EXPLORING CRITICAL FAILURE MODES IN THE RAIL ENVIRONMENT AND THE CONSEQUENTIAL COSTS OF UNPLANNED MAINTENANCE

P.D.F. Conradie¹, N.F. Treurnich²
Department of Industrial Engineering
Stellenbosch University, South Africa
¹prasa@sun.ac.za
²nicotr@sun.ac.za

ABSTRACT

This study explores in-service failure modes for rolling stock in the rail environment, identifies the most critical failures and explores the consequential cost of these failure modes. Rolling stock is maintained according to maintenance plans with a major goal being the prevention of in-service failures, but due to the nature of the equipment not all failures can be prevented. In-service failures normally result in train delays or the cancellations of trains not only disrupting commuter services but also causing financial losses.

The typical failures of rolling stock are analysed using data from the facility maintenance management system. The critical failure modes are identified and classified according to cause, severity, consequence and frequency parameters. A decision model is employed to classify the criticality of the failure modes.

The most prominent critical failure modes are analysed to determine root causes, to conclude the investigation. Areas are identified where the focus of future investigation and planned maintenance will have the most significant impact.
1 INTRODUCTION

In asset intensive organisations, strategic asset management is a key to the success of the organisation. Assets fail because of a number of reasons at random times and intervals, and disruption as a result of these failures has consequential effects that cannot always be quantified. Assets should therefore be maintained and operated to predetermined standards and best practices to minimise the impact of failures.

The concepts of Asset Management and Asset Maintenance are often misunderstood. Asset Management refers to the activities and principles over the whole life cycle of the physical asset, to achieve the stated output [1]. Asset Maintenance is part of Asset Management and refers to the actions to retain an item in, or restore it to a state in which it can perform a required function [1].

In this paper the Western Cape branch of a South African passenger rail company is used as a case study, and in-service failures of rolling stock are explored to determine the contribution of the different failure modes to reliability. Delays of trains not only cause disruption to the service but have a financial implication on the business. The contribution of each failure mode to the overall loss is calculated and a cost model is developed to quantify cancellations and delays as a result of these failures.

The condition and performance of other assets and departments have a significant influence on the performance of the rolling stock. There are a number of interfaces that can be taken into account, such as the wheel-rail interface, the pantograph-overhead interface as well as the human interface. For simplicity these influences are excluded from the scope of this paper.

2 BACKGROUND

In South Africa the passenger rail operator also own the infrastructure, unlike railroads in other countries where there are different asset owners. This should be a benefit to the rail operator but because of an imperfect track record on punctuality, the company has been stigmatised as “unreliable” over the past few years, which had a negative impact on rail passenger numbers. Maintenance of the rolling stock fleet and the prevention of failures have an influence on reliability and punctuality, which can be controlled and managed. The contribution thereof will be discussed further.

The service provided by a rail operator can be rated in terms of many parameters and from different perspectives. One such parameter from the rail passenger point of view is the punctuality of the train service, which is probably the most widely used reliability measure in practice [3]. Reliability from the rail operator point of view can be defined around functionality or ability of the rolling stock to perform its duty, and in this paper more emphasis will be placed on this perspective. Punctuality of a train service is a key performance indicator (KPI) valued highly by the rail passenger [5][6], and for the evaluation of a railway system, objective measures like cancellations, delays and the number of realized connections between trains can be measured [5]. Cancellations and delays are also used as key measures of the reliability by the company in our case study.

Many studies have been done on reliability in the railway industry and the effect thereof on rail passengers. In a study Kingham [6] found that reliability is a key factor that will convince vehicle drivers to switch to public transport. According to Cox et al [7] increased passenger density and overcrowding as a result of delays can contribute to low productivity and efficiency in tired workers. Rietveld [8] examined the unreliability of public transport and the effect thereof on travel time, Olsen and Haugland [4] studied influencing factors on train punctuality while Skuce [9] investigated rail passenger satisfaction in the South African environment. These studies found punctuality to be a main reason why people are not making use train services.
Before examining failures and their consequences, a common frame of reference with regard to the principles of failures and reliability need to be defined. A failure occurs whenever a system or component no longer operates within the expected or designed specification, which includes breakdowns and out-of-specification performance. A general definition for reliability is the probability that an item will perform a required function without failure under stated conditions for a stated period of time [10].

