Renewal decisions from a Life-cycle Cost (LCC) Perspective in Railway Infrastructure:

An integrative approach using separate LCC models for rail and ballast components

António Ramos Andrade

M.Sc Thesis in:

Civil Engineering

President:
Supervisor: Prof. Paulo Manuel da Fonseca Teixeira
Member:

September 2008
E assim nas calhas da roda  
Gira, a entreter a razão,  
Esse comboio de corda  
Que se chama coração.

(Interpretation by Roy Campbell)

Thus to beguile and entertain  
The reason, does he start,  
Upon its rails, the clockwork train  
That's also called the heart.

(Autopsicografia by Fernando Pessoa, 1930)
Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Professor Paulo Teixeira, for believing in my capacities, introducing me to the complex world of Railways and encouraging me throughout the research. Moreover, I would like to thank him for his stimulating suggestions and his priceless time.

I would like to thank all my friends who continuously supported me and had to listen to boring discussions and arguments about my work, especially Catarina, Mariana, Luiz and Rita. A special thanks to Pedro for teaching me that evaluation is not based on your potential, but on your results.

I am also very grateful to my colleagues for their support and enthusiasm, especially Mariana, Ricardo, Edgar, José Pedro, Joana, Miguel, Silvia, André and Nuno.

Moreover, I would like to thank Rosario for her attentive reading and prompt suggestions and contributions on improving my English.

I would also like to extend my gratitude to my beloved grandmother and grandfather, for their interesting conversations on History, Art, Science and Math, but most of all for transmitting me values and standards that I am proud to respect.

Finally, I would like to express my unconditional love towards my parents and my brother João, for investing in my education and motivating me into the world of knowledge, and also for listening to my arguments and contributing to their improvement.

António Andrade
September, 2008
Lisbon, Portugal.
Abstract

The aim of the present research is to develop an integrative approach to support renewal decisions of Railway infrastructure components, namely ballast and the rail. Having in mind that the Railway infrastructure sector has faced drastic transformations due to liberalization of the sector, more conscious and transparent decision-making processes are needed, where Life-cycle Cost (LCC) analysis may be crucial to overcome new challenges. In the present work, rail and ballast LCC models are developed separately. For the rail LCC model, an existing model is used, whereas for the ballast LCC model, a new model is proposed based on a track geometry degradation model. Regarding the assessment of the uncertainty associated with LCC estimations, an innovative approach is put forward, where a step back is taken to the typical approaches, inserting uncertainty earlier in degradation model parameters, rather than RAM (Reliability, Availability and Maintainability) parameters. Nevertheless, in LCC ballast model, potential uncertainty related to unavailability costs proves to be more relevant on renewal decisions, rather than uncertainty related with tamping costs assessed by Monte Carlo simulation. A straight-forward approach that integrates both LCC models based on the construction of hypothetical scenarios where distinct renewal tonnages are used to achieve an optimal renewal strategy is developed.

Keywords: LCC, Railway infrastructure, maintenance, renewal decisions, uncertainty.
Resumo

O objectivo deste trabalho é desenvolver uma abordagem integrada para apoiar as decisões de renovação das componentes da infra-estrutura ferroviária, nomeadamente o balastro e o carril. No contexto das profundas transformações que o sector ferroviário atravessou, devidas à liberalização do sector, são necessários processos de tomada de decisão mais conscientes e transparentes, nos quais a análise de custos por ciclo de vida (CCV) pode ser crucial para enfrentar estes novos desafios. Neste trabalho, os modelos de CCV do carril e do balastro são desenvolvidos separadamente. Para o modelo de CCV do carril, é utilizado um modelo existente, já para o modelo de CCV do balastro, é proposto um novo modelo baseado num modelo de degradação geométrica da via. Considerando a avaliação da incerteza associada às estimativas de CCV, é proposta uma abordagem inovadora em que, face às abordagens usuais de inserir a incerteza nos parâmetros de RAM (Fiabilidade, Disponibilidade e Manutibilidade), se dá um passo atrás, inserindo previamente a incerteza nos parâmetros dos modelos de degradação. Contudo, no modelo de CCV do balastro, a potencial incerteza relacionada com os custos de indisponibilidade sugere ser mais relevante nas decisões de renovação, do que a incerteza relacionada com os custos do batimento avaliados pela simulação de Monte Carlo. Uma forma expedita para integrar ambos os modelos de CCV, baseada na construção de cenários hipotéticos com diversa tonelagem de renovação, é proposta a fim de atingir uma estratégia óptima de renovação.

Palavras-chave: CCV, infra-estrutura ferroviária, manutenção, decisões de renovação, incerteza.
List of figures

Figure 2.1 - Typical relationship between dependability and LCC for the operation and maintenance phases, IEC 60300-3-3 (2004).

Figure 2.2 - Distribution of construction cost of a new railway line to the various components of the railway system, Profillidis (2006).

Figure 2.3 - General factors influencing the costs and performance of the rail infrastructure, Zoeteman (2004).

Figure 2.4 - Structure of the Life-cycle Cost DSS – calculation steps and data needed, Zoeteman (2004).

Figure 3.1 - Ballasted track components, Profillidis (2006).

Figure 3.2 - Analysis principle applied to an hypothetical track geometry deterioration, Esveld (2001).

Figure 3.3 - Theoretical failure probability over lifetime, Zoetmann (2004).

Figure 3.4 - Evolution of longitudinal level of an FS test section, Esveld (2001).

Figure 3.5 - Types of failure progression, Esveld (2001).

Figure 3.6 - Schematic survey of maintenance and renewal process, Esveld (2001).

Figure 4.1 - Life-cycle Cost of rail per MGT for a 1-km section.

Figure 4.2 - Sensibility of the present value of Life-cycle Cost of rail per MGT for a 1-km section to different discount rates (1%, 2%, 4% and 6%), for an annual accumulated tonnage of 12 MGT.

Figure 4.3 - Sensibility of the present value of Life-cycle Cost of rail per MGT for a 1-km section to different annual accumulated tonnages (12 MGT, 24 MGT, 36 MGT and 48 MGT), for a discount rate of 4%.

Figure 4.4 - Example of track level development on 10 successive sections of 200 m on FS, Esveld (2001).

Figure 4.5 - Total Life-cycle Cost of Tamping operations per MGT of a 100 km plain track section ($\text{TLCC}_{\text{tamping}}$).

Figure 4.6 - Total Life-cycle Cost of ballast per MGT of a 100-km plain track section ($\text{TLCC}_{\text{ballast}}$).

Figure 4.7 - Total Life-cycle Cost per MGT of rail and ballast components for a 100-km plain track section.
List of tables

Table 1.1 - Measures for long-distance travel (passenger and freight) adapted from EP (2008).

Table 2.1 - Construction costs (values for 2006) of high-speed tracks constructed during recent years, Profiliidis (2006).

Table 3.1 – Track maintenance activities (Patra, 2007)

Table 4.1 - Parameter values for rail defects used in the example analyzed, Zhao et al (2006).

Table 4.2 - Coefficients of the model used in the example analyzed, adapted from Zhao et al (2006).

Table 4.3 - Example of tamping costs for different lines, CENIT (2008).

Table 4.4 - Renewal scenarios for ballast and rail with respective total life-cycle cost per MGT (ballast + rail).
Definitions and Acronyms

ABC – Activity Based Costing
ATW – Alumino-Thermic Welds
CWR – Continuous Welded Rail
DSS – Decision Support System
EU – European Union
EP – European Parliament
FMEA – Failure Modes and Effects Analysis
GHG – Greenhouse Gases
HD – Horizontal Defect
ICT – Information and Communications Technology
ICE – InterCityExpress
IEC – International Electrotechnical Commission
IM – Infrastructure Manager
ISO – International Organization for Standardization
LCA – Life-Cycle Assessment
LCC – Life-Cycle Costing / Life-Cycle Cost
LD – Longitudinal Defect
M&R – Maintenance and Renewal
MAD – Mean Administrative Delay
MGT – Million Gross Tons
MLD – Mean Logistic Delay
MRT – Mean Repair Time
MTTF – Mean Time To Failure
NDT – Non-Destructive Testing
P-F – Potential failure to Functional failure
PPMA – Performance Payment Mathematical Algorithm
PSI – Present Serviceability Index
RAM – Reliability, Availability and Maintainability
RAMS – Reliability, Availability, Maintainability and Safety
RCF – Rolling Contact Fatigue
SD – Standard Deviation
SJT – Scheduled Journey Time
STRAIT – Straightening of Rail welds by Automated Iteration Technique
S&C – Switches and Crossings
TD – Transverse Defect
TGV – Train à Grande Vitesse (High Speed Train)
UIC – Union Internationale des Chemins (International Union of Railways)
USA – United States of America
Symbols

\( \alpha \) - Shape parameter of Weibull distribution

\( \alpha, \beta \) - Guérin material parameters

\( a_0 \) - Settlement increase rate in mm/decade

\( a_1 \) - Mean Settlement for a traffic load \( T_r \) in mm

\( \beta_j \) - Probability that a defect \( j \) can be detected through inspection

\( c_0 \) - Rate of increase of standard deviation of longitudinal defects in mm/decade; Rate of deterioration of standard deviation of longitudinal defects in mm/MGT

\( c_1 \) - Standard Deviation of longitudinal defects for a traffic load \( T_r \) in mm; Initial quality measured after renewal or tamping operations in mm

\( c_i \) - Cost of Inspection for one cycle

\( c_R \) - Renewal Cost per km

\( c_d \) - Cost of repairing a defect that has been detected through inspection

\( c_f \) - Cost of a functional failure

\( c_g \) - Cost of a grinding operation

\( c_x \) - Cost of a derailment

\( c_{\text{geom.insp}} \) - Cost of geometric inspection per inspection and maintenance section

\( c_{\text{renewal}} \) - Cost of ballast renewal per maintenance section

\( c_{\text{tamping}} \) - Cost of tamping for a 200-meter section

\( C \) - Constant controlling the rate of growth of deformation

\( C(T) \) - Rail Life-Cycle Cost per MGT for a rail life-cycle \( T \)

\( C_0(T) \) - Present Value of \( C(T) \)

\( CC \) - Construction Costs

\( d \) - Maximum Elastic Deflection

\( d_0 \) - Horizontal defect increase rate in mm/decade

\( d_1 \) - Mean horizontal defect for a traffic load \( T_r \) in mm

\( D \) - Minimum Wheel Diameter of the freight train in m

\( DC \) - Delay Costs

\( \varepsilon \) - Total Permanent Strain; Probability of an accident being caused by a rail failure
\( \varepsilon_1 \) - Permanent Strain after the first load cycle

\( \eta \) - Scale parameter of Weibull distribution

\( F_i \) - Axle Load \( i \)

\( F_{eq} \) - Equivalent Axle Load

\( \gamma(q) \) - Probability that a surface-initiated defect cannot be removed by grinding

\( g_0 \) - Rate of increase of standard deviation of track twist in mm/decade

\( g_1 \) - Standard Deviation of track twist for a traffic load \( T_r \) in mm

\( G_j(t) \) - Cumulative Distribution Function of the P-F interval for the \( j \)th type of defect

\( h \) - Total Number of rail defects

\( HD \) - Horizontal Defect

\( JT_{tr} \) - Journey Time for a train type \( tr \)

\( K_1, K_2 \) - Shenton model parameters

\( k_{fr} \) - Coefficient taking into account effects of both load and wear provoked by freight bogies

\( k_1 \) - Coefficient taking into account wear resulting from traction vehicles

\( \lambda_a(t) \) - Hazard rate of a defect occurring within each ATW at time \( t \)

\( \lambda_{RCF}(t) \) - Hazard rate of an RCF when no grinding is being undertaken

\( LCC \) - Life-Cycle Cost

\( LCC_{geom\_insp} \) - Life-Cycle Cost of geometric inspection per maintenance section

\( LCC_{renewal} \) - Life-Cycle Cost of ballast renewal per maintenance section

\( LCC_{tamping} \) - Life-Cycle Cost for tamping operations

\( LCC_{tamping\_k} \) - Life-Cycle Cost for tamping operations for the \( k \)th maintenance section

\( LD \) - Longitudinal Defect

\( L_{inf} \) - Alarm values of track defects

\( L_{sup} \) - Upper values of track defects

\( m \) - Number of inspections in the time period \( (0,T) \)

\( m_a(T) \) - Mean settlement for an accumulated traffic load (tonnage) \( T \)

\( m_{HD}(T) \) - Mean horizontal defect for an accumulated traffic load (tonnage) \( T \)

\( MC \) - Spot Maintenance Costs

\( \nu_a(t) \) - Hazard rate of an ATW defect within a section of rail at time \( t \)
$\nu_b(t)$ - Hazard rate of a type B defect within a section of rail at time $t$

$\nu_f(t)$ - Hazard rate of a defect appearing in a section of rail at time $t$

$\nu_{RCF}(t)$ - Hazard rate of a rolling contact fatigue defect within a section of rail at time $t$ when no grinding is being undertaken

$n_0$ - Number of ATWs in a section of rail at time $t = 0$

$n_{insp}$ - Number of inspections per year

$N$ - Load Cycle

$N_d(t)$ - Expected Number of Defects that have been detected by inspections within $(0, t)$ in a section of rail

$N_f(t)$ - Expected Number of Failures at time $t$ in a section of rail

$OC$ - Organizational Costs

$OH$ - Operational Hours per day

$P_c$ - Maximum Axle Load with Wheels of Diameter $D$ in tons

$r$ - Discount Rate

$r_0$ - Increase Rate of $c_0$

$r_1$ - Increase Rate of $c_1$

$RC$ - Periodic Maintenance and Renewal Costs

$\sigma$ - Standard Deviation of longitudinal defects

$\sigma_{lim}$ - Limit Standard Deviation of longitudinal defects

$s_f$ - Interval of Inspection

$s_g$ - Interval of Grinding

$sd_{LD}(T)$ - Standard Deviation of longitudinal defects for an accumulated traffic load (tonnage) $T$

$sd_{lim}$ - Limit value of Standard Deviation of longitudinal defects

$sd_{TD}(T)$ - Standard Deviation of transverse defects for an accumulated traffic load (tonnage) $T$

$sd_{twist}(T)$ - Standard Deviation of track twist for an accumulated traffic load (tonnage) $T$

$S_{fr}$ - Coefficient related to train Speed for freight traffic

$S_{p}$ - Coefficient related to train Speed for passenger traffic

$SD$ - Standard Deviation

$SJT$ - Schedule Journey Time

$\tau$ - Settlement
\( t \) - Amount of traffic passing over rails in million gross tonnes (used to refer to time)

\( T \) - Rail life-cycle; Accumulated Tonnage between tamping operations

\( T_{\text{ballast}} \) - Accumulated tonnage at first ballast renewal

\( T_{\text{ballast}} \) - Accumulated tonnage at second ballast renewal

\( T_{\text{disruptive}} \) - Disruptive tonnage (accumulated tonnage limit above which tamping operations become disruptive.

\( T_f \) - Fictive Tonnage or Notional

\( T_{fr} \) - Mean Daily Freight Tonnage Hauled in gross tons

\( T_g \) - Real load for daily freight traffic

\( T_i \) - Accumulated Tonnage for a maintenance section when the \( i \)th tamping operation is performed

\( T_{i\text{cum}} \) - Accumulated tonnage till the \( i \)th tamping operation

\( T_{\text{lim}} \) - Tonnage Interval between two successive maintenance sessions (tamping operations)

\( T_n \) - Cumulative Tonnage at year \( n \)

\( T_p \) - Mean Daily Passenger Tonnage Hauled in gross tons; Real load for daily passenger traffic

\( T_r \) - Traffic load of 2 MGT

\( T_{\text{rail}} \) - Accumulated tonnage at rail renewal

\( T_{\text{renewal}} \) - Accumulated tonnage at ballast renewal

\( T_{tf} \) - Mean Daily Tonnage of Tractive Unit used in freight traffic in gross tons

\( T_{th} \) - Theoretical Traffic Load in gross tons

\( T_{tp} \) - Mean Daily Tonnage of Tractive Units used in passenger traffic in gross tons

\( T_{\text{year}} \) - Annual Accumulated Tonnage in MGT

\( TD \) - Transverse Defect

\( \text{TLCC}_{\text{ballast}} \) - Total Life-Cycle Cost of ballast component per MGT for the 100-km plain track

\( \text{TLCC}_{\text{ballast} + \text{rail}} \) - Total Life-Cycle Cost of ballast and rail components per MGT for the 100-km plain track

\( \text{TLCC}_{\text{geom},\text{insp}} \) - Total Life-Cycle Cost of geometric inspection per MGT for the 100-km plain track

\( \text{TLCC}_{\text{rail}} \) - Total Life-Cycle Cost of rail component per MGT for the 100-km plain track

\( \text{TLCC}_{\text{renewal}} \) - Total Life-Cycle Cost of ballast renewal per MGT for the 100-km plain track
\( TLCC_{tamping} \) - Total Life-Cycle Cost of tamping operations per MGT for the 100-km plain track

\( TLCC_{unavailability} \) - Total Life-Cycle Cost of unavailability per MGT for the 100-km plain track

\( TPH_{tr} \) - Number of Trains per hour per track for a train type \( tr \)

\( TRACKS \) - Total Number of Tracks

\( u_0 \) - Rate of increase of standard deviation of transverse defects in mm/decade

\( u_1 \) - Standard Deviation of transverse defects for a traffic load \( T_r \) in mm

\( V_{max} \) - Maximum Permissible Speed in km/h

\( V \) - Track Speed in km/h

\( y \) - Vertical settlement

\( z(T,x) \) - Average of track elevations between the internal and external rails, corresponding to a traffic load \( T \) at a kilometric position \( x \)

\( z_{int} \) - Track elevation of the internal rail

\( z_{ext} \) - Track elevation of the external rail

\( z_{th}(T,x) \) - Theoretical value or initial value of \( z(T,x) \)
# Index

Acknowledgements.............................................................................................................. ii
Abstract................................................................................................................................. iii
Resumo .................................................................................................................................. iv
List of figures............................................................................................................................ v
List of tables............................................................................................................................. vi
Definitions and Acronyms...................................................................................................... vii
Symbols................................................................................................................................. ix

## Chapter 1

**Introduction**....................................................................................................................... 1

1.1 - Vertical separation: a challenge of Railway infrastructure ........................................ 1

1.2 - European policy framework ....................................................................................... 2

1.3 - Life-cycle Costing (LCC): an instrument to a more efficient railway sector .......... 4

## Chapter 2

**Life-cycle Costing (LCC)**.................................................................................................. 6

2.1 - LCC: a support for decision-making and cost structure analysis ............................. 6

2.2 - International Standard IEC 60300-3-3: Application Guide ........................................ 7

2.3 - Main obstacles to LCC implementation in Railways ................................................. 12

2.4 - LCC in other transport infrastructure: roadway infrastructure .............................. 13

2.5 - LCC in Railways: the particular Zoeteman model .................................................... 16

## Chapter 3

**Degradation and Maintenance of Railway Infrastructure**.............................................. 27

3.1 - Railway Infrastructure components .......................................................................... 27

3.2 - Degradation models .................................................................................................... 29

3.2.1 - Degradation of track geometry .............................................................................. 32

3.2.2 - Degradation of rails ............................................................................................... 40

3.3 - Maintenance and Renewal activities ......................................................................... 41

3.3.1 - Inspection ............................................................................................................... 42

3.3.2 - Maintenance ........................................................................................................... 45
Chapter 4

Practical example: Rail and Ballast renewals from a life-cycle cost perspective .......... 49
  4.1 - Rail LCC model .......................................................................................................... 49
  4.2 - Ballast LCC model ....................................................................................................... 55
  4.3 - Integrating rail and ballast LCC models ........................................................................ 64

Chapter 5

Conclusions and Future Research ................................................................................. 68
  5.1 - Conclusions .................................................................................................................. 68
  5.2 - Limitations .................................................................................................................... 69
  5.3 - Future Research ............................................................................................................. 70

References ............................................................................................................................. 72

Appendix 1 ............................................................................................................................. 75
Chapter 1

Introduction

This introductory chapter intends to give an inside look of the Railway sector, starting with a brief historical introduction, highlighting the necessity to revitalize Railways through vertical separation. Moreover, it identifies new challenges of European Railways towards a more sustainable European transport market, where Life-cycle Costing (LCC) analysis is an effective instrument towards a more efficient railway sector, supporting decision-making processes.

1.1 - Vertical separation: a challenge of Railway infrastructure

In a modernized and globalised world, the mobility of people and goods has been increasing, even above economic growth. While millions of people all around the world commute in urban transport networks, millions of products move throughout complex logistic chains. In both systems, Railway as a mode of transport plays an important role.

In the last two centuries, the Railway has changed radically, not only in terms of design, manufacturing, etc; but also in terms of its image in society. In 1825, when the first connection opened, linking Stockton to Darlington, Railway was seen as a symbol of progress, explored by private companies, as it was a profitable sector at the time. However, in the 20th century, the railway sector faced a serious financial crisis, culminating with America's Great Depression. It increasingly became regarded as a societal burden, losing its attractiveness throughout the years, mainly because of the introduction of more competitive modes like the car and the plane.

After the Second World War, nationalized railway companies represented a valuable asset to society, operating in a monopoly industry with non-profitable purposes similar to other infrastructures and utilities owned by the state, such as telecommunications, mail delivery, water and energy supply. However, in the 1960s and 1970s, the Railway rapidly lost its market share to road and air, and annual losses increased in such a dramatic way that revitalizing the railway sector worldwide became necessary. Adopted solutions mostly consisted of privatization of the entire railway system or parts of it. The United States and Japan took the lead, and soon the European Union followed this trend.

The relation between the different stakeholders involved in the Railway sector changed drastically with the introduction of the EU directive 91/440. Its aim was to facilitate the adoption of Community railways to the needs of the Single Market and to increase efficiency. This
required a separation between the provision of transport services and the operation of infrastructure in the railway systems, while improving the financial structure of infrastructure managers by introducing separate accounts between both activities and the state. It is important to refer that though the separation of accounts was compulsory, organizational and institutional separation was optional. This fact may have triggered regulation and competitive problems that European markets are still facing today. Note that this directive also highlights the importance for member states to retain general responsibility for the development of the appropriate railway infrastructure, while excluding from its scope all the railway systems with solely urban, suburban and regional services.

With separate accounts implemented, the infrastructure manager (IM) would charge a fee to railway operators for using that infrastructure. The rules for determining this fee should be set by member states. The calculation of this fee is still controversial, demanding more knowledge on life-cycle cost of the railway infrastructure. However, it is considered reasonable that this fee should take into account mileage, the composition of the train, specific requirements as speed, axle load, degree or even period of utilization.

In fact, the European railway policy for vertical separation of this sector can be pointed as a good example of “framing integration”, as suggested by Knill and Lehmkul (1999). That basically means that the regulatory policy does not prescribe any concrete institutional requirements, nor does it modify the institutional context for strategic interaction, but affects domestic arrangements indirectly by altering the beliefs and expectations of domestic actors. Therefore, railway industry has a whole may benefit from the opportunities of moving from a ‘stock economy’ to a ‘flow economy’ through physical, operational and commercial integration, towards effective rail interoperability with other transport modes.

1.2 - European policy framework

In a continuous quest for a more cost-competitive system, European railways have walked a long path of standardisation and knowledge sharing in order to invert the current railway position in the European transport system. In fact, asymmetric regulation in the transport sector is needed since the so-called internalization of external costs in the roadway transport is not implemented, which makes consumers pay much less than its real cost (especially when considering environmental impacts of CO₂ emissions). Moreover, to become more competitive than road transport, interoperability between national systems must be achieved, improving the seamless movement of trains across Europe, while reducing delays at border crossings and extra costs associated.

The 2001 White Paper emphasized the importance of safeguarding efficient mobility for people and goods as the central element of a competitive EU internal market. This key
document of the EU’s strategy on transport policy proposed a list of measures to break the link between economic and traffic growth, to promote modal shift and combat the unequal growth of the various modes of transport. In fact, the document sees rail transport as the key to achieve modal rebalance and mitigate the dependency on road transport, particularly in the case of passengers. In the case of goods, the railway has been losing its market share from 21% in 1970 to 8% in 1998 which, compared to the 40% market share that rail haulage still represents in the USA, reveals that there is still in Europe great potential in railway transport to explore. In order to increase its competitiveness, the EU has tried to create a network of railway lines dedicated exclusively to the service of goods, while partly transferring passenger traffic to ‘new’ high performance high-speed rail networks. Nevertheless, interoperability between national railways and other modes of transport seems to be more important even than segregated traffic. For example, one major obstacle to effective operation usually discussed is the difference of gauge between the Iberian infrastructure (1,668 m) and the international one (1,435 m). The new high-speed line will contribute to the alignment of the gauge, improving the links between Spain and Portugal and the rest of the European network. In addition to interoperability and increasing competitiveness of intermodal transport, other measures such as creating a genuine market in rail transport, making optimum use of infrastructure while guaranteeing rail safety standards and assuring the necessary modernisation of services, are also important strategic goals.

More consciousness of the importance of sustainability in the European transport policy is also evident, as energy and environmental aspects of the transport policy became crucial for the challenge of climate change. In fact, according to the last European Environment Agency report (2008), transport accounts for 21% (878Mt CO₂ equivalent) of all EU-15 greenhouse gas (GHG) emissions. Road transport contributes with the biggest share, accounting for 19% of all EU-15 GHG emissions, while Railways only represent 0.1% of the total GHG emissions (6 Mt CO₂ equivalent). Although the focus to face this challenge is the road transport, through technological improvements in vehicle and fuels (increasing energy efficiency), or even new methods of charging and taxation (internalisation of external costs of transport), the revitalisation of railways is still seen as an essential element to invert the status quo. The next table sums up some measures proposed to the European railway sector:

<table>
<thead>
<tr>
<th>Measures</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail interoperability</td>
<td>Improve seamless movement of trains across Europe</td>
</tr>
<tr>
<td>Harmonised regulation systems</td>
<td>Provide fair competition for rail operators</td>
</tr>
<tr>
<td>Rail efficiency</td>
<td>Increase technical unit efficiency of rail travel</td>
</tr>
<tr>
<td>Rail passenger service quality</td>
<td>Increase quality (rolling stock, ICT)</td>
</tr>
<tr>
<td>Intermodal facility for passengers and freight</td>
<td>Develop service integration (operating facilities)</td>
</tr>
<tr>
<td>Rail capacity</td>
<td>Improve rail capacity in key corridors and rail bottlenecks</td>
</tr>
</tbody>
</table>

Table 1.1 - Measures for long-distance travel (passenger and freight) adapted from EP (2008).
1.3 - Life-cycle Costing (LCC): an instrument to a more efficient railway sector

Due to the liberalization of the railway sector and the consequential increase of government regulations and restrictions, assessing infrastructure performance has become a necessity for infrastructure managers (IMs), regulator entities and even operators (as costumers of the infrastructure). Restrictions on RAMS operation level (Reliability, Availability, Maintainability and Safety) demand systematic management, in order to improve the quality and transparency of decision-making, but also to increase rail quality in terms of service. In fact, RAMS management lacks a life-cycle cost perspective, so that decisions on maintenance and renewal of infra components consider long run costs. Therefore, a Decision Support System based on Life-cycle Costing (LCC) Analysis should be developed, balancing short and long term costs with performance (revenues), opposing alternative design and maintenance strategies in order to minimize (long run) total cost, while testing the robustness of those decisions under different operational conditions. Having said that, this work will try to understand better the details on Life-cycle Costing analysis of the Railway Infrastructure, focusing on maintenance operations and degradation processes, while evaluating the risks associated, towards a symbiotic model between RAMS and LCC.

*Railway design and maintenance from a life-cycle cost perspective*¹ is believed to have both societal and scientific relevance, which is beneficial for taxpayers and travellers, as resultant decisions on investments/operation/maintenance are more transparent and cost-effective. Transparency is decisive to build up confidence in relationships between the different actors of the Railway sector (IMs, operators, regulatory entity...).

As Zoeteman formulated as his research objective in his PhD thesis (2004):

> ‘to develop and test an approach that can immediately assist the rail infrastructure sector in identifying and adopting design and maintenance strategies, which take ‘life-cycle’ impacts into account (within the required levels of technical and functional performance)’

Decisions on design options and maintenance strategies must be supported by an economic analysis, assessing costs and performance associated. Therefore, the railway sector, as a capital-intensive industry, needs a life-cycle approach. However, models developed so far are usually hostage of expert opinion, as their inputs are based on expert estimations (Espling, 2007). This fact may not be a disadvantage in terms of quality of LCC estimations considering past experience of the decision-makers, though it may result in slower processes very often, as experts are outsiders of the IM’s organisation. Moreover, expert judgment is mostly applied when is difficult to obtain the required data, or where great precision in prediction may not be

---

¹ Zoeteman PhD thesis title.
required, or even in situations considered highly uncertain, though further research on linking all various theories for handling different types of uncertainties from expert judgment versus models’ uncertainties is needed (Booker, Anderson and Meyer, 2000).

Therefore, LCC as an economic instrument to support decision-making will definitely contribute to a more cost-efficient Railway sector, adapted to the new challenges of a competitive and sustainable European Transport sector.
Chapter 2

Life-cycle Costing (LCC)

This chapter discusses in detail LCC analysis as a support for decision-making and cost structure analysis, based on an International Standard: an application guide to LCC analysis for a product in general. Furthermore, main obstacles to LCC implementation are discussed and a small description of LCC analysis in Roadways is added. Finally, an example of an LCC model developed by Zoeteman for Railway infrastructure is analysed in detail.

2.1 - LCC: a support for decision-making and cost structure analysis

Life-cycle Costing analysis is a process that involves structuring costs, but most of all assessing the costs of a product throughout its different life-cycle phases, so that consumers are more informed, contributing to a more conscious decision-making process. In fact, one important output of a Life-cycle Costing analysis is the identification of cost drivers, meaning the activities that most contribute to the overall life-cycle cost. Apart from identifying cost drivers and representing a useful cost structuring framework, LCC determines cost-effectiveness of alternative investments and business decisions. Therefore, it may represent a valuable input for a multicriteria model to support decision-making.

Other life-cycle analysis may contribute with different approaches. For instance, Life-cycle Assessment (LCA) compares environmental performance of alternative product systems to meet the same end-use function, from a broad and societal perspective, evaluating all processes causally connected to the physical life-cycle of the product (Norris, 2001). Note that this evaluation include the entire pre-usage supply chain, use and the processes supplying use, end-of-life and the processes supplying end-of-life steps. Moreover, no monetary units are addressed in LCA, like in LCC when evaluating costs and benefits directly impacting the decision maker, but instead, the main flows considered (e.g. pollutants, resources, inter-process flows of material) are quantified in physical units (e.g. mass, energy, volume).

Although LCC and LCA are each one designed to provide answers to very different questions, they both may contribute to better decision-making. Therefore, LCC and LCA may be performed separately, and afterwards they may be integrated through a multicriteria decision-making model (Norris, 2001).
2.2 - International Standard IEC 60300-3-3: Application Guide

The International Electrotechnical Commission (IEC) is a worldwide organization whose object is 'to promote international co-operation on all questions concerning standardization in the electrical and electronic fields'. To pursue international uniformity, IEC is responsible for publications, such as International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications and Guides, which compile recommendations for international use. IEC collaborates closely with the well-known International Organization for Standardization (ISO). One of the IEC Publications is the International Standard IEC 60300-3-3, titled Application guide – Life-cycle costing (2004). Many important aspects described in this standard may contribute to build up an LCC model for the Railway Infrastructure. In the next paragraphs, relevant aspects are discussed.

Products today are required to be reliable and safe, but should also be easy to maintain throughout their useful lives. When one makes the decision to purchase a product, one should not be influenced by the product's initial cost alone (acquisition cost), but also by the product's expected operating and maintenance cost over its life (ownership cost) and the disposal cost. Moreover, in order to be cost competitive, producers should design their products to meet requirements and optimize acquisition, ownership and disposal costs. This optimization may even result in adopting different strategies during life-cycle phases (e.g. maintenance strategies).

Therefore, Life-cycle costing (LCC) is defined as the process of economic analysis to assess the total cost of acquisition, ownership and disposal of a product.

This analysis provides important inputs in the decision-making process for different life phases of the product. Thus, by evaluating different design, operating, maintenance and disposal strategies, LCC can be optimized and a cost-effective solution achieved.

LCC is also used to assess the costs associated with a particular activity, for instance, the effects of different maintenance concepts/approaches, to cover a specific part of a product, or to cover only a selected phase or phases of a product's life-cycle. This is usually referred in literature as Activity Based Costing (ABC), but here is integrated in a life-cycle analysis. Note that performing LCC for a particular activity can be very helpful depending on the actors involved. In the Railway infrastructure, its manager may be only interested in assuring the lowest operation and maintenance cost, while obeying to safety and availability restrictions defined by the regulatory entity.

It is important to note that in many industries, design options influence greatly maintenance costs and therefore, life-cycle costing as an integrated analysis to evaluate all
costs throughout the product's useful life is more reliable to make conscious decisions in the life-cycle phases. As will be seen later on, Railway infrastructure is not an exception.

Although the need for formal application of the LCC process to a product will normally depend on contractual requirements, it plays a very useful input to any design of decision-making process. Therefore, LCC should be integrated with the design process, to the feasible extent, to optimize product characteristics and costs. For example, many products in the cold drink industry have design characteristics that take into account transport requirements. These design options are usually evaluated from a life-cycle perspective, which may even include inverse logistics and recycling. In fact, all the different alternatives and options during each life phase of a product should carry an LCC analysis in order to evaluate indirect costs in other phases.

As the product needs to be reliable during life phases, dependability related costs should be included in the analysis. Dependability is defined as the collective term used to describe the product's availability performance, and its related costs should include system recovery cost, preventive maintenance cost and consequential cost.

![Diagram](image.png)

**Figure 2.1** - Typical relationship between dependability and LCC for the operation and maintenance phases, IEC 60300-3-3 (2004).

The figure 2.1 above illustrates the relationship between dependability and LCC during operation and maintenance. Dependability is mainly composed by Reliability, Availability and
Maintainability (RAM). Usually, Safety is included in this group of three aspects, forming RAMS. However, in Railways, Safety is almost always considered an a-priori condition and indirectly guaranteed by good performance levels of the other components of RAMS.

Dependability performance is conditioned by failures, which may be mitigated or repaired through preventive or corrective maintenance with associated costs that should be included in LCC. Furthermore, maintenance activities may demand some investment in logistic support (e.g. maintenance facilities). Finally, consequential costs are those incurred when a product is unavailable, including warranty cost, liability cost, cost due to loss of revenue and costs for providing an alternative service. Further consequential costs can be considered, evaluating adverse impacts on the company's image, reputation and prestige, though they may be more difficult to estimate. Later on in this thesis, it will be seen that in the case of Railways, performance payment regimes contracted between the IM and the regulator entity are mainly the consequential costs considered in LCC.

Life-cycle Cost analysis contributes to achieve optimal product reliability, corresponding to the lowest life-cycle cost, evaluating the trade-off between higher acquisition costs and lower maintenance and support costs. In order to be realistic, an LCC model should represent to the better extent possible the characteristics of the product, its intended use environment, maintenance policy as well as any constraints or limitations (one possible limitation is any financial constraint). Moreover, it should be comprehensive, including every factor relevant to LCC, though simple enough to be useful to support quick decision making, updating or any future modification. Therefore, it should be flexible and designed in such a way that permits evaluating specific elements of LCC independent from the others.

In practical terms, a simple LCC model is basically an accounting structure that contains mathematical expressions to estimate the cost associated with each of the cost elements constituting the LCC.

All empirical relationships, elements and other constants and variables of the model should be included so that it can become more reliable. Comparisons with any existing LCC model or even historical data of similar products collected during its life-cycle may be suitable to validate the model. Another aspect to incorporate in the model is the incurred costs for functional changes or product improvements. This is particularly important in civil infrastructures as they are usually designed to resist over a very long life period, where technology adaptations and improvements should be considered.

Detailed LCC breakdown into its constituent cost elements bring more precision to life-cycle cost estimation and permit to control the origins of risk associated with it better. Obviously, breakdown structure complexity depends largely on the complexity of the product. A designing
approach involves categorizing every cost in cost elements based on work/product category, life-cycle phase and applicable resource(s). Other LCC breakdown structures are possible, dividing life-cycle cost into two parcels: recurring and non-recurring costs\(^2\), or even fixed and variable costs.

Therefore, estimating costs may be based on three methods: engineering cost method, analogous method or parametric method. In fact, one or more of these methods may be used together. Engineering cost method involves examining the product component by component, part by part and using standard established cost factors such as manufacturing estimates properly updated to the present time by the use of appropriate factors (e.g. consumer price index components). Analogous cost method may be the quickest and simplest way to estimate costs. It is based on historical data from a similar product, updated as well to reflect cost escalation. At last, the parametric cost method uses parameters and variables to develop cost estimation relationships in form of equations. Some assumptions are made on the values of the system based on historical data or expert opinion. For example, the rate of failures in a system is a parameter used to estimate the cost of corrective maintenance of that system. This method to its extreme complexity may include typical stochastic processes analysis, assuming some randomness in the parameters and variables used. Note that the analogous cost method can also be improved by using time-series forecasting as an example, decomposing it in different trend lines.

LCC should be seen as a multidisciplinary activity. Therefore, analysts must be familiar with typical cost elements of the product in question, but also financial principles. Using different sources, which may include representatives of both the suppliers and customers, contribute to better estimate the different costs throughout the product life-cycle.

LCC modelling should also assess uncertainties and risks, performing a sensitivity analysis in order to identify cost drivers. The use of maximum and minimum values may be extremely useful to set boundaries in the total life-cycle cost. More accurate and complex analyses can be performed, as it will be seen later on, such as the Monte Carlo simulation to deal with uncertainty. In fact, many factors contribute to uncertainty and risk, such as lack of information, introduction of new technology, or even political and economical circumstances (including legislative changes). Elements like predicted inflation rates, labour, material and overhead costs, may cause additional uncertainty.

In order to make LCC analysis coherent and ensure that it can be easily understood by decision-making actors, it should begin with the development of a plan addressing the purpose and scope of the analysis. This plan should identify the analysis objectives, for example:

\(^2\) Recurring costs are regular costs incurred for each item produced or each service performed, while non-recurring costs are unusual costs, unlikely to occur again, also called extraordinary costs.
determine LCC for a specific product in order to support planning, contracting, budgeting or similar needs; or evaluate the impact of different alternatives such as design options or maintenance policies; or even identifying major contributors to the LCC of a product.

Factors such as the operating environment and maintenance support for the different life phases of the product, or even the time period considered in the analysis may influence greatly the outputs of LCC. Moreover, all the assumptions and limitations such as performance requirements need to be evaluated as they might restrict the range of acceptable options.

Detailed LCC models should take into account the availability of data and also the degree of sensitivity required to accomplishing the analysis objectives within the available time.

Therefore, the main steps to perform Life-cycle Costing are:

1) Collect data for all cost elements considered in the breakdown structure;
2) Perform LCC analysis of product operating scenarios;
3) Report analyses trying to identify optimum support scenarios;
4) Identify cost drivers;
5) Quantify any differences in product performance, availability or any constraint that may affect the applicability of any scenario considered;
6) Categorize and summarize LCC model outputs;
7) Analyze the robustness of the model, performing a sensitivity analysis.
8) Review LCC model outputs against the objectives defined in the analysis plan.

Note that in step 7) sensitivity analysis should pay particular attention to some assumptions made related to the time value of money, quantified by the discounting rate used in the analysis.

After the steps described above, all this process should be documented in a report that allows decision-makers, users and other interested parties to easily consult; whereas it is highly recommended that a review, if possible independent, should be carried out to confirm the correctness and integrity of results.

Moreover, updating the LCC model so that it can be exercised throughout the life-cycle of the product is extremely important in cases where different strategies of maintenance should be adopted in distinct life phases, in order to reduce the total life-cycle cost.

Railway infrastructure as a product has its own specifications. Since railway infrastructure has a long life span and investments are very costly, maintenance decisions should be made from a life-cycle cost perspective. LCC is particularly important in order to
develop an optimal maintenance strategy, as the IM is responsible for ensuring a certain RAMS level, performing the necessary maintenance operations. Otherwise, the IM will be penalised by the regulatory entity, based on a performance payment scheme.

2.3 - Main obstacles to LCC implementation in Railways

LCC requires favourable conditions for acquiring reliable data in practical time for decision-making and, sometimes, consistent empirical data on maintenance costs and on degradation of different components of the infrastructure are missing or the action to collect them is extremely time-consuming compared to expert opinion-based inputs.

Although decision-makers face the unavailability of data, civil infrastructures are still assets that degrade slower than mechanical equipment, providing them with more time and reliable data. LCC analysis should start in an early stage assessing then the uncertainties attached.

In some industries, LCC is not so used because consequential costs, in the event of an asset breakdown, are difficult to estimate and can be several times its capital value. However, in civil infrastructures, LCC may be an appropriate means to apply the principle of user-payer through fair taxation, aggregating the necessary societal support among stakeholders (citizens, private sector, operators, end consumers). In Railway infrastructure, LCC application still lacks a more integrative perspective.

In her PhD thesis, Ulla Espling (2007) proposes to identify and describe factors influencing the development of a proactive strategy for operation and maintenance of infrastructure, and then formulating a conceptual maintenance framework, taking into account factors such as outsourcing and partnering, benchmarking and risk management. In fact, Espling found that including partnering in maintenance strategy between different actors brings significant advantages: cost reduction, time savings (e.g. reduction on train delays by increasing reliability and availability) and improved project quality. Moreover, Espling argues that maintenance contracts can be improved by including incentives for contractors in achieving quality standards and by sharing more reliable data between IMs and contractors.

Patra (2007) presented the results of the survey conducted within the INNOTRACK project, of RAMS and LCC work on rail tracks in Europe, trying to find out the incorporated rules and standards that are currently being used by IMs and the models and tools used for RAMS and LCC analysis. He concluded that LCC and RAMS are still at a very early stage of implementation for railway infrastructure. Patra also defends that RAMS and LCC validation
after installation of track is necessary in order to check RAMS and LCC targets laid in the design phase. However, no clear method is proposed.

In the European project IMPROVERAIL (2003), other important obstacles to LCC in Railway management are discussed. Although the increasingly competitive business environment, declining resources and an ever-increasing need to obtain value for money will inevitably lead to the wide adoption of LCC in Railway, there is still a lack of infrastructure LCC models and the knowledge at the technical-economic interface is still insufficient. In fact, the project handbook (IMPROVERAIL, 2003) identifies relevant areas for knowledge acquisition and decision support, such as: the effect of standards and quality levels on cost, planning and optimization of asset-replacement strategies, asset condition monitoring and prediction for tailor-made maintenance processes, the understanding of mechanisms and quality-aspects of various maintenance processes and inspection tools, or even the commercial assessment of the impact of rolling stock quality on infrastructure LCC among others.

Therefore, studying Life-cycle Cost became extremely important as one of the European commission objectives is reducing the life-cycle cost of railway infrastructure by 30%. Therefore, this study may be useful to enhance performance of the Portuguese Railway Infrastructure manager, so that it will gradually achieve excellent performance, with best practices in terms of design and maintenance from a life-cycle perspective, but also more conscious and transparent decision-making processes and the necessary evaluation of risks and uncertainties inherent to investments of this nature.

2.4 - LCC in other transport infrastructure: roadway infrastructure

Life-cycle cost as an economic analysis to support investment decisions started to be performed for other transport infrastructures before it was performed for Railways. This section will discuss some important issues on LCC in Roadways, but it does not intend to give detailed information on how to perform LCC analysis in Roadway infrastructure. Nonetheless, it highlights the types of costs included, while discussing the necessity to evaluate them in the railway infrastructure.

In Roadways, especially in pavement design, life-cycle cost analysis has walked a long way compared to Railways, partly because the Railway system is more complex, but also because earlier liberalization of the market turned actors into more conscious decision-makers.

In order to provide better investment decisions and serve as a guidance of good practices and recommendations in conducting life-cycle cost analysis in pavement design, the US Department of Transportation published an interim technical bulletin (FHWA, 1998). The
procedural steps involved in conducting Life-cycle Cost analysis for design purposes, include establishing alternative pavement design strategies, estimating performance periods, activity timing, agency costs, user costs and finally computing the net present value so that conclusions can be drawn. While evaluating competing design and associated maintenance strategy alternatives, LCC analysis should focus on the costs that are actually different.

Pavement cost structure includes most frequently initial construction cost, major maintenance cost, rehabilitation cost, and salvage value (residual value and serviceable life). The salvage value, meaning the remaining value of the investment at the end of the analysis period, may be included as a negative cost. There are two main components associated with salvage value: residual value and serviceable life. More sophisticated analyses may include user delay costs, but some public agencies do not consider them, because they are not direct costs ("out of pocket"). Note that the quantification of all these costs is determined by using available previous construction and maintenance projects historic data.

LCC structure assesses user costs: delay, vehicle operating and crash costs. User costs are heavily influenced by current and future roadway operating characteristics, which are related to the current and future traffic demand, facility capacity, timing, duration and frequency of maintenance works, inducing capacity restrictions, congestions and speed reductions. In fact, user costs are all those costs incurred by the highway user over the life of the project.

To quantify user costs, we have to assume some hypotheses. Firstly, to determine user delay costs, we have to assume a value for time. This may be quite controversial, because people value time very differently depending on many factors (e.g. purpose of travel). The bulletin analysed comprehend delay costs calculated based on different vehicles and travel category (personal, business, truck drivers) and trip type (local or intercity).

Performance life for the initial pavement and sub-sequential rehabilitation activities has a major impact on LCC analysis results. It affects the frequency of agency intervention on the highway infrastructure, increasing agency costs and user costs during periods of maintenance activities, especially delay costs. Performance life expectation can be estimated through pavement inspection data and historical experience, based on the process of pavement degradation. Therefore, operational pavement management systems should supply accurate pavement condition and traffic volumes. As this work will confirm later on, traffic volumes (usually expressed in tonnage) are also used as the main explaining variable in Railway track degradation models.

Note that in Roadways, very rarely maintenance activities imply complete unavailability of the infrastructure. Most of the times, only one lane is closed for maintenance, which reduces capacity and may cause congestion with the associated delay cost. As in Railways, unplanned
maintenance usually implies complete unavailability of the infrastructure (at least in one way), whereas planned maintenance takes place during the night, and thus it is not traffic disruptive. Only in cases of double tracks in the same way, unplanned maintenance may not disturb traffic depending on operating conditions.

The proposed life-cycle cost structure includes, among others, crash costs. These costs may be relevant in Roadways but not taken into account in Railways, as accidents and crashes are more frequent in roadways. For example, the Japanese Shinkansen system has an unblemished safety record, carrying 275 million passengers per year without a major accident in its 44-year history, whereas the European high-speed rail network had a similar remarkable record, with no deaths or serious injuries until 1998, when the InterCityExpress (ICE) train connecting Munich-Hamburg crashed in Eschede. This accident occurred on a conventional intercity track, not a high-speed infrastructure, which may raise concerns that operating on tracks not built specifically for high-speed may pose higher risks (Harrison, 2000). Although the train is regarded as the safest mode of transport, the German accident shall remind us that safety can never be taken for granted. Note that costs associated with natural disasters are also not included in this analysis, as their probability of happening is very low.

Agency costs are those related to the initial design cost added to the subsequent rehabilitation costs. They typically include all the preliminary engineering studies, contract administration, construction supervision and construction costs, as well as future routine and preventive maintenance, resurfacing and rehabilitation cost. Moreover, agency costs may also include operating costs such as: pump station energy costs, tunnel lighting and ventilation costs.

In terms of probabilistic analysis, the state of the art of LCC in Roadways is far more advanced and sophisticated when compared to Railways. For instance, there are some important studies on quality of fit tests using past data of previous projects to estimate components of LCC depending on pavement material, usually following Lognormal distributions (Tighe, 2001). This source of knowledge sustains Monte Carlo simulations for more precise assessment of LCC, and therefore, a more comprehensive risk-based appraisal.

Another important aspect discussed is the fact that minimizing LCC from the perspective of the Infrastructure Manager may not coincide with the desirable societal optimum, as many external costs may not be considered. Note that no cost on CO₂ emissions is assessed in the analysed bulletin, though it can be argued that this cost may not vary significantly for different design and maintenance options.
2.5 - LCC in Railways: the particular Zoeteman model

Most Infrastructure Managers (IMs) lack an Asset Management System to support track maintenance decisions, from a life-cycle cost perspective, in order to minimize total long term costs of ownership. In fact, evaluating consequential costs of alternative maintenance regimes, risks and related costs may help the decision-making process, in particular that of deciding on the use of preventive maintenance in order to postpone renewals and reduce traffic disruptions.

In this section, the Zoeteman LCC model will be discussed - a decision support system named LifeCycleCostPlan, which involves considering degradation models of infra components. Other LCC models based on historic cost and performance data exist, but they lack in integrating RAMS analysis in them.

Zoeteman proposes a model to quantify Life-cycle Cost \(\text{LCC}\) of the Railway Infrastructure, from the perspective of an Infrastructure manager. It includes construction costs \(\text{CC}\), spot maintenance costs \(\text{MC}\), periodic maintenance and renewal costs \(\text{RC}\), delay costs \(\text{DC}\) and organizational costs \(\text{OC}\). Every single cost is estimated for a specific year \(y\) and discounted to the base year considering a constant discount rate \(r\), throughout the life-cycle.

\[
\text{LCC} = \sum_{y=0}^{n} \frac{TC(y)}{(1+r)^y} = \sum_{y=0}^{n} \frac{CC(y) + MC(y) + RC(y) + DC(y) + OC(y)}{(1+r)^y} \tag{1}
\]

Construction costs \(\text{CC}\) are those incurred during an initial phase, including all costs in order to build the infrastructure (design, materials, labour and equipments, etc). Several factors influence the construction cost of a new railway line. Firstly, among these factors are all features of the design and layout characteristics (number and size of bridges and tunnels, investment in substructure quality, expropriation costs especially in urban areas, number of switches, crossings and electrical substations, etc per kilometre). Moreover, design will directly influence the life-cycle costs via the costs of construction, and indirectly via the initial and inherent design quality, which will influence maintenance specifications. That is mainly why design and maintenance strategy are steering variables in the conceptual model analyzed afterwards. Secondly, building conditions such as labour cost and safety legislation, which vary from country to country, site accessibility, construction method (prefab or in-situ) and unit rates for materials, machines and personnel influence construction costs. Finally, as financial conditions and budget constraints may influence design options, they are included as an external variable that restrains the amount of construction work. Note that financial conditions also restrain the amount of maintenance work and can seriously compromise the availability of infrastructure.
As an example, to have an idea of construction costs, Profillidis compiled the cost data from tracks for high-speed trains built during recent years, from data of UIC and constructors (Table 2.1). They can give a first rough estimation of the construction cost of a new high-speed railway line. Moreover, Figure 2.2 illustrates a typical distribution of construction costs of a new railway line to the various components of the railway system.

<table>
<thead>
<tr>
<th>Country</th>
<th>Line</th>
<th>$V_{\text{max}}$ (km/h)</th>
<th>% on ballast</th>
<th>% on concrete slab</th>
<th>% of tunnels</th>
<th>% of bridges</th>
<th>Construction cost per km (million of €)</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>'TGV Méditerranée'</td>
<td>350</td>
<td>100%</td>
<td>-</td>
<td>6.5%</td>
<td>12.7%</td>
<td>16.95</td>
</tr>
<tr>
<td>Spain</td>
<td>Madrid – Barcelona</td>
<td>270 - 300</td>
<td>100%</td>
<td>-</td>
<td>26.8%</td>
<td>3.4%</td>
<td>6.1</td>
</tr>
<tr>
<td>Germany</td>
<td>Cologne – Frankfurt</td>
<td>300</td>
<td>-</td>
<td>100%</td>
<td>26.5%</td>
<td>4.3%</td>
<td>21.7</td>
</tr>
<tr>
<td>Italy</td>
<td>Rome – Naples</td>
<td>300</td>
<td>100%</td>
<td>-</td>
<td>17.8%</td>
<td>24.0%</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Table 2.1 - Construction costs (values of year 2006) of high-speed tracks constructed during recent years, Profillidis (2006).

As Table 2.1 shows construction cost per kilometre for high-speed tracks have a significant dispersion. The German line linking Cologne to Frankfurt costed around 21.7 M€ per km, three and a half times the cost of the Spanish line Madrid-Barcelona (6.1 M€ per km). The distribution of construction cost of a new railway line to the various components presented below shows the predominance of civil engineering projects (including subgrade, expropriations, tunnels and bridges), representing around 55% of the total construction cost, whereas the track itself represent around 15%, even less than electrification, signalling and telecommunications, which totals 18%.

![Figure 1.2 – Distribution of construction cost of a new railway line to the various components of the railway system, Profillidis (2006).]
Spot maintenance costs \( (MC) \) differ from periodic maintenance and renewal costs \( (RC) \). Spot maintenance is not planned and scheduled in advance throughout the infrastructure lifecycle, whereas periodic maintenance is supposed to be planned in advance.

Zoeteman highlights that the costs of ongoing maintenance, meaning spot maintenance costs, are usually small compared to the costs of periodic maintenance and renewal. In fact, spot maintenance costs are those costs spent to repair local defects manually or supported by small machines (e.g. spot tamping by portable tamping machines); whereas periodic maintenance usually involve more sophisticated machinery (e.g. Plasser and Theurer 09-CSM).

According to Profillidis, Infrastructure maintenance costs in general comprise maintenance and renewal of track (rails, sleepers, ballast) and subgrade; maintenance of electrification, signalling and telecommunications facilities and substations; maintenance of tunnels and bridges; and also maintenance of platforms. All the different components of the railway infrastructure, here understood in a wider definition, will start deteriorating once the operational phase commences, more slowly or rapidly depending on the specific component. Maintenance costs\(^5\) per year of 44 300 €/km of track is reported for France, and a value of 56 500 €/km is reported for the Netherlands (monetary values for 2006). This cost is distributed through the various maintenance components: 65% for track and platforms, 30% for electrification, signalling, telecommunications and substations and the remaining 5% for bridges and tunnels (Profillidis, 2006).

Many factors concur to influence maintenance costs. For instance, initial design quality is a determining factor that influences the rate of deterioration, and therefore, the laying of track should be carried out very carefully, verifying every geometrical parameter (e.g. transverse slope 3-5% of the subgrade). Other external variables influence the quality of geometry & structure, and therefore maintenance costs (whether or not planned), such as transport concept, physical environment, design and maintenance specifications and maintenance and building conditions.

Delay costs \( (DC) \) basically depend on the delivered reliability and the cost consequences of the delay incurred (lost revenues and increased operating costs). Moreover, a performance payment regime may be included, prescribing penalty rates. Another way to assess delay costs is developing a cost-benefit model considering three costs: direct income loss of the operators, direct extra costs for using buses or rerouting traffic and societal extra costs.

\(^5\) Maintenance cost is here understood as the total cost of adding spot maintenance costs and periodic and renewal costs.
As for organizational costs \((OC)\), they should be included only if they change significantly due to a different decision alternative. Organizational costs are the annual flat costs required for inspection, maintenance planning and incident response organisation. Profillidis (2006) considers a similar cost: operation cost of infrastructure, estimated per year at 1.25 \(€/\text{train.km}\) (values of 2006), which includes traffic management (92% of total operation costs) and schedule planning (8% of total operation costs).

Zoeteman builds up a conceptual model from the perspective of a separated Infrastructure manager (IM). It takes into account four types of variables: steering, external, internal and effect variables. Steering variables are all those factors that an IM can directly influence, as external variables are all those that are not under the IM’s control. For instance, financial conditions or even transport concept, in which railway infrastructure is included, may be controlled by other actors. Internal variables make part of the maintenance and renewal process, quantifying planned maintenance volume and failure performance. This process will then influence the effect variables, overall costs and RAMS performance level. The Figure 2.3 presents the conceptual model, identifying general factors that influence costs and performance of the rail infrastructure.

As mentioned before, separated Infrastructure Managers have two key steering variables: design and maintenance strategy. Design specifications will influence the amount of construction work and the quality of geometry & structure, whereas maintenance strategy will influence directly the planned M&R (Maintenance and Renewal) volume and failure performance and expenditures on infrastructure. As the initial quality of geometry influences drastically its deterioration, design affects indirectly the planned M&R volume too.

Financial conditions (budget constraints) are a decisive external variable, as they may influence not only the amount of construction work but also define the actual realised M&R volume and may increase deterioration and failures if necessary maintenance expenditures on infrastructure are postponed. This distinction between the planned maintenance and renewal volume and the volume actually produced is important to include in this model, as these two work volumes may differ significantly. In fact, under budget restrictions, life-cycle cost is also a suitable method for prioritising projects or maintenance activities.

Other external variables are design and maintenance specifications, building conditions, transport concept, physical environment and maintenance conditions. Note that design and maintenance specifications refer to restrictions imposed by maintenance and safety standards, which may be part of the overall maintenance policy or simply legally required. They intend to guarantee safety, riding comfort and noise levels. Whereas maintenance conditions refer to work methods and equipment available, repair concept (full repair or a quicker temporary repair
Figure 2.3 - General factors influencing the costs and performance of the rail infrastructure, Zoeteman (2004).
method) or even the possibility to cluster maintenance works with associated gains of efficiency because of effects of scale and scope.

The M&R process, generated by the steering variables and conditioned by the external variables, results in overall levels of ownership costs and reliability, called effect variables. In fact, effect variables also include planned availability and noise, vibrations, safety and riding comfort. The M&R process is visualised through the internal variables mentioned before, which include the amount of construction work, quality of geometry and structure, planned and maintenance/renewal volume, produced volume and failure performance and expenditures on the infrastructure.

Planned availability level may influence consequential costs. If preventive tasks cannot be performed in non-operative time, they would need track possessions, reducing availability of track and cause losses of revenue or even penalties. Other costs included are those caused by temporary speed restrictions during and after M&R works, responsible for train delays. Repetitive bad performance may lead to losses of goodwill from the end-users, representing other costs that are not included. Note that the degree of traffic disruption caused by M&R works is influenced by the track layout (availability of crossover switches) and traffic density (maximum train speeds, distribution of speeds, and the mix of passenger/freight traffic).

Based on the conceptual model, Zoeteman identifies five calculation steps of the Decision Support System (DSS). The picture bellow presents them, showing the respective data needed for the calculation steps on the left and right sides: Construction and maintenance data, on the left; and Financial and transport data, on the right.
Step 1 – Estimating the loads on the infrastructure

As quality degradation is modelled as the loss of quality per unit of time or load, the first calculation step estimates the load on infrastructure elements in terms of accumulated tonnage (based on the number of train passages and the type of train). In fact, accumulated tonnage is a common measure used to determine the deterioration of the track quality, providing an indication of when maintenance and renewal are needed. Therefore, a theoretical traffic load ($T_{th}$) should be calculated as the following expression shows:

$$T_{th} = S_p \cdot (T_p \cdot k_t \cdot T_{tp}) + S_{fr} \cdot (k_{fr}T_{fr} + k_t \cdot T_{tf})$$

In which: $T_p$ is the mean daily passenger tonnage hauled [gross tons]; $T_{fr}$ is the mean daily freight tonnage hauled [gross tons]; $T_{tp}$ is the mean daily tonnage of tractive units used in passenger traffic [gross tons]; $T_{tf}$ is the mean daily tonnage of tractive units used in freight traffic [gross tons]; $k_{fr}$ is a coefficient that takes into account effects of both the load and wear provoked by freight bogies (normally equals 1.15); $k_t$ is a coefficient that takes into account wear resulting from traction vehicles (normally equals 1.40); $S_p$ and $S_{fr}$ are coefficients related to train speed, $S_p$ for passenger trains and $S_{fr}$ for freight trains (varies from 1.00 to 1.50, assuming higher values for higher speeds).
The formula above converts the loads of various trains into equivalent passenger train loads, and then takes into account the effect of different speeds. Another alternative formula to incorporate the effect of different running speeds and vehicles in tonnage is considering a fictive tonnage called notional ($T_f$), given by the expression:

$$T_f = T_p \cdot \frac{V_{\text{max}}}{100} + T_g \cdot \frac{P_D}{18D}$$

In which: $T_p$ is the real load for daily passenger traffic; $T_g$ is the real load for daily freight traffic; $V_{\text{max}}$ is the maximum permissible speed [km/h]; $D$ is the minimum wheel diameter of the freight train [m]; $P_D$ is the maximum axle load with wheels of diameter $D$ [gross tons].

The alternative formula above explicitly incorporates the effects of different maximum speeds ($V_{\text{max}}$) for passenger trains, axle-loads ($P_D$) and wheel diameters ($D$) for freight trains. Both formulas can be calculated with the help of UIC leaflet 714, in order to convert the reference timetable into infra loads (Zoetemann, 2004).

Therefore, an essential input for the load calculation is a reference timetable, specified for different periods (years) in order to include forecast traffic growth or decline. It defines the number of trains per day, the types of train and locomotives, and the average number of train-sets in a day. Note that this timetable is not only useful for estimating traffic loads, but also needed for calculating schedule journey time ($SJ\dot{T}$) on the analysed railway line section. $SJ\dot{T}$ is the sum of journey times for all trains in a specified time interval (an year) and can be used in performance payment regimes as a basis for calculating the infrastructure availability (RAMS performance). It should also include the average agreed time window (slots) for planned maintenance.

$$SJ\dot{T}(y) = 365 \cdot TRACKS \cdot OH \cdot \sum_{tr} TPH_{tr} \cdot JT_{tr}$$

In the formula above, the schedule journey time($SJ\dot{T}$) is calculated for both directions together, considering the journey time ($JT$) for a single train and the total number of tracks ($TRACKS$). Its calculation is based on the operational hours per day ($OH$) and number of trains per hour per track ($TPH$) for all train types ($tr$). To estimate $SJ\dot{T}$ for a time interval of one year, we multiply $SJ\dot{T}$ for an average day by 365 days.

Based on the reference timetable, indicating the type of trains and axle-loads that run the track, total accumulated tonnage for the period of the reference table can be estimated (typical for a day). Multiplying this accumulated tonnage by 365, we can calculate the accumulated tonnage for a year. More accurate estimations may be achieved if the period of the reference table considered is broadened to a week, including possible differences of operation
in the average week analyzed (shorter trains in the weekends, for example). These values should be added year after year, since the beginning of operation, and include traffic growth predictions (materialized in more loads per year).

**Step 2 – Estimating the periodic maintenance volume**

Periodic maintenance and renewal activities can be estimated for each major infrastructure component (rail, ballast, sleeper, fastenings) based on forecast loads. The initial quality design, the deterioration rate and the M&R thresholds need to be known so that M&R intervals for each component can be calculated\(^4\). Maintenance thresholds are limits specified in terms of infrastructure parameters or in calendar-based or tonnage-based thresholds.

As carried loads usually differ for different segments of the same line section, residual lifespans and the timing of major M&R are only valid for distinguished and homogeneous infrastructure segments. Therefore, it is extremely important to separate the line into different sections. This separation depends on design characteristics (e.g. ballast track section or slab track section), but also on operating conditions (network analysis). In fact, this process is usually referred to as harmonisation and it will be discussed later in more detail.

Note that some deterioration processes of different components are interrelated, either through physical relationships or even through a conscious policy of harmonising a priori the maintenance of components in time and space, so that M&R intervals coincides and M&R works are made simultaneously (clustering of activities), benefiting from gains of scale in maintenance operations.

After estimating residual lifespans and necessary intervals for performing major M&R activities, the amount of maintenance work is estimated, and the respective number of shifts is distributed over the years. Additional input data as productivity rates (depending on machinery available) is necessary to calculate the duration of every maintenance work, permitting the construction of maintenance teams and the allocation of resources (materials, human, equipments). Therefore, operation research techniques may be helpful to schedule and manage maintenance activities.

**Step 3 – Estimating maintenance costs and possession hours**

Total costs for periodic maintenance are estimated based on the work shifts defined in step 2. To calculate these costs, data on unit costs of fuel and materials, manpower and machines need to be known. As the volume of M&R work is defined, resulting possession time and speed restrictions are deducted. Speed restrictions are necessary in order to expose

\(^4\) In chapter 3, mathematical expressions will show how to estimate M&R intervals.
gradually the infrastructure to the dynamic traffic loads after maintenance, as lateral track resistance drops considerably to 20%-40% of the initial value (Esveld, 2001). Speed restriction can be included in different formats, such as days or hours or even number of train passages. Usually, it is defined as the time interval to accumulate 2 million tons. After that, the speed limit may be lifted to its original value before tamping operations.

Periodic maintenance & renewal work is usually not disruptive, which means that it is performed in the ‘natural gaps’ of the timetable, not implying delays and associated costs. However, maintenance & renewal work, above a certain limit and concentrated in time, may become disruptive, as it may not fit in night shifts.

In step 3, the impacts of small maintenance and failures are also included. They are assumed to be partly related to the cumulative tonnage or years in service of the infrastructure components and partly independent from traffic loads, meaning they are more seen as time-based. FMEA (failure modes and effect analysis) technique is used to collect input data on failure process of each infrastructure component, if historical data is not available. Resulting costs, possession and speed restriction hours are then included per ton carried or per year. These minor maintenance activities often take place in non-operative hours, and therefore they should be carefully studied in order to assess planned possessions correctly.

**Step 4 – Estimating infrastructure availability and reliability**

The first three steps provide a set of estimates on individual M&R costs, and planned and unplanned track possessions as well as speed restrictions over the years. This step focuses on the estimation of availability and reliability of infrastructure and their consequences. In order to evaluate the consequences of traffic disruption (in terms of delayed and/or cancelled trains), a mathematical or simulation model based on the reference timetable should be used. Resulting delays are influenced by the track layout (availability of passing tracks, switches, or alternative routes), the braking and acceleration of the trains, and the feasible headways (signalling system). In fact, these factors contribute decisively to the flexibility of the infrastructure operation.

Several methods, some more refined than others, may be used to estimate delay times. As in many analyses, a trade-off between time available for analysis and model complexity used should be evaluated. Simple mathematic expressions can achieve rough estimations. However, they do not take into account many effects, included in more advanced dynamic analyses, as the rerouting of traffic or the ‘knock-on’ impacts of delayed trains. Note that these effects depend largely on the robustness of the operative timetable.
Total train delays can be roughly estimated by considering the sum of the time differences between passing time in case of a set speed restriction and the usual passing time (at usual maximum speed permissible) for every train delayed.

**Step 5 – Estimating total life-cycle costs**

The final step is to estimate the overall life-cycle costs for each decision alternative, which requires evaluating the impacts of system availability and reliability and the cash flows for every cost as expression (1) showed. Note that financing costs are usually included indirectly, by using higher discount rates for discounting the cash flows.

Costs and missed benefits incurred by the IM from disrupted transport operation should be included in delay costs. Therefore, a cost-benefit model may be developed, or even a regime of penalty rates to reflect those costs and revenues. These regimes are based in a performance payment mathematical algorithm (PPMA), usually defined by the regulator entity, depending on the contractual regimes between IM and operators. The PPMA attributes penalties to the IM on the basis of a calculated ‘availability level’, meaning the ratio of total train delay in minutes and the total scheduled journey time calculated for a period of time. Usually, penalty regimes have progressive rates, for example, basic penalty rates range from 500 €/min for minutes of delay in the interval from 99% to 96% availability to 3500 €/min for the interval between 92% and 90% (Zoetemann, 2004). These progressive regimes are incentives for the IM to provide a reliable railway system. Note that availability is not calculated on the basis of actual delay times (which are also influenced by third-party entities like the operators), and recovery scenarios should be simulated on the basis of a fixed timetable, so that a standard delay recovery period is deducted for each of the delayed trains.

In conclusion, Zoeteman proposes a model assuming a single rational stakeholder to manage the railway system, to quantify overall life-cycle costs. Note that assuming a multi-actor process would demand further complexity and requirements such as bridging interests, multi-perspective research and trustworthy analysis, mainly because different stakeholders may have distinct interests on LCC. Furthermore, a participative policy analysis contributes decisively for decision support, and its acceptability among stakeholders. In fact, this goal is drastically simplified if we assume a single actor – a separate Infrastructure manager, as Zoeteman does, and thus to some extent, it is a strong contribution to encourage ‘life-cycle thinking’ in decision-making processes among railway actors.
Chapter 3

Degradation and Maintenance of Railway Infrastructure

This chapter will discuss Degradation and Maintenance aspects of the Railway infrastructure. It begins with a brief description of infrastructure components, and afterwards a theoretical perspective of degradation models is included, discussing in particular the degradation of track geometry and the degradation of rails. Moreover, Maintenance and Renewal activities are identified, highlighting the importance of inspection in Maintenance.

3.1 - Railway Infrastructure components

Railway infrastructure components normally consist of rails, sleepers, fastenings, ballast, sub-ballast and subgrade. Sometimes, as in tunnels, the typical ballast bed is substituted with concrete slabs resting on track foundations. As Figure 3.1 below shows, two subsystems of a ballasted track are distinguished: the superstructure, composed of rails, sleepers, ballast and subballast; and the subgrade, composed of a formation layer and a base. The superstructure supports and distributes train loads and is subjected to periodical maintenance and replacement; whereas the subgrade on which the train loads are transferred to after superstructure, should not be subjected in principle to maintenance interventions\(^5\).

![Figure 3.1 – Ballasted track components, Profiliidis (2006).](image)

In a simplistic way, the classical railway track (which refers to the ballasted track) consists, therefore, of a flat framework made up of rails and sleepers, which is supported on ballast. Rails and sleepers are connected by fastenings. These components and other structures such as switches and crossings are considered as part of the railway track.

\(^5\) Maintenance of subgrade may be performed in exceptional cases, for example, if the geometry deterioration rate is very high and there is strong evidence that this is caused by bad condition of subgrade.
The subgrade or substructure consists of a formation that includes slopes, verges, ditches and any structure with them. It must confer sufficient bearing strength and stability, providing good drainage of rain and melted snow from the ballast bed. Sometimes, an intermediate layer is placed between ballast bed and formation to improve filtering characteristics, also contributing to better distribute loads, while ensuring protection against frost and separating the coarse-grain ballast from the fine sand (if necessary recurring to synthetic material like geotextile).

The ballast bed consists of a layer of loose, coarse grained materials, which can absorb considerable compressive stresses, usually made of crushed stone (grading: 30/60 mm). Note that though vertical bearing strength of the ballast bed is significant, lateral bearing strength is reduced. Moreover, another important aspect to consider is the contamination caused by weathering of the ballast material or by the upward penetration of fine particles (clay mixture), hindering water drainage, which may increase degradation rate of vertical settlements. Therefore, subballast protects the subgrade top from the penetration of ballast stones, while conferring better drainage of rainwater and distributing stresses to the subgrade.

Sleepers are the track components positioned between rails and ballast, ensuring appropriate load transfer and distribution from the rails to the ballast. Moreover, sleepers help to maintain constant rail spacing (specified by the track gauge) and rail inclination (1/20 or 1/40), while contributing with an adequate mechanical strength in vertical and horizontal directions. In electrified lines, they also provide the electrical insulation of each rail from the other. Large variation in design characteristics exist. In fact, sleepers can be made of concrete, timber or even steel, and for concrete sleepers they may assume different shapes: as the monoblock or the twin-block sleepers.

The rail may be seen as the most important component of track structure, as it accommodates the wheel loads distributing them over the sleepers. In fact, the rail guides the wheel, providing a smooth running surface and distributing accelerating and braking forces by adhesion. Moreover, it acts as an electrical conductor on an electrified line, conducting signal currents essential to the proper functioning of the signalling system. Although rail profile has evolved through time, the Vignole rail (or flat-bottom rail) is the standard profile used. It is derived from the I-profile in which, for the purpose of support and guidance, the upper flange is converted to form a rail head. The rail profiles most extensively used in Europe are the UIC 54 and the UIC 60\(^6\). Therefore, the flat-bottom rail can be divided into three parts: the head, the web and the foot. The rail head provides a high wear margin and its shape ensures good contact with the wheel, whereas the rail foot provides the necessary stability of the rail and load distribution to the sleeper. Note that the rail must mechanically ensure the required moment of inertia and adequate stiffness against bending and buckling.

\(^6\) The numbers refer to the rounded weight in Kg per meter.
3.2 - Degradation models

Predicting future degradation of infrastructure components is a key element in maintenance planning. A track deterioration model is intended to describe how an infrastructure facility changes from one condition to another within a specified time period. Therefore, a comprehensive track deterioration model must include track component deterioration models, in which each component condition should be monitored through time. As component condition evolves randomly through time, probabilistic and mechanistic deterioration models may be used to model component deterioration.

There are mainly two types of probabilistic models: continuous probability distributions (state based) and Markov-models (time-based). With the continuous probability distribution it is possible for a track section to be in a certain state, knowing the elapsed time since the last maintenance activity, whereas a Markov model distributes the infrastructure in different condition classes over time. These models usually assume that condition measurements through inspection reveal the true track condition without errors, though latent Markov processes may be used to overcome this assumption, so that imperfect inspections can be modelled.

Other degradation models have been developed, based on engineering principles or mechanistic laws of physics, though they may be hard to quantify and are hostage of real life and reliable data to calibrate their parameters. Mechanistic degradation modes are handled separately for each different component and ideally for distinct degradation modes. For example, ballast fouling and ballast deterioration should be analyzed separately and wear, corrugations and crack formation also demand separate degradation modes, which might be included in a comprehensive component degradation model. Later on in chapter 4, another approach will be discussed to include uncertainty in track deterioration models, based on mechanistic degradation models, though no distinct degradation modes are explored.

Some integrated approaches have been tried, developing integrated computer-based tools for the prediction of track behaviour under changing traffic conditions. Some authors showed that increasing axle load and train speed accelerates track degradation in general and rail wear, in particular. Moreover, sub-grade stiffness and ballast depth were found to be the most important parameters for track roughness (Zhang et al, 2000).

---

7 In fact, Markov models usually are given by stochastic processes that benefit from a simplifying property: future states of a system, only depend on the present state of the system and do not depend on the all system past history.
In order to build degradation models, regression techniques are used to estimate condition state as a function of a number of independent variables. Therefore, a prerequisite to reliable degradation models is a comprehensive historical database so that model parameters can be calibrated and potential correlations between variables explored. Nevertheless, regression analysis has shown little success on the estimation of track degradation. As an example, a report on US trials to estimate track degradation identified fifteen important causal factors, building up a multiple regression model including five explanatory variables with a low $R^2 = 0.49$ (Bing and Gross, 1983 cited by Andersson, 2002).

In fact, in the past, degradation models ended up being simplified down to a linear representation, so that feeble computers of that time could cope with them. Moreover, at that time few condition data were available (especially in digital format), which resulted in simplistic representations of track behaviour (Jovanovic, 2004). Therefore, new challenging future developments in deterioration models are expected, as more comprehensive databases of component condition related with rail operation and design characteristics are implemented in IM organizations.

To build up a track quality degradation model, a condition-parameter, measuring track quality should be defined. Moreover, a list of maintenance activities that influence the behaviour of this parameter must be drawn, distinguishing essential activities from temporary activities. Essential activities are those dramatically influence the behaviour of a certain condition-parameter (e.g. ballast renewal to the track geometry behaviour), whereas temporary activities are those performed several times, changing temporarily the value of a condition-parameter, but whose efficiency decreases over time (e.g. tamping operations).

**Figure 3.2 - Analysis principle applied on an hypothetical track geometry deterioration, Esveld (2001).**
As the Figure 3.2 illustrates, over time, as the track grows older, several things change. The efficiency of tamping, meaning the intensity of quality increase represented by the vertical drop, decreases. Another change is the ‘deterioration rate’, i.e. the slope of the line defined by measured points (during linear phase), which may be represented by using different angles ($\alpha_1$ and $\alpha_2$). These two events have their impact on the required tamping frequency, becoming higher and higher, i.e. the time period between two tamping works (tamping cycle) becomes shorter. Eventually, tamping frequency becomes so high that it is no longer seen as efficient and rather ballast renewal is appropriate. Three phases of deterioration are distinguished in the figure above: rapid deterioration (a), linear deterioration (b) and again rapid deterioration (c), resembling the ‘bathtub’ curve.

In fact, theoretically the ‘bathtub’ curve is often used to model the chance of failure over time. As Figure 3.3 illustrates, in the first and last life phases, the failure rate is higher. The first phase is one of infant mortality, sometimes with undesirable designing problems contributing for failures. The system tends to reach a ‘steady-state’, with mostly random failures, and finally, the failure rate increases again as the system becomes older and obsolete, suggesting ‘wear out’ so that renewal in needed.

![Bathtub curve diagram](image)

Figure 3.3 - Theoretical failure probability over lifetime, Zoetmann (2004).

The Weibull distribution is generally considered appropriate to model times-to-failure of electronic and mechanical components, equipments or systems, which exhibit failure rates similar to the ‘bathtub curve’ over time. In fact, the Weibull distribution is a very versatile continuous probability distribution, defined with two parameters: a shape parameter ($k$) and a scale parameter ($\lambda$). Its density function is given by the expression:

$$f(x, k, \lambda) = \begin{cases} k \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k}, & x \geq 0 \\ \frac{k}{\lambda} & x < 0 \end{cases}$$
The exponential distribution is also very commonly used to model times-to-failure random variables, but for systems with constant failure rates. Its density function is given by the expression:

\[ f(x) = \lambda e^{-\lambda x} \]

In fact, the exponential distribution models the time between consecutive failures in a Poisson process. Later on, some degradation models integrated in LCC models for rail and ballast will be discussed, using both distributions described above.

Many degradation models are usually based on track parameters to quantify different degradation states and distinct degradation modes. There are mainly two different classes of track parameters: geometrical parameters and mechanical parameters. Geometrical parameters (e.g. track defects) degrade much faster than mechanical parameters and are usually reversible through corrective maintenance, while mechanical maintenance parameters cannot be easily restored without parts replacement (rails, sleepers, fastenings). In this context, it is important to distinguish track defects from rail defects. Track defects, as typical geometrical parameters, are deviations between the actual and theoretical values of geometrical track characteristics, whereas rail defects are of a mechanical and microscopic nature and in most cases non-reversible, due to initial manufacturing imperfections (internal discontinuities of rail) which may give rise to rail fatigue. Later on in chapter 4, some typical rail defects will be included in a stochastic model to assess the economic life of a rail.

3.2.1 - Degradation of track geometry

Track geometry degradation can be quantified by using geometrical parameters deterioration. Normally, five track defects are considered: the longitudinal defect, the transverse defect, the horizontal defect, the gauge deviations and the track twist. A brief description of each track defect is presented below.

The longitudinal defect (LD) is the difference between the theoretical and the real value of track elevation, given by the expression:

\[ LD = z_{th}(T, x) - z(T, x) \]

In which \( z(T, x) \) is the average of track elevations between the internal and external rail, corresponding to a traffic load \( T \) (since the last track maintenance), at a kilometric position \( x \); whereas \( z_{th} \) is its theoretical value (initial track elevation).
The transverse defect ($TD$) is the difference between the theoretical and the real value of cant, given by the expression:

$$TD = (z_{int} - z_{ext})_{th} - (z_{int} - z_{ext})$$

In which $z_{int}$ is the track elevation of the internal rail and $z_{ext}$ is the track elevation of the external rail. Both correspond to a traffic load and a kilometric position, for which $TD$ is calculated. Note that the cant is the difference between the track elevations of the internal and external rails.

The horizontal defect ($HD$) is the horizontal deviation of the real position of the track from its theoretical positions. It depends on the transverse track effects and on the particularities of the rolling stock.

The gauge deviations are not only affected by the particularities of the rolling stock, but also by the mechanical properties of track materials. The track gauge is defined as the distance between the inner sides of the rails, measured 14 mm below the rolling surface. Gauge deviations have some tolerance, for standard gauge (1.435 m) varies between minus 1 mm or plus 3 mm than its standard value.

Track twist is defined as the deviation of one point of the track laying from the plane defined by the other three. Note that along straight and circular sections (where cant is constant) the four points of the track laying should lie in the same plane. Therefore, track twist may be seen as the transverse defect per unit length ($\Delta l$), given by the expression:

$$Track \ twist = \frac{\Delta TD}{\Delta l}$$

Degradation of track geometry greatly depends on ballast degradation. In fact, the ability for ballast to keep stable geometry depends on initial ballast quality, physical state and loading characteristics. Ballast materials must satisfy several requirements and properties. Moreover, it must resist breakdown through fracturing and also resist attrition through wear with neighboring ballast particles, while it anchors the sleepers contributing to the necessary lateral resistance. Apart from that, it must confer good mechanical stability, providing maximum friction through angular and rough surfaces of ballast particles.

When a track is loaded, especially to high-frequency load variations, it may result in non-elastic deformations of the ballast and substructure layers. When unloaded, the track will not return exactly to its original position. Different parts of the track will accumulate these small non-elastic deformations in different ways. This phenomenon is usually referred to as differential
settlement. The severity of these settlements caused by the repeated traffic loading depends on the quality and behaviour of the ballast.

Therefore, two major phases are distinguished in track settlement occurrence. One is directly after tamping (as track position is readjusted), in which track settlement is relatively fast until the ballast is consolidated (gaps between ballast particles are reduced). The second phase is much slower and the relationship between track settlement and time (or load) is approximately linear. This second phase is governed by several basic mechanisms of ballast and subgrade behaviour, such as continued volume reduction caused by particle rearrangement, particle fracture or even abrasive wear at contact points with other particles. Moreover, during this second phase, inelastic recovery on unloading may cause permanent deformations, and also substructure layers may penetrate into ballast voids and cause changes in track levelling, or even sleepers may sink into the ballast bed.

Methodologies to calculate the permanent deformation response have been developed. Making the assumption that the ballast starts from an uncompressed state, the total permanent strain $\varepsilon$ after load cycle $N$ of a series of identical load cycles is given by:

$$
\varepsilon = \varepsilon_1 [1 + C \cdot \log(N)]
$$

In which: $\varepsilon_1$ is the permanent strain after the first load cycle ($N = 1$) and $C$ is a dimensionless constant controlling the rate of growth of deformation.

Another early model trying to simulate track settlements was suggested by Shenton (1985 cited by Dahlberg, 2003 and Öberg, 2006). According to the author, the logarithmic settlement law above (proportional to $\log(N)$) may be considered reasonable over a short period of time, but it might significantly underestimate the settlement in the case of large numbers of loading cycles. Therefore, based on laboratory and filed experiments, Shenton suggested a settlement law where the settlement is proportional to the fifth root of the number of load cycles, which would fit well with site measurements up to $10^6$ load cycles. To obtain an acceptable fit for larger values of load cycles, a linear term was added to the settlement law, resulting in the following expression:

$$
y = K_1 N^{0.2} + K_2 N
$$

In which: the constants $K_1$ and $K_2$ are model parameters, whose numerical values depend on a number of factors such as axle load, rail section, sleeper spacing, and track and foundation stiffnesses.
Shenton identified axle load as one of the most important factors influencing the settlement, finding that at low train speeds, a linear relationship between axle load and settlement is a reasonable approximation. In fact, tri-axial tests revealed that the effect of higher loads dominates the track settlement phenomenon, and lower loads (below half of the maximum load) had no influence on the track settlement. This remained valid even when the number of low loads was around 90% of the total number of load cycles, implying that locomotive axles might be the main cause for track settlement. For \( N \) different axle loads \( F_i \), Shenton suggested an equivalent axle load \( F_{eq} \) given by the expression:

\[
F_{eq} = \left( \frac{\sum_{i=1}^{N} F_i^5}{N} \right)^{0.2}
\]

Fröhling (1997, cited by Dahlberg, 2003) explored a new approach to calculate differential track settlement due to dynamic wheel loading and spatially varying track support conditions. Based on measured results, it was found that the differential track settlement was dominated by the spatial variation of the track stiffness.

Guérin (1999, cited by Dahlberg, 2003) used a scaled-down model to investigate the ballast and subground settlement. The settlement is divided into two phases: one of ballast compaction, with a relatively large ballast settlement rate, and the other with a steady-state settlement rate. The length of the first phase depends strongly on the quality of the pre-compaction of the sample, and it comprises from 50 thousand cycles up to almost 1 million loading cycles. For the second phase, Guérin modeled the settlement \( \tau \) per loading cycle \( (N) \) as a function of the maximum elastic deflection \( d \) of the ballast and subground sample during the loading cycle as:

\[
\frac{d\tau}{dN} = \alpha d^\beta
\]

In which: \( \alpha \) and \( \beta \) are material parameters, determined by linear regression.

As described above, this phenomenon is rather complex. Immediately after tamping, large settlements occur and their non-uniformity develops irregularities due to inhomogeneities in subgrade support conditions. This results in differential settlements which lead to the development of irregularities in the wavebands experienced by the rolling stock.

Many investigations were carried out by the Office for Research and Experiments (ORE, cited by Esveld, 2001), in order to understand the fundamentals of the deterioration mechanism and controlling this phenomenon. The examination of data available from a number of administrations showed that the factors governing the rate of deterioration were not obvious and
unknown factors in the track were the most critical in determining both the average quality and the rate of deterioration.

Original tests on the rate of deterioration of track geometry were carried out by ORE committee D 117 (Esveld, 2001). Although the results were not very conclusive, track quality on relaying was identified as the most important factor. More measurements on track geometry were recorded and other main conclusions were drawn:

- After the first initial settlement, both vertical quality and alignment deteriorate linearly with tonnage (or time) between maintenance operations;
- The rate of deterioration varies drastically from section to section even for apparently identical sections carrying the same traffic;
- No statistical evidence proved a market effect on the quality or on the rate of deterioration by the type of traffic or track construction;
- The rate of deterioration appears to be a constant parameter for a section of track regardless of the quality achieved by the maintenance machine;
- In general, tamping machines improves the quality of a section of track to a more or less constant value.

Moreover, three main causes were identified for geometry deterioration: the random settlement of the ballast; the lack of straightness of the rails and the variation in the dynamic loads along the track caused by vehicles. Most significant dynamic loads could come from unsprung masses, which are responsible for wavelengths in the track of less than 1 m. Tamping cannot correct these short wavelengths, only grinding or corrective welds by straightening are appropriate.

Further research was conducted to confirm existing findings and to search for new information, so tests sections on different railways were monitored. These sections excluded switches and crossing in order to simplify the problem under analysis. Three aspects of track geometry development were reconfirmed:

- The large and fairly constant improvement due to successive tamping operations;
- The linear trend of quality deterioration over the period between the tamping operations;
- The lack of effect of tamping operations on the rate of deterioration.

However, it is important to highlight that certain sections have a higher rate of deterioration, which may be caused by singular features (rail bridges, level crossing), or local geometry faults present from the start of the operation, as well as sub-layers of inferior quality formation, or even welds of inferior quality.
Figure 3.4 illustrates an example taken from Modern Railway Track of a 10-year old section on FS which shows an important behaviour: the recurrence of faults arising from the presence of structures under the ballast.

![Figure 3.4 - Evolution of longitudinal level of an FS test section, Esveld (2001).](image)

Decisive parameters for maintenance decisions vary on the track speed. Tracks are usually classified in four categories, depending on train speed: High-speed tracks (V > 200 km/h); rapid-speed tracks (140 km/h < V < 200 km/h); medium-speed tracks (100 km/h < V < 140 km/h) and low-speed tracks (V < 100 km/h).

For low- and medium-speed tracks, the absolute defect values are the critical and determining parameters, as safety depends on them, whereas for medium-, rapid- and high-speed tracks, the decisive parameters are those determining passenger comfort, as at this speed ensuring a high level of passenger comfort implies ensuring traffic safety.

For these track defects at higher speeds, two limit values are specified: alarm values and upper values. Alarm values (L_{inf}) of track defects are values which, when reached, require programming of maintenance intervention, whereas upper values (L_{sup}) of track defects should not be reached, otherwise deterioration may become irreversible. Therefore, the decision for maintenance should be taken between the limits L_{inf} and L_{sup}. Another limit is the emergency value, which imposes immediate speed reduction until maintenance teams intervene.
Standard deviation of longitudinal, transverse and horizontal defects for a determined length of track is specified for different Infrastructure managers, depending on the track speed. Higher speeds require lower limit values than lower speeds.

As for medium- and low-speed tracks, maintenance decisions are taken based on the maximum values of defects recorded by the recording vehicle. Two limits are distinguished: intervention limits and acceptance limits. Intervention limits are values of track defects that, when reached need intervention and track maintenance, whereas acceptance limits are values of track defects that need to be above the left value after execution of track maintenance. Note that, as it is practically impossible to attain a geometrically perfect track, this limit acts as a control quality parameter of maintenance work.

In order to schedule maintenance works and accurately estimate costs involved, more knowledge of the progress of track defects and of the influencing factors is needed. Tests and statistical analysis has shown that a defect present in a track after maintenance progresses rapidly up to a critical traffic load of 2 million tons, beyond which defect progress is considerably slower.

The evolution of track defects is given by empirical semi-logarithmic expressions. The mean settlement of track is given by the following empirical formulas:

\[ m_e(T) = a_1 + a_0 \cdot \log \frac{T}{T_r} \]

In which: \( a_1 \) is the mean settlement for a traffic load \( T_r \) (with values between 5-15 mm); \( a_0 \) is the settlement increase rate (mm/decade), mainly depending on subgrade quality, with mean values 2-6 mm/decade and \( T_r = 2 \cdot 10^6 \) tons.

The standard deviation of longitudinal defects is of particular use for medium- and high-speed tracks, and may be estimated by the expression:

\[ sd_{LD}(T) = c_1 + c_0 \cdot \log \frac{T}{T_r} \]

In which: \( c_1 \) is the standard deviation of longitudinal defects for a traffic load, with mean values 1.0-1.35 mm; \( c_0 \) is the rate of increase for standard deviation of longitudinal defects as a function of traffic load, with mean values 0.1-0.2 mm/decade and \( T_r = 2 \cdot 10^6 \) tons (Profillidis, 2006).
By inverting the empirical expression above, an estimation of the interval between maintenance sessions (tamping operations) can be achieved based on a limit of traffic load $T_{\text{lim}}$:

$$T_{\text{lim}} = 2 \times 10^6 \times 10^{-\frac{\text{sd}_{L,D}^\text{lim} - c_1}{c_0}}$$

Since the parameter $c_0$ is almost constant, the determining factors for tonnage interval $T_{\text{lim}}$ between two successive maintenance sessions are the terms $\text{sd}_{L,D}^\text{lim}$ and $c_1$, the latter amounting to the track condition after maintenance. In the case of medium- and low-speed tracks, average values of longitudinal defects are used, instead of the standard deviations.

Transverse defects have a pattern of evolution very similar to the expression of standard deviation of longitudinal defects:

$$\text{sd}_{T,D}(T) = u_1 + u_0 \cdot \log \frac{T}{T_r}$$

In which: coefficients $u_1$ (mean value 1.2 mm) and $u_0$ (mean value 0.1-0.4 mm/decade) are defined similarly to $c_1$ and $c_0$ (Profillidis, 2006).

Horizontal defects follow an empirical law of evolution that may be also approximated by a semi-logarithmic formula of the traffic load. However, it has shown deviations and a large dispersion.

$$m_{H,D}(T) = d_1 + d_0 \cdot \log \frac{T}{T_r}$$

In which: coefficients $d_1$ and $d_0$ are defined as in the equation of the mean settlement, with mean values $d_1 = 0.6 - 1.0$ mm and $d_0 = 0.15 - 0.30$ mm/decade (Profillidis, 2006)

The evolution of gauge deviation is difficult to determine because it depends on various parameters, such as subgrade and rolling stock type.

The evolution of the standard deviation of track twist is also of semi-logarithmic form:

$$\text{sd}_{\text{twist}}(T) = g_1 + g_0 \cdot \log \frac{T}{T_r}$$

In which: coefficients $g_0$ (with mean values 1.0 - 2.0 mm/decade) and $g_1$ (with mean values 0.2 - 1.0 mm/decade) have a rather large dispersion and are defined similarly to the equations before (Profillidis, 2006).
As a conclusion of the evolution of track defects, their degradation follows a semi-logarithmic formula, whose main variable is the average traffic load. Note that the apparent disparity between the two analyses (linear trend versus semi-logarithmic) is just apparent. In fact, without including a quick settlement and a rapid deterioration of track immediately after tamping, the deterioration rate generally displays a linear trend between two maintenance operations, as the semi-logarithmic expression contemplates a linear phase, with an almost constant slope (deterioration rate).

3.2.2 - Degradation of rails

Wear and fatigue in rails are major contributors to rail deterioration. Wear is a form of degradation which cannot be completely eliminated, but can be reduced by lubrication. Also rail grinding is a recognized way of reducing RCF defects, as well as permitting corrugation removal.

Rail defects detected by the NDT cars (non-destructive testing) are then verified by hand-held ultrasonic equipment, and prioritization of defects is conducted so that corrective maintenance is performed on severe defects. Priority of defects may be based on several factors, such as track geometry, traffic type, traffic density, axle load, age of rails, defects history, rail material, curvature, yearly and total accumulated MGT. Low-priority defects are recommended for planned maintenance in the form of grinding, rail welding, rail section rectification/replacement, whereas high-priority defects are immediately recommended for unplanned maintenance in the form of rail welding or rail section replacement. For example, defects detected by the signalling system are generally severe and need immediate attention. Thus, unplanned maintenance is carried out to solve these defects.

Kumar, in his thesis (2008), discusses reliability analysis in railway infrastructure. He identifies the different factors influencing railway infrastructure, developing a framework for classification of degradation and failure data for reliability and cost modelling, while structuring a maintenance optimization model to arrive at an optimal maintenance policy for a component during its useful life. In fact, a great deal of literature is available on maintenance cost-effective models to improve track performance. However, these models lack a contribution to maintenance and design from a life-cycle cost perspective.

Degradation models are simulated separately. However, as Kumar (2006) points out in his approach to develop a framework to rail break prediction, maintenance aspects such as tamping and ballast cleaning may influence rail degradation. In fact, more generally it is admissible to argue that separated degradation models for a specific component are influenced
by the maintenance activities of other infra components. For example, in ballast renewal
sleepers may crack during this maintenance activity, shortening their life expectancy, and
therefore sleeper degradation models should include the impact of other components’
maintenance activities.

In fact, Kumar (2006) develops an approach to predict rail failures, which will help to
optimize maintenance activities (inspection, grinding, rectification/replacement and/or welding),
evaluating the trade-off and finding a balance between maintenance costs and consequential
costs due to derailment risk. Consequential costs may comprehend infrastructure unavailability
costs, traffic delay costs, penalties imposed by transport authorities, loss of assets,
environmental impact (e.g. transport of hazardous material) or even loss of human lives. Note
that the estimation of these consequential costs may depend on cultural aspects and
particularly how society evaluates the risk of accident and loss of human lives, which may be
quite controversial and highly subjective. However, not to consider the consequential costs
derived from this subjectivity is worse than dealing with it. Therefore, a requirement to an
effective rail maintenance procedure is the assessment of the rail failure rate.

3.3 - Maintenance and Renewal activities

Maintenance and renewal are main aspects that should be included in a life-cycle cost
analysis. Engineers usually assign a performance index that indicates the current performance
level of a structure. For instance, the present serviceability index (PSI) for pavement or the β-
index for bridges. Maintenance assures that the performance index does not drop below a
specified level. Note that condensing the performance of a structure in a unique index may be
rather simplistic and, in a way, it clusters a huge amount of explaining factors for performance
change, which may cause trouble in estimating performance index between major
rehabilitations.

Concerning Railways, track maintenance means the total process of maintenance and
renewal required to ensure that the track meets safety and quality standards. Ideally, this goal
should be achieved at minimum cost (Esveld, 2001). Note that minimizing maintenance and
renewal costs may not be equivalent to minimizing life-cycle costs, as the alternative considered
with lowest maintenance and renewal cost may have other relevant costs as construction costs,
for example. In fact, that might be an argument for defending that the construction of these
infrastructures should be delivered to consortiums, which are also held responsible for the first
period of operation and maintenance as the first 40 years, for example.

To some extent, the need for maintenance *per se* is rarely questioned since it naturally
follows from an investment, and it is mainly ‘captured’ in the investment decision (design
alternative). However, the uncertainty about the future and the fact that some events may
develop in another way than expected, justifies the implementation of a track maintenance
management system as an important tool to handle strategic decision making, even in the ideal
case of optimising life-cycle cost a-priori for the infrastructure.

On the role of maintenance, Profillidis (2006) argues that ‘with respect to safety,
maintenance should be preventive; regarding comfort, maintenance should be corrective; and
finally, as regards the economic aspects of maintenance, an optimum solution should be
sought, so as to ensure a satisfactory safety margin and prevent a quick degradation of track
quality’. In fact, efficient management of track maintenance depends largely on data acquisition
through manual or automatic inspection, but also on track deteriorating or degradation models
and the consequent planning of maintenance activities, taking into account the cost and
availability of resources. The next section will focus on the importance of inspection to
maintenance and renewal activities.

3.3.1 - Inspection

Measuring the system and controlling the quality of the infrastructure through
inspections is essential to ensure safety and quality standards. Inspection involves more than
just the inevitable measurement of defects for correction in maintenance works, as the
measurement of accelerations to guarantee passenger comfort, measurement of forces or even
inspection of rolling stock. However, the objective of inspection is not only to assure the non-
existence of any faults that might lead to accidents, but also to monitor successive degradation
in infrastructure in order to prevent faults and to provide the infrastructure manager with
information for short and long term planning of maintenance activities (Andersson, 2002).

Note that maintenance quality will greatly depend on the quality of available monitoring
systems of track inspection. Monitoring techniques are the base for providing more insight into
the infrastructure components’ behaviour over time (degradation models). Note that only
accurate inspection will guarantee the development of accurate degradation models, so
techniques used should be compatible with the expected accuracy. Therefore, as an ultimate
goal, Railway infrastructure monitoring should be processed in a fast enough and continuous
way to allow consecutive monitoring runs, and provide more accurate forecasting of track
condition and consequent maintenance planning.

Particular structures, such as tunnels and bridges may benefit from dedicated
monitoring systems. Tunnels benefit from laser, visual and thermal scanning systems such as
the TS 360 BP Scanner, whereas bridges, as one of the most critical elements, benefit from
bridge management systems. They provide continuous monitoring and allow more cost effective
maintenance, repair and rehabilitation actions, while ensuring safety standards at the same time. Visual inspection frequency varies depending on speed limit and daily train tonnage from a few times a week to once a month, though any exceptional circumstance (as very hot weather) may demand extra inspections.

Apart from tunnels and bridges, dedicated monitoring systems are lacking for substructure, whose monitoring is often underestimated. In fact, it significantly affects track maintenance costs, as its inaccessibility makes maintenance works focused on treating the symptoms rather than the causes, till deterioration rates become so high that it is no longer acceptable, and major corrective maintenance to substructure layers is necessary. Substructure conditions are an important input to split the track into maintenance sections. This splitting of the track into quasi-homogeneous sections should be based on a combined evaluation of geotechnical characteristics (thicknesses of layers and their condition) investigated by ground penetrating radar technique, track stiffness and its change over time estimated by measuring track deflection under wheel loadings.

In fact, this splitting process of track into quasi-homogeneous maintenance sections should be performed based on the infra data and deterioration rates of track components. Esveld (1990) argues that for optimum use of information and control of renewal and maintenance processes the deterioration data should be condensed in such a way. Therefore, infra data is understood as all the information on: type of rail, type of sleeper, type of fastenings, type of pad and renewal data, type of ballast, renewal and cleaning date, CWR or jointed, traffic type, speed spectrum, annual tonnage or even design geometry. As mentioned before, other factor contributing to track splitting into maintenance sections is the difference between deterioration rates. Therefore, due to geometry deterioration the track should be split into very short sections (100 to 200 meter-length), whereas for the purpose of rail failure prediction some relaxation is needed, producing maintenance sections of 5 to 10 kilometre-length.

Switches and crossings (S&C) should also have dedicated monitoring systems, as they sometimes represent a significant amount of the total M&R budget, around 25% according to Esveld. However, no general methodology for S&C management has ever been fully developed and defined, which may be blamed on the enormous diversity of technical solutions, especially if compared to the plain track case. This general methodology should comprehend the definition of critical elements that cause failures and the definition of key parameters that describe S&C condition independent of its technical condition. Thus, a condition-based maintenance management may be pursued, substituting the usual predefined fixed time interval management. Nevertheless, user-friendly software that bring together data on rail geometry through sophisticated measurement and inspection vehicles, like SwitchView, is currently being used to facilitate decision-making in S&C maintenance management.
The demand for inspections varies significantly depending on different drive factors, such as maximum train speed, traffic intensity, type of traffic, climate and environmental conditions, technical standards of the infrastructure components, age and quality, etc. Inspection management systems may divide the rail network into classified sections depending on inspection demand, prioritizing their need for inspection.

The scheduling of inspection activities is set at fixed time intervals, independently of the life phase of the component inspected and its degradation rate. Note that inspection cost increases when its frequency is higher (or the inspection interval is shorter). On the other hand, the risk or loss caused by failure will decrease when the inspection frequency is higher. Therefore, appropriate scheduling of inspection activities is important for a reduction in maintenance cost. In fact, some authors argue that inspection activities should be scheduled according to the estimated risk of failure. (Zhao et al, 2006). Since inspection usually requires a certain amount of human interaction and involves the necessary allocation of equipment, they are often quite expensive. Therefore, a large potential of savings in the long-run is still at reach to be explored, as optimizing the scheduling of inspection activities in order to minimize life-cycle cost may contribute to a more cost-effective approach towards maintenance.

Another important aspect to ensure cost-effectiveness of inspection activities is guaranteeing that there are no superfluous sensors to detect failures, depending on the type of failure progression. Four types of failure are summarized in the Figure 3.5, showing their respective progression over time. The philosophy inherent to this approach is considering that some failures are so sudden and have such a rapid progression that there should be no inspection system (unless a time-continuous one) to monitor degradation but rather failure.

Figure 3.5 – Types of failure progression, Esveld (2001).
Rails are one of the most important track components to ensure safety standards. Commonly, rail inspection consists of ultrasonic rail inspection, focused on detecting internal failures. Ultrasonic trains have replaced manual inspection, though switches and transitions for moveable bridges, junctions and sidings are normally inspected with ultrasonic hand equipment. These ultrasonic trains, like the UST 96, are equipped with a computer-controlled measuring system and an on-board data analysis system, and can be used on all standard gauge lines. Ultrasonic inspection is scheduled once a year for ordinary tracks, but it can be executed twice or even four times a year depending on daily tonnage and specificities of the track (S&C or transitions of moveable bridges).

However, ultrasonic rail inspection is usually included in track recording cars, combining other track quality measurements and overhead wire inspection. For example, the UFM 120 represents a universal measuring car designed for the track inspection of European normal gauge lines enabling, with a single run, the measurement of the relevant quality data of track geometry, overhead wire and rails as well.

Data collected from track recording cars will generate standard deviations, often per 200 m section in various wavebands. Track geometry measurements are gathered to calculate track defects, and serve as the basis to predict maintenance sessions. However, as standard deviations are hard to interpret for non-experts, quality indices are often built based on standard deviations. Later on, these aspects will be discussed in detail, though no quality indices will be built.

3.3.2 - Maintenance

Railways as a system that requires maintenance through its life-cycle is neither simple and obvious nor easy to anticipate. On the one hand, traffic process tends to increase track defects by the track-rolling stock interaction and therefore destabilize/degrade the system as a whole. On the other, maintenance process strives to reduce defects and restore the track to its previous good conditions. These two trends should be in equilibrium through rational scheduling of maintenance works while guaranteeing RAMS levels of performance and operation, optimizing mechanical equipment and assigning priorities of intervention, at minimum life-cycle cost. Note that the performance of a system, meaning RAMS level, may change due to variations on demand, and not only due to component degradation processes. Demand variations in Railway Infrastructure may be translated in the increase of axle load and traffic density, increasing the amount of accumulated tonnage for the same period of time, and therefore, decreasing time between maintenance activities, which may cause decrease of infrastructure availability.
Figure 3.6 - Schematic survey of maintenance and renewal process, Esveld (2001).

Figure 3.6 above gives a schematic summary of the various components that make up the maintenance and renewal process. In both maintenance and renewal works, manual and mechanical activities are included separately.

Some of the usual mechanized maintenance actions are:

- **Tamping** - correct level, cant and alignment – using tamping machines;
- **Ballast regulating** - establish correct ballast profile – using ballast regulators;
- **Ballast stabilizing** - compact ballast – using stabilizers;
- **Rail grinding** - remove corrugations and grind welds – using rail grinding machines;
- **Joint straightening** - straighten welds – using STRAIT method;
- **Ballast cleaning** - clean ballast bed – using ballast cleaners.

Track maintenance can be also divided into: rail geometry, track geometry, track structures, ballast bed, level crossings and miscellaneous. Maintenance of track geometry can in turn be subdivided into (incidental) spot maintenance, in other words, repair of local irregularities, and systematic maintenance (mechanized), which is carried out with heavy track maintenance machines.

Normally, spot maintenance is performed manually, supported by small machines, consisting of levelling and tamping using vibrating compactors or using jacks to raise the track and then filling the space with ballast material in order to rectify track level at crossings or other short sections of track.

Patra in his thesis (2007), tried to deal with the uncertainty related to maintenance life-cycle cost using the uncertainty related to RAM parameters. In fact, he argues that uncertainty
in Railway LCC estimation is due to the statistical characteristics of Reliability, Availability and Maintainability parameters (RAM). Therefore, Patra used a methodology combining Design of Experiment and the Monte Carlo simulation to assess the uncertainty linked with maintenance LCC. In fact, Patra distinguished two origins of uncertainty linked with LCC estimation. Level I uncertainty is due to the external risk of the LCC analysis, as penalties imposed by traffic operators on IM due to train delays, traffic disruption or derailment, whereas Level II uncertainty is due to the external risk associated with LCC, originated from uncertainty on RAM parameters. Nevertheless, note that RAM parameters also impact level I uncertainty as penalty regimes are based on RAMS performance.

As maintenance costs are the most complex cost component of the rail infrastructure, an accurate LCC estimation should focus on the uncertainty of Maintenance costs. Patra divided the maintenance cost modelling into three major components: preventive maintenance cost, renewal cost and corrective maintenance cost; and characterized each maintenance activity in condition-based or time-based maintenance triggering. Nevertheless, best practices in maintenance strategies try to change time-based maintenance to condition-based. Table 3.1 shows the maintenance activities considered in his analysis:

<table>
<thead>
<tr>
<th>Maintenance strategy</th>
<th>Maintenance activity</th>
<th>Maintenance trigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preventive maintenance</td>
<td>rail grinding</td>
<td>Time</td>
</tr>
<tr>
<td></td>
<td>Tamping</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>rail lubrication</td>
<td>time</td>
</tr>
<tr>
<td></td>
<td>ballast cleaning</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>track inspection</td>
<td>Time</td>
</tr>
<tr>
<td>Renewal</td>
<td>rail renewal</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>ballast renewal</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>sleeper renewal</td>
<td>Condition</td>
</tr>
<tr>
<td></td>
<td>fasteners renewal</td>
<td>Condition</td>
</tr>
<tr>
<td>Corrective maintenance</td>
<td>rail replacement</td>
<td>Failure</td>
</tr>
</tbody>
</table>

Table 3.1 – Track maintenance activities (Patra, 2007)

To assess the uncertainty in LCC, Patra assumed probability distributions for MTTF (mean time to failure) and for MTTR (mean time to repair) for high and low rails, and estimated the probability distribution of LCC performing a Monte Carlo simulation, creating random values for each RAM parameter and calculating the respective LCC for each pair of RAM parameters.

Although Patra focuses his efforts in RAM parameters uncertainty and its influence in LCC estimations, the further work (developed in chapter 4) will try to insert uncertainty analysis through the uncertainty related with degradation model parameters. Therefore, the present work
proposes to give a step backwards, introducing earlier uncertainