Delta) = \frac{C_d E[N_f(\Delta)] + C_f + C_f E[N_f(\Delta)]}{T + D + D_E[N_f(\Delta)]} \tag{2}
\]

where \( E[N_d(\Delta)] \approx E[N(t_i, t)] \) and \( E[N_f(\Delta)] \approx E[N(t_i, t)] \) for all \( i \) when \( i \) is very large.

From equations (1) and (2) we can see that \( r(h) \) and \( \nu(t) \) are the two important formulas to be derived.

We first show that without any inspection interventions, then a defect originates in time interval \((u, u+du)\) will become a failure at time interval \([t, t+dt]\) after a delay time \( h \), see Fig. 3.

![Fig.3. Defect origination and delay time h](image)

The approximated expected number of defects within \([u, u+du]\) is \( \lambda(u)du \) when \( du \) is small and the approximate expected number of failures due to these defects in \([t, t+dt]\), defined as \( \nu(u, t)dt \), when \( dt \) is small, is given by

\[
\nu(u, t)dt = \lambda(u)du(\nu(t) – \nu(u))dt \tag{3}
\]

Integrating equation (3) with respect to \( u \) from 0 to \( t \) and divided it with \( dt \) gives

\[
\nu(t) = \int_0^t \lambda(u)(\nu(t) – \nu(u))du \tag{4}
\]

where \( \nu(t) \) is the rate of the arrival of failures at time \( t \).

Applying the same principle and now assuming that there are inspection interventions following the assumptions made earlier, see Fig. 4.

![Fig.4. Defect origination with inspections](image)
For those defects arose in \([u, u+du)\), \((t_i, t_i+d)\) and eventually become failures in \([t_i, t_i+d)\), \((t_i, t_i+d)\), there will be \(i-I\)-negative inspections without identifying the defects with varying probability of \(1-r(t_i-u)\), \(k=1,\ldots,i-1\), depending on the distance since \(u\). This is modelled by \(\lambda(u-t_i)\int_{t_i}^{t_i+d} (1-r(t_i-u)) f(t-u) du\).

For those defects arose in \([t_i, t_i)\), the expected number of failures given \(u\in[t_i, t_i)\) is given by \(\lambda(u-t_i)\int_{t_i}^{t_i+d} (1-r(t_i-u)) f(t-u) du\). Integrating \(\lambda\) over its respective intervals, \(\lambda(t_i, t_i)\), \(n=1,\ldots,i-I\) and \(\lambda(t_i, t_i)\), we have the rate of arrival of defects is given by

\[
\nu_i(t) = \sum_{i=1}^{\infty} \lambda(u-t_i) \left[ (1-r(t_i-u)) f(t-u) du + \int_{u}^{t_i} \lambda(u-t_i) f(t-u) du \right]
\]

for \(t\in[t_i, t_i)\).

Similarly it can be proved that \(\nu_i(t) \approx v(t)\) when \(i\) is large. Given \(v(t)\) is available, it is straightforward that the expected number of failures over \([t_i, t_i)\) is given by

\[
E(N_{f}(t_i, t_i)) = \int_{t_i}^{t_i+d} \nu_i(t) dt
\]

The expected number of defects found at an inspection point, say, \(t_i\) is also a Poisson variable with the mean given by,

\[
E(N_f(t_i)) = \sum_{i=1}^{\infty} \lambda(u-t_i) \left[ (1-r(t_i-u)) r(t_i-u)(1-F(t_i-u)) du + \int_{u}^{t_i} \lambda(u-t_i) f(t-u) du \right]
\]

Because of \(v_i(t) \approx v(t)\), it follows that \(\approx E(N_f(t_i, t_i)) \approx E(N_f(t_i, t_i))\) when \(i\) is large. We can also show that \(E(N_f(t_i)) \approx E(N_f(t_i))\) when \(i\) is large. This provides a steady state condition so that equations 1 and 2 can represent the expected downtime and cost per unit time over an infinite horizon. It can be shown that when \(\lambda(u); t_i = \lambda\) and \(r(h) = r\), equations 6 and 7 reduce to

\[
E(N_f(t_i, t_i)) = \int_{h}^{\infty} \lambda(h) F(h) dh \text{ and } E(N_f(t_i)) = \int_{0}^{\infty} \lambda(h) F(h) dh
\]

as given in Wang [14]. In equations (6) and (7) we assumed that the downtimes caused by failures during \([t_i, t_i)\) are small compared with the duration of the interval, and therefore are ignored. This allows the integration over the whole interval. However, the downtime of failures must be there in equation (1) since it is the objective to be minimised. This assumption is really for a computation purpose and can be relaxed at the expense of more mathematics.

