An Approach to Improving the Performance of Rail Systems in a Design Phase

A practical example from the High Speed Line South tender

*Arjen ZOETEMAN¹, Egbert BRAAKSMA²
Delft University of Technology

Faculty of Technology, Policy and Management
Jaffalaan 5, PO Box 5015, 2600 GA Delft
Tel: +31-15-278-1135
Fax: +31-15-278-2719
E-mail: a.zoeteman@tbum.tudelft.nl

Summary
In the last few years the importance of maintenance in Design and Construct contracts for railway infrastructure has increased. Managing bodies put increasingly higher requirements to contractors for delivering availability, reliability and low costs of ownership. With regard to the Dutch high-speed line (HSL South), a so-called Infrastructure Provider Contract encompasses the design, construction, and financing of the total rail system, including a maintenance period of 25 years. Due to the fact that maintenance used to play a minor role in infrastructure projects, successful examples of optimising investment, maintenance costs, availability and reliability as a whole are almost absent. In this paper an approach to deal with the requirements is discussed.

Keywords:
Infrastructure Provider contracts, FMEA, life cycle costs, decision support system

¹ Researcher Life Cycle Costing for Rail Infrastructures at TU Delft
² Head of Innovation, Strukton Railinfra
1. Introduction

Setting performance requirements is becoming more and more common practice in the management of railway infrastructure. The recent development of EN50126, the European Standard for Reliability, Availability, Maintainability and Safety of Railway Systems, shows the changing attitude towards maintenance of railway assets [1]. Maintenance of railway assets is not regarded anymore “as something that needs to be done”, but more and more as a professional business delivering very important products for rail operations:

1. **Availability**: the time that the infrastructure is available for operations per calendar period. A part of the time the infrastructure is out of service due to planned possessions (preventive maintenance). Another part of the time the infrastructure can be unavailable due to infrastructure failures (corrective maintenance), possession over-runs or external factors, such as vandalism and bad weather.

2. **Reliability**: the time that the infrastructure is available for operations during the operation periods agreed. In other words, with regard to the reliability only the unplanned maintenance and repair is considered. The reliability depends on e.g. the asset quality and maintainability, the amount of preventive maintenance, and the failure restore times.

3. **Safety, noise, vibrations and riding comfort**: these aspects are covered in design criteria, maintenance thresholds (e.g. geometry control limits), inspection and failure response strategies (e.g. inspection frequencies and speed restrictions). The use of Safety Cases forms an important instrument to analyse the railway safety provided.

The current railway environment can explain the changing attitude towards infrastructure maintenance: the separation of infrastructure and operations and the subsequent introduction of user charging have triggered the need to specify the performance of the infrastructure and develop performance regimes. Infrastructure failures will disrupt the business of the transport operators and they will demand a guaranteed infrastructure performance\(^3\). On the other hand, preventive maintenance will require more planned unavailability. A **performance (payment) regime** should deal with the trade-off, considering the long-term costs of ownership for the infrastructure owner (life cycle costs) as well. In the Performance Regime the impacts of traffic disruptions are (often only partly) reflected in the penalties assigned. However, the performance required can be completely different for e.g. a regional line, a conventional main line with mixed passenger and freight traffic or a dedicated high-speed line. The design (e.g. track alignment, signalling and power supply systems), the required maintenance frequencies and the operational impacts (costs and revenues) of failures will vary a lot. The performance delivered therefore depends on many local factors, and ‘planning the infrastructure performance’ can be a tough job [2].

Recently RAMS requirements (Reliability, Availability, Maintainability and Safety) are increasingly being introduced in the design and construction phase, and contractors often have to maintain the assets during the first operational years. In this paper the consequences of the Infrastructure Provider Contract for the Dutch high-speed line are discussed. This contract has been an important change in the tender-policy in the Dutch railway sector: the contractor has to design, build, finance and maintain the complete railway assets (not including the substructure). In 2030 the assets will be handed over to the Dutch State with a guarantee period of five years. In section 2 the Infrastructure Provider Contract is considered. In section 3.1 an approach, taken in one of the consortia to analyse and optimise the rail system performance, is described. Section 3.2 deals with the obstacles faced and the results of the analysis. Section 4 contains the conclusions.