Todinov [11] suggests a theoretical framework that link reliability and losses from failures, and he demonstrates, contrary to conventional reliability analysis, that maximising the reliability of a system does not necessarily minimise the losses from failures. The aim is therefore to optimise reliability by taking all factors into account. If maintenance is effectively planned and executed, reliability can increase that can result in less corrective maintenance (CM) required with more opportunity for preventative maintenance (PM) [12].

The effectiveness of a maintenance plan is a key issue, and Ahren & Parida [13] define Maintenance Performance Indicators (MPIs) as a benchmarking tool to measure effectiveness. A typical application of MPI will be to express CR as a ratio to PM and compare to the industry benchmark of 20%. This ratio can indicate the effectiveness of the maintenance planning that has a direct influence on reliability. Pham and Wang [14] define the effectiveness of maintenance around the concept of “imperfect maintenance” where the degree to which the operating conditions of an item is restored by maintenance.

In the background aspects discussed so far it is argued that reliability and failures are two aspects that are critical and must be managed, that they are related and have an influence on each other. It is important to manage and optimise both as a key for business success but care needs to be taken when doing so.

2.1 Cost Of Maintenance

It was mentioned earlier that asset management has a big influence on the success of the organisation, and part of asset management is the maintenance thereof. In a case study by Ahren & Parida [15] it is mentioned that “maintenance is one of the largest controllable expenditures for the railway industry, as it can reduce cost and improve equipment effectiveness, reliability and performance”. Unfortunately the maintenance department is not seen as value adding to most organisations and regarded as an expense account, which is a popular target for cost reduction programs [15].

In a study by Cavalcante and Almeida [16] they argue that if the average cost of CM is larger than the cost of PM, then only it will be beneficial to predict a failure as there will be a cost saving. Salonen et al [17] comment in a study that the cost of maintenance can be classified in direct cost, indirect cost and non-realized revenue cost, where the latter refers to the loss of income because of reduced sales, missed deadlines, as well as other related factors. Both authors mainly take into account the direct cost of maintenance associated with a failure, but because of a failure there are second line consequences or a ripple effect that can be quantified as follows:

- the change in perception of the customer and future support willingness
- the indirect cost to the customer
- the cost to the economy
- loss in market opportunity

The most prominent of these effects are included in the cost model in the next section.

A popular maintenance strategy is the PM strategy that is used to increase equipment lifetime, reduce downtime and reduce the risk of failures. In our case study a PM strategy is also used and train sets are scheduled on a time base in order to perform maintenance. In-service failures result in CM which is a reactive maintenance approach, triggered by the unscheduled event of equipment failure [18] that should be avoided. Tsang [20] also points out that the reasons for performing PM are to prevent failure, detect the onset of failure
and to discover hidden failure. Higgens [12] agrees with Tsang that failures should be analysed, root causes identified and managed according to the maintenance plan that will reduce the number of instances of CM, and to convert findings of the analysis into PM to prevent failures.

A method to analyse failures and determine the root causes is the Failure Mode and Effects Analysis (FMEA), that is defined as “a procedure by which each potential failure mode in a system is analysed to determine the results or effects thereof on the system, and to classify each potential failure mode according to its severity” [10]. It can be expanded to FMECA by including criticality of the failure modes in the analysis.

3 METHODOLOGY

In South Africa the safety criteria for rail operations are governed by the Rail Safety Regulator (RSR) [19]. One requirement of the RSR is that maintenance data must be recorded and maintained. This is one reason why the case study uses a Fleet Maintenance Management System (FMMS) [20] to capture maintenance data, and accurate reports can be extracted by means of GQL (General Query Language).

In this study maintenance reports extracted from the FMMS were manipulated and analysed. From the analysis it was determined which groups of components cause the most in-service failures, and a correlation between the number of failures and the duration of failures were established. Passenger and various other statistics were used to determine:

- the cost of a train delay
- the cost of a train cancellation

These were used to determine the consequential cost of each failure mode.

The critical failure modes were classified according to cause, severity, consequence and frequency, and a decision model is employed to classify the criticality of these failure modes.

4 DATA ANALYSIS

4.1 The effect Of In-Service Failures On Punctuality And Reliability

As discussed in the literature study, punctuality is widely used as a reliability measure. Railroads define their punctuality as the probability that a train will arrive at the final destination within a delay of less than a certain margin. An international margin of 5 minutes is used, however in some countries like Switzerland, Netherlands and France 3 minutes is used. It is important to note that this measure only refers to the arrival at the final destination and intermediate stops are neglected [8].

Table 1: Punctuality Comparison Between Different Rail Operators

<table>
<thead>
<tr>
<th>Country (post 2009)[21]</th>
<th>Operator</th>
<th>Margin used</th>
<th>Probability of arriving within margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switzerland</td>
<td>SBB</td>
<td>3 minutes</td>
<td>95.0%</td>
</tr>
<tr>
<td>Germany (Aug 2011)[21]</td>
<td>Deutsche Bahn</td>
<td>5 minutes</td>
<td>93.0%</td>
</tr>
<tr>
<td>Austria (Jan-Jun 2011)[21]</td>
<td>OBB</td>
<td>5 minutes</td>
<td>96.5%</td>
</tr>
<tr>
<td>United Kingdom (UK) (Jan-Feb 2011)[21]</td>
<td>London Underground</td>
<td>10 minutes</td>
<td>96.4%</td>
</tr>
<tr>
<td>Country</td>
<td>Operator</td>
<td>Margin used</td>
<td>Probability of arriving within margin</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------------------------</td>
<td>-------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>South Africa, Johannesburg (March 2012) [22]</td>
<td>Bombela Operating Company</td>
<td>3 minutes</td>
<td>98.98%</td>
</tr>
<tr>
<td>South Africa, Cape Town(2011)</td>
<td>Case study</td>
<td>5 minutes</td>
<td>84.5%</td>
</tr>
</tbody>
</table>

In Table 1 the punctuality of the company in our case study is compared to other rail operators and it can be seen that the company uses the international norm of 5 minutes, but has room for improvement. There can be many reasons for this and in this paper it is assumed that delays can have an effect on punctuality, whether the delay is as a result of failures or operational in nature. The paper will focus further on in-services failures and explore the effect on cancellations and delays.

It is not within the scope of this paper to determine the root causes of punctuality, but the contribution of in-service failures on punctuality and reliability is significant.

4.2 The Consequence Of In-Service Failures

In the literature background a failure is defined as a system or component which no longer operates within the expected or designed specification. When a failure occurs while the train is in-service, these failures can result in cancellations or delays, and is referred to as in-service failures.

As discussed earlier, cancellations and delays are KPI’s used by the company in our case study to measure performance. An in-service failure can cause a cancellation and/or delay, it often happens that the knock-on effect of one failure is catastrophic on the service delivery. The contribution of each department to the overall cancellation and delays were calculated and compared.

4.3 Contribution Of Each Departmental To Cancellations And Delays

Information obtained from the Train Operations department was analysed with a special focus on the contribution of the Rolling Stock department to cancellations and delays. In 2011 there were more than 193 000 train trips and based on this number more than 15% of the trains were delayed (table 1) and 12% cancelled [23].

It was questioned by the researchers whether the number of trains delayed is an accurate enough measure. The reason for questioning this measure was because the duration of delays will be different dependant on the severity and the cause of the delay. It was found however that for 2011 the two measures correlate within 0.2% to 5.5% for the different departments. For the Rolling Stock department the variance was 1.2% and it was concluded that either the number of trains delayed or the duration of delays can be used.

Furthermore it was found that in 2011 the Rolling Stock department was responsible for 24% of train delays (ranked second) and 78% of cancellations (ranked first) [23]. Nearly 25% of the cancellations caused by the Rolling Stock department occurred in-service. The rest of the cancellations were as a result of train sets that were not available for service, mainly due to maintenance and vandalism.

From these figures it is evident that the Rolling Stock department has a large influence on the punctuality and reliability of the train service.

4.4 Trends In Rolling Stock Department

Cancellations and delays caused by the Rolling Stock department since 1998 were investigated. In this section the statistics around cancellations and delays are discussed.
4.4.1 Delays

In the previous section the conclusion was made that either the number of trains delayed or the duration of delays can be used as a delay measurement. In Figure 1 historical information on the number of trains delayed ($D_{\text{trains}}$) is compared to the time delayed ($D_{\text{time}}$) [23].

The graphs were superimposed which highlight the following:

- In Figure 1 it is unclear why the trend in delays is upwards. The following factors are presented as contributing factors:
  - The trains are becoming older and the majority of the fleet of 406 trains were built between 1958 and 1985. Typical problems are that parts fail more often, lead times are longer, the number of suppliers are limited and spare parts are more difficult to obtain.
  - Business processes are inefficient and outdated.
  - Loss of technical expertise

- Also in Figure 1 the trends for both data sets ($D_{\text{trains}}$ and $D_{\text{time}}$) display an increasing trend. This is a significant perspective on the data and it implies that if 2011 is compared to 1998, $D_{\text{trains}}$ (number of trains delayed) is 5 times more whereas $D_{\text{time}}$ (time trains were delayed) is 8 times more. It can be seen that both graphs follow similar fluctuations until 2008, where after $D_{\text{time}}$ increase at a faster rate than $D_{\text{trains}}$. This increase can be explained by Figure 2 where the average time of a train delay ($D_{\text{time}}/D_{\text{trains}}$) more than doubled from 1998 to 2008.

- While the trends of $D_{\text{time}}$ and $D_{\text{trains}}$ seem to be exponential, the trend of the average delay per train ($D_{\text{time}}/D_{\text{trains}}$) displays a flattening trend (Figure 2) that is a positive sign in that it is starting to normalise. In the opinion of the authors, the reasons contributing are:
  - From 2009 to 2011 $D_{\text{time}}$ stayed almost the same while $D_{\text{trains}}$ increased with 35%. This can be interpreted as:
    - the number of trains that are delayed is increasing at a faster pace $D_{\text{trains}}$
    - The deduction can also be made that $D_{\text{time}}$ is stabilising although still on a upward curve

![Figure 1: Delays Caused By The Rolling Stock Department Expressed In Number Of Trains Delayed, And In Time Delayed [23]](image-url)
In conclusion it can be noted that delays have increased approximately tenfold compared to a decade ago and should be listed as a KPI for the Rolling Stock department. Delays are more frequent and also becoming longer and should be closely monitored.

4.4.2 Cancellations

The number of cancellations caused by the Rolling Stock department displays an increasing trend as shown in Figure 3 and has increased more than tenfold since 1998. Although this is of concern, not enough information could be obtained to quantify the extent of cancellations. Cancellations are more complex than delays as the unavailability of one train set can result in a number of cancellations which is more difficult to manage.

4.5 In-Service Failures For The Rolling Stock Department

4.5.1 In-Service Failure Process Flow

To understand the discussion on the analysis that follows, the in-service failure process flow is presented (refer to Figure 4). When the train driver experiences an in-service failure, the driver reports the failure to the Train Operations Control Centre (TOCC). As a first line
maintenance person the driver will evaluate and attempt to resolve the failure, but because of a lack of technical experience or knowledge, it can happen that the driver reports a failure that can be resolved by himself or describe the fault incorrectly. This can result in the improper corrective action being initiated.

After the failure is reported and a job is registered, two parallel, non-related process lines are followed.

1. The failure is attended to by the Failure Department, the reason for the failure noted and the job closed
2. Management investigates the failure and allocates it to the relevant department who will take ownership and responsibility of the consequence of the failure. Taking ownership implies that the cancellations or delays are added to the totals for the department, as described in the previous section.

An analysis was done on in-service failures [24] of 2011 for the region where 85 train sets are running and nearly 1600 in-service failures were experienced. The company grouped the components that fail frequently in component groups (refer to group descriptions in Table 2). Based on the fault codes, the occurrence of each component group was calculated. One of the key drivers of this process is that the correct fault code must be generated before the job is closed and it must be noted that 29.5% of failures have no fault codes which makes the statistics questionable and not totally reliable, but the results are nevertheless a good indication of trends.

4.5.2 Contribution And Frequency Of Component Groups To In-Service Failures

The contribution of each component group was calculated in terms of cancellations and delays, and the results are summarised in Table 2.
Table 2: Contribution Of Each Component Group On Cancellations And Delays [24]

<table>
<thead>
<tr>
<th>Group</th>
<th>Group description</th>
<th>Column I</th>
<th>Column II</th>
<th>Column III</th>
<th>Column IV</th>
<th>Column V</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Electronic Control Equipment</td>
<td>47%</td>
<td>48%</td>
<td>18.0 min</td>
<td>44.3min</td>
<td>31%</td>
</tr>
<tr>
<td>P</td>
<td>High Voltage and switch equipment</td>
<td>23%</td>
<td>23%</td>
<td>18.4 min</td>
<td>50.4 min</td>
<td>30%</td>
</tr>
<tr>
<td>M</td>
<td>Traction/Auxiliary machine and controls</td>
<td>14%</td>
<td>14%</td>
<td>19.3 min</td>
<td>52.2 min</td>
<td>13%</td>
</tr>
<tr>
<td>O</td>
<td>Brake gear</td>
<td>2%</td>
<td>2%</td>
<td>17.5 min</td>
<td>29.7 min</td>
<td>5%</td>
</tr>
<tr>
<td>B</td>
<td>Cab and Saloon doors</td>
<td>6%</td>
<td>5%</td>
<td>21.1 min</td>
<td>46.5 min</td>
<td>6%</td>
</tr>
<tr>
<td>A</td>
<td>Air related</td>
<td>4%</td>
<td>5%</td>
<td>18.1 min</td>
<td>58.6 min</td>
<td>4%</td>
</tr>
<tr>
<td>G</td>
<td>Pantograph</td>
<td>3%</td>
<td>3%</td>
<td>18.9 min</td>
<td>54.0 min</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>Other components</td>
<td>1%</td>
<td>2%</td>
<td>13.9 min</td>
<td>20.5 min</td>
<td>4%</td>
</tr>
</tbody>
</table>

18 min ave 46 min ave

Explanation of Table 2:

- 85% of the number of delays (column II) is caused by failures in component groups E, P and M. It can also be seen that these 3 groups contain all the electronic and electrical components, and these components normally follow the bathtub failure curve. This may explain why so many in-service failures are experienced as some components are old, failures cannot always be predicted and the failure happens suddenly.

- Although contributing only 5% to the number of delays, Group B caused the longest average time per delay (column III, 21.1min). It can be contributed to the fact that cab and saloon doors are regarded as safety items by the RSR and must be repaired if not working. Therefore a train will not depart before the doors are fixed.

- Air related failures (group A) contributed also 5% to the number of delays, but took the longest to repair at 58.6min (column IV).

- Between groups E, P and M, group M causes both the longest time per delay (19.3min) and the longest time per event (52.2min), thus the complexity of failures and the severity of failures can be illustrated by this.

- Group A is the most complex failure group to repair although not contributing to a large number of delays or cancellations.
It can be seen that the ratio for the groups based on the time delayed (column I) or based on the number of delays (column II) correlated within 1%. This means that any one of the two ways to express delays can be used.

Detail statistics are as follows:

- 16% of failures were rectified by the driver attributing to 11.7% delays. This implies that the failures which the drivers can fix are small and are fixed quickly.
- 4.2% cancellations are caused by drivers
- If the Pareto principle is applied, 20% of the failures are responsible for 64% of delays measured in time
- Cancellation show a similar trend to that of delays in that the same 3 component groups caused 74% of cancellations

From the analysis it is evident that 3 component groups (E,P and M) cause most of the cancellations and delays. The Traction/Auxiliary machine and controls (Group M) is a critical group to explore further as the consequence and severity of this component group are relatively more than the other groups.

4.6 Cost Of Each Component Group To The Total Cost Of In-Service Failures

The total cost of a cancellation and a delay was calculated, and various sources were used to obtain statistics on:

i. Average number of passengers on a train [25]
ii. Number of passengers who commute to work, who earn a salary [25]
iii. Average number of trains per day [25]
iv. Demographic distribution of the population who use train as a mode of transport [26]
v. Average income per population group [27]
vi. Number of delays and cancellations for Rolling Stock department [25]
vii. Average train fares [27], duration of trips on Metrorail [28] and on a typical bus [29]

There are a number of assumptions in this cost calculation, and there are also a number of cost factors excluded from this calculation such as:

- The direct cost to repair the in-service failure is not taken into account
- The indirect cost to recover for lost production is excluded
- Non-realized revenue were taken into account, except for the second line consequential or ripple effect costs, as defined earlier
- The average number of passengers of a train was used. In peak hour all passengers on the train is assumed to be commuters and there are many more commuters on the train than average, therefore the cost per delay in peak hour is at least an order higher than on average.

Taken the above into account, the average cost of a delay is calculated. If a train is delayed it is assumed that the commuter will lose a portion of his salary. By using statistics i-vi it was calculated that the cost of a delay will be R506 per minute. Considering an average delay of 20 minutes per train, the average cost per delayed train is R10 000.

Using the same statistics, the cost of a cancellation is calculated, but the following is taken into account:

- Revenue lost because of the cancellation
- Direct loss due to higher cost of alternative transport for the passenger
- Lost wages because of longer travel time on the alternative transport
- Lost wages because of time wasted to arrange and waiting for alternative transport

It was concluded that the estimated average cost of a cancellation is R56 175.
With these cost figures, cancellations and delays of each component group can be expressed in monetary terms.

5 CONCLUSION AND RECOMMENDATIONS

In this paper an analysis of in-service failures in a South African passenger rail company has been conducted, and various statistics and ratios considered. Results and trends were obtained and presented without full detail to protect the intellectual property of the company.

The correlation between maintenance, cancellations, delays and punctuality were discussed, and the effect thereof on reliability cannot be ignored. While this article highlights significant statistics on historical information, it should be further investigated to predict future trends. Historical information alone is not suitable as a prediction of future trends, and by using suitable mathematical models future behaviour can be predicted which can result in less in-service faults.

The analysis of the in-service failures displayed noteworthy trends. The most evident trend is that cancellations and delays are increasing over the years. This trend can partly be attributed to ageing trains, emphasising the need for appropriate maintenance strategies for the age of the equipment. The authors are of the opinion that the increasing trend is not fully explained by the ageing trains and it appears as if this study identifies the existence of opportunities for more optimised maintenance.

The findings on the contribution of each component group are significant as it is a skew distribution and the list of top failures stays almost the same every month. It is recommended that the component groups be redefined to focus on more detailed failures, especially the top failure groups. Another suggestion is to change from component groups to functional groupings where a system belongs to a group. Reliability Block Diagrams can be beneficial to determine dependencies between different functional groupings.

There was an inadequate amount of information to do a full analysis on root causes because the information in the maintenance log was incomplete in terms of fault codes. It is recommended that the fault codes be entered on the job cards as a compulsory field, and once the groups are established the root causes can be analysed from these codes.

6 REFERENCES


[22] Newsletter, Bombela Operating Company (Pty) Ltd. BOC’nGauge. 2012(6)


[28] Metrorail Fare Calculator. [Online] [2011] [access 2012, April]; Available http://www.capemetrorail.co.za/_fare-calc/fares_index.htm