Because of the involvement of non-constant probabilities for defect identification, equations (6) and (7) cannot be evaluated analytically and as a result, they must be calculated numerically.

We present in the next section a numerical example to demonstrate the modelling idea.

5. Numerical example

The parameters used in this example are chosen according to our experience in past delay time modelling, see Tab. 1

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>α</th>
<th>(D_1)</th>
<th>(D_2)</th>
<th>(D_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>0.01</td>
<td>0.001</td>
<td>0.9</td>
<td>0.2</td>
<td>0.001</td>
<td>0.01</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

Note that \(\lambda(u; t_i) = A - Be^{-c(u-t_i)}\), \(r(h) = a \text{ and } f(h) = \text{constant}\).

The delay distribution was chosen as the exponential distribution with scale parameter \(\alpha\), but surely other distributions can be used since our calculation was done numerically anyway. Using the parameter values in Tab. 1 and equations (1), (6) and (7), the expected downtime per unit time in terms of the maintenance interval is shown in Fig. 5.

Fig.5. Expected downtime per unit time

It can be seen that \(\Delta t=28\) is the optimal maintenance interval. To see the impact of model parameters on the chosen optimal maintenance interval, we selected two extreme cases for a test. First we tried to set \(B=0\) and \(c=0\) which is the case of constant defect arrival, \(\lambda=A=0.02\) and constant defect identification probability, \(r=a-b=0.7\). This corresponds to the case of a higher defect arrival and lower defect identification compared with the case in Fig. 5. The second is by setting \(C=0\) and \(b=0\) which is the case of \(\lambda=A-B=0.01\) and \(r=a=0.9\) corresponding to a lower defect arrival and high defect identification. Fig. 6 shows the result of the expected downtime per unit time in terms of the maintenance interval. It can be seen that in the first case the expected down time is higher as expected and the optimal maintenance interval is shortened to \(\Delta t=21\). For the second case the optimal maintenance interval is lengthened to \(\Delta t=35\) with a relatively lower expected downtime.

Fig.6. Expected downtimes in case 1 and case 2
From Fig.s 5 and 6 we can conclude that different set-ups of the rate of the arrival of defects and the probability of defect identification do impact upon the expected downtime and the determination of the optimal maintenance interval. This illustrates the necessity of carefully examining the set-ups of both the rate of the arrival of defects and the probability of defect identification. Most importantly, it confirms our initial modelling assumption that the effectiveness of PM and variable defect identification ability do influence the final optimal decision and the expected downtime. Of course this is only a case example and one may argue that if different combinations of the model parameters were chosen the result may not be this obvious. We agree with this point and it is not our intention to show that the PM effect and variable defect identification probability must be considered. The final choice of model must be done through a rigorous statistical testing using some criterion function such as AIC, Baker and Wang [1].

To demonstrate graphically that the expected numbers of failures and defects identified at planned maintenance epochs are constant when $i$ is large, we used the same model parameters in table 1 and calculated the above two statistics using equations 6 and 7 shown in Fig.s 7 and 8. It can be seen that only after $i=4$ both statistics become almost constant.

6. Conclusions

This paper developed and discussed two new contributions using the delay time concept in the modelling and optimisation of planned maintenance. It showed that by allowing a non-constant rate of the arrival of defects after a PM and a variable probability of defect identification, the decision model responded to the changes and produced what we expected. The contributions made resulted from our observations in maintenance practice but also followed our common sense. The model developed can also be used for evaluating the effectiveness of PM activities and the efficiency of defect identification through the examination of the parameters within $\lambda(u); t_i$ and $r(h)$. The numerical example confirmed the impact of these changes made upon the maintenance decision.

There is a considerable scope for maintenance modelling to impact productivity upon current maintenance practice. This paper reported the use of only one modelling concept for modelling maintenance and inspection practice, but the potential is still to be fully explored. The delay time concept is a natural one within the maintenance engineering context. More importantly, it can be used to build quantitative models of the maintenance practice of plant items, which has proved in practice to be valid. The theory is still developing, but so far there has been no technical barrier to develop DT modelling principle to any plant items studied.

However, there are many obstacles to be overcome before putting the model developed in this paper into practice. A notable one is the estimation of model parameters, particularly the parameters within the rate of the arrival of defects and the probability of defect identification. There are several methods available to estimate the delay time model parameters, Wang (2008a), but with more parameters to be estimated than the conventional delay time models, it requires sufficient failure and maintenance data to get any significant and valid estimates.

**********

The research reported here is partially supported by EPSRC under grant numbers EP/G023042/1 and EP/F038526/1

**********

7. References


Prof. Wenbin WANG, Ph.D.
Centre for Operational Research and Applied Statistics
University of Salford
Salford M45 4WT, UK
e-mail: W.Wang@salford.ac.uk