---

\(^3\) On the other hand, the infrastructure manager will also set and monitor its requirements to for instance the quality of the rolling stock in order to minimise the quality deterioration of its assets (e.g. avoiding ‘flat wheels’ and too heavily loaded wagons).
2. The HSL South Infrastructure Provider Contract

The High Speed Line South (HSL South) is the Dutch part of the high-speed link from Amsterdam to Brussels, Paris and London. In 2005 the operations (both domestic and international) on the HSL South should commence. The new tracks, with a total length of almost a hundred kilometres, will be designed for 300 kph. In fact it concerns two different track sections, since the trains will use the conventional network at Rotterdam. For the northern section in particular considerable investments will be made for the substructure: a piled concrete slab structure will support the rail system. A bored tunnel under the Green Heart of Holland is included. The civil works and the connecting infrastructure to the conventional network are (geographically) split up into seven contracts. The rail system, i.e. tracks and switches, power supply and signalling systems, electrical and mechanical (auxiliary) equipment, and some sound barriers, is tendered as a whole in the so-called Infrastructure Provider Contract.

The Infrastructure Provider (IP) has the obligation to Design, Build, Finance and Maintain the rail system. During the first 25 operating years the Dutch State will pay a quarterly fee to the IP. The IP will be penalised for a shortfall in performance: for a low asset condition and especially for train delays and cancellations so-called non-availability deductions will be made. The calculation of the penalties is based on the Performance Payment Mathematical Algorithm (PPMA) [3].

Since the consortia already had to submit their bids in 2000 without knowing the exact traffic conditions, the Dutch State has developed a so-called Reference Timetable. This timetable shows the forecasted trains for the years 2005-2030, completely specified as a timetable. This ‘hypothetical timetable’ will be used to calculate the train delays and cancellations. In other words, if in reality an infrastructure failure in e.g. 2008 may not have caused any problems, it is still possible that the IP is penalised based on the Reference Timetable. Delays resulting from faults of the train operator are, of course, not included: the Journey Time Deviation is the train delay ‘collected’ on the two high-speed track sections as a result of the faults and actions of the Infrastructure Provider. A train cancellation is included in the magnitude of a 3-hour delay. Consequential delays of trains scheduled later in the timetable (after an initial delay caused by failing equipment) are included. For the measurement of performance a so-called Performance Simulation Model will be developed by the IP that is to be approved by the Dutch State. This model is part of the Agreement between the IP and the State and will calculate the penalties for the failures and possession over-runs. These penalties are deducted from the so-called Performance Fee, which is the quarterly amount of money paid by the Dutch State to the IP.

The assignment of penalties by the PPMA is rather complicated. Without going too much into details, the Availability is calculated on a daily and monthly basis. The non-availability is the total Journey Time Deviation divided by the total Scheduled Journey Time (SJT). For an availability level below 99% penalties are assigned with a progressive penalty rate per train delay minute (from e.g. 500 to 3500 Euros - indicative). If the availability is lower than 90% on a single day, this is considered to be a Not Available Day (NAD); the part of the Performance Fee allocated to that day will not be paid. Figure 1 gives an impression of the relation between the annual availability level and the amount of penalties not considering the impact of the irregular occurrence of failures during the year. As figure 1 shows, the number of trains in the reference timetable for the HSL South does not influence the total amount of penalties.

It should have become clear by now that the availability of the rail system forms a critical aspect in the design and maintenance concept of the High Speed Line South. This means that the consortia bidding for the contract needed to have insight into the performance of their design and had to develop a maintenance strategy already during the Tender Phase (Oct. 2000-March 2001).
The most important criterion for the Dutch State was the net present value of the total Performance Fees to be paid, including the impact of penalties. Therefore the question is posed: which system availability can be provided in order to realise the lowest total costs of ownership for the Dutch State?

It can be assumed that the engineering and maintenance efforts increase progressively to provide a higher availability level, whereas the returns on this investment become smaller: the penalties saved (reflecting the operators' loss of revenue) decrease progressively for a higher availability level, as was described earlier in this section. This means that each percent more availability results in less savings in penalty costs, i.e. the amount of money available for investments. Figure 2 shows this relationship graphically: it becomes clear that somewhere there must be an optimal level of investment in design and maintenance.

The penalty rates applied in the Infrastructure Provider Contract will stimulate aiming for a high availability level [4], but to select the most cost-effective and robust design and maintenance solutions a systematic approach is required. Such an approach is being dealt with in the next section.
3. A Systematic Trial-And-Error Approach

3.1 The approach taken

The risk of failures, the maintenance needs and the failure restore time can be influenced by many factors. For instance: the quality and maintainability of materials and components, the track alignment, the planning of renewals, the monitoring equipment installed, the number of maintenance depots and failure response teams, the availability of spare parts, the site accessibility and the quality of the rolling stock. Some of these factors are hard to control by the Infrastructure Provider, whereas others can be controlled very well. Therefore the question is: how much improvement of the availability can be realised by a specific design or M&R measure and what does it cost? Maybe – for instance - an investment of 1 Euro extra in signalling equipment realises a higher availability than 1 Euro extra in track components.

In order to systematically analyse the availability and the costs of ownership for every part of the rail system (Track, Signalling and Power Supply Systems), insight is needed into:

- failure frequencies (MTBF: Mean Time Between Failures);
- failure restore times (MTTR: Mean Time to Restore Services);
- costs of engineering, procurement and construction (EPC costs); and
- costs of maintenance and renewals (M&R costs).

At the start of the Tender Phase the engineers of the consortium were working in separate teams on the design issues related to their own technical discipline (Track - T, Signalling - S, and Power Supply - PS). A separate team covered maintenance and renewal issues. Three steps should be taken in order to include the penalties in the Performance Fee:

1. **Composing so-called Incident Matrices per Technical Discipline (T, S, PS):** these matrices contain all the failures, their failure modes and their consequences. In fact, this is a Failure Mode Effect Analysis with quantified estimates on frequencies and restore times per failure type.

2. **Forecasting Availability Levels with (statistical) Confidence Intervals:** in a next step a Monte Carlo simulation should deliver the predictions on the availability levels based on the Incident Matrices (with the help of a first version of the Performance Simulation Model, which contains the timetable, the track features and the rolling stock performance).

3. **Calculating the Performance Fee required:** in the final step the estimation of the performance penalties is combined with the design and M&R costs in the Financial Model. Other financial variables, such as tax and interest, are included in order to estimate the cash flows.

However, in order to be able to analyse and optimise the system availability resulting from the plans of the different engineering teams in an early phase running through this cycle one analysis cycle was not sufficient. An approach, provocatively called the Optimisation Process, consisted of regular plenary meetings of the design teams and the M&R team [5]. In these meetings the different design and M&R choices should be considered in order to realise an optimal (i.e. high) availability level. After every discussion the teams had the obligation to include the measures agreed upon in their Incident Matrices. Before every meeting the impact on the availability and costs of ownership (including penalties) was estimated. In order to show the outcomes and the contribution of each technical discipline to non-availability and costs, a Life Cycle Cost decision support system (DSS) was used [6]. This DSS was able to instantly produce availability estimates under a number of assumptions: it was based on the average failure frequencies (MTBFs) and a flat spread of delay minutes within a year. Therefore an underestimation of the delay minutes was likely, but it was sufficient for analysing the contribution of different failure types to the penalties and the result of different design and M&R strategies.
Due to the fact that the number of failure types was rather large, these were categorised on their consequences (duration of the speed restrictions) and the technical subsystem they are part of. Figure 3 shows some examples of these **Incident Profiles** for illustration purposes. T1 could for instance contain a number of failure types, such as a rail break.

![Example Incident Profiles](image)

**Figure 3: Examples of Incident Profiles (ONLY ILLUSTRATIVE)**

The main input for the DSS are the **traffic forecasts** from the Reference Timetable, the penalty regime (PPMA), the **interest rate**, the design, maintenance and renewal costs per year, the **average delay and number of Not Available Days (NADs) per Incident Profile** (calculated by the Performance Simulation Model), and the **failure frequency per Incident Profile**.

On a computer screen the effects of a change in frequency of incidents can be shown directly. On this screen the savings or increases in penalties can be shown (figure 4), which are the margins available for investment. Basically, there are two possibilities for improving the availability:

- Reduce the number of failures in a category;
- Transfer failure types to a category with less speed reduction (due to better incident management).

![Incidents Direct Fastening: availability costs include penalties and NA](image)

**Figure 4: Amount of penalty for each Incident Profile (ONLY ILLUSTRATIVE)**

This should be traded off against the total costs for design, maintenance and renewal, which is shown in a separate computer screen (figure 5). With the help of the on-line use of the DSS and the regular meetings of representatives from the different technical disciplines (T, S, and PS) trial-and-error of all kinds of design and M&R measures should systematically improve the total costs.
### Life cycle cost breakdown with Direct Fastening

<table>
<thead>
<tr>
<th></th>
<th>Power Supply (MIO ?)</th>
<th>CCC / Signalling (MIO ?)</th>
<th>Track / CW (MIO ?)</th>
<th>E&amp;M (MIO ?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discounted costs of NADs (MIO ?)</td>
<td>22.57</td>
<td>46.38</td>
<td>46.07</td>
<td></td>
</tr>
<tr>
<td>Discounted penalty costs (MIO ?)</td>
<td>16.89</td>
<td>30.69</td>
<td>22.65</td>
<td></td>
</tr>
<tr>
<td>Discounted possession costs (MIO ?)</td>
<td>0.00</td>
<td>0.00</td>
<td>16.20</td>
<td>0.00</td>
</tr>
<tr>
<td>Discounted renewal costs (MIO ?)</td>
<td>40.00</td>
<td>79.00</td>
<td>271.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Discounted maintenance costs (MIO ?)</td>
<td>74.76</td>
<td>107.19</td>
<td>137.76</td>
<td>128.81</td>
</tr>
<tr>
<td>Construction costs (MIO ?)</td>
<td>94.00</td>
<td>190.80</td>
<td>271.25</td>
<td>153.85</td>
</tr>
<tr>
<td>% penalty excl. NAD</td>
<td>23.71%</td>
<td>43.08%</td>
<td>33.21%</td>
<td></td>
</tr>
</tbody>
</table>

![Histogram showing breakdown of costs](image)

**Figure 5: Ownership cost and non-availability per discipline (ONLY ILLUSTRATIVE)**

In figure 6 the whole calculation process is summarised: the technical designs chosen and the M&R strategy result in direct costs (construction costs, M&R costs and some costs for track possessions). Further, they will result in a certain failure occurrence. The engineers quantify the expected failures in the so-called Incident Matrices. In the matrices the failures are categorised according to different Incident Profiles. For each Incident Profile the average delay is calculated with a preliminary Performance Simulation Model. These average delays, together with the failure frequencies and the distribution over the years is further used in the Life Cycle Cost DSS. This DSS contains the PPMA and the data on construction and M&R costs, and calculates the total costs of ownership (life cycle costs during the first 25 years). Although it may look somewhat complicated, the process can be performed quickly and is repeated after every improvement in the design, the M&R strategy and/or the Incident Matrices.

### 3.2 Observations on the process and the results

The Optimisation Workshops started in December 2000; during the first two months it proved that the Incident Matrices were of a low quality. Failure types were missing, failure frequencies, restore times and speed restrictions were modified after every session, and all this resulted in a high ‘volatility’ of the penalty costs caused by the different technical systems. However, it led to the recognition of the importance of the Performance Payment Regime among the participants. Furthermore, it helped in starting discussions between the construction and maintenance staff. During the last two months more experts were involved, in order to develop robust Incident Matrices and to obtain more benchmark data from European high-speed lines. All participants considered the approach successful in realising realistic availability forecasts, but it proved to be very hard to explicitly show the relations between investments and ‘system availability’ [7].
Nevertheless, the robustness of a number of design changes was tested with the DSS and implemented in the design and M&R strategy, such as:

- faster response times to incidents and improved scenarios for incident handling (such as allowing a bidirectional traffic regime in case of failing switches on the connection to the city of Breda);
- implementing redundant systems if possible (Signalling);
- installing monitoring devices for the switches;
- developing quicker repair methods for rail breaks (e.g. for embedded rails); and
- (research proposals on) a reduction of cross-over switches.

4. Conclusions

Although ‘life cycle costs’ and ‘availability’ are well-known terms in the railway sector, a professional approach to optimising costs and availability of the rail system as a whole is far from available. A systematic analysis of the costs and availability impacts on the long run is still an exception in a design phase. During the Tender Phase of the High Speed Line South it proved that most engineers were not used to do this type of analysis. As one risk analyst involved put it: "I regularly participate in workshops to identify and prioritise risks, but the Optimisation Workshops were something new for me". Also the changes made to the Performance Payment Regime during the Tender Phase by the Dutch State showed (in the opinion of the authors) that it was a learning process for the buyer of ‘Availability’ as well.

The major advantage of the Optimisation Workshops, and the use of the decision-support system, was the clear identification of the contribution of different system components to the ‘Non-Availability’. The graphical display proved to trigger the discussion among the engineers in the consortium and this helped to overcome the many obstacles, such as the poor communication between construction and maintenance staff and the reliability of the maintenance data.
A result of the Optimisation Process was the reduction of the failure frequency and restore times estimated in the Incident Matrices for track, power supply and signalling, being developed based on the Failure Mode Effects Analysis approach. Analysing the availability of a railway system is not easy as it concerns stochastical processes. To explicitly show the relations between investments and an improvement of the availability proved to be the toughest job.

Some doubts on the reliability of data remained however due to the fact that it concerned the use of innovative technology (e.g. slab track and the ECTS/ERTMS train control system), and partly because it was difficult to obtain benchmark data during the tender. It is expected that an application in the maintenance organisation of a conventional rail network should be more successful in this perspective, and that this could lead to important new insights as well.

ACKNOWLEDGEMENT

Herewith we express our gratitude to the management of the FASTRAIL Consortium, and the companies Strukton Railinfra, Alstom and Volker Stevin Rail & Traffic for the research possibilities provided as well as the opportunity to present the research work.

BIBLIOGRAPHY

3. Ministry of Transport, HSL South Project Organisation Invitation to Tender. HSL South Project Organisation, Utrecht, 1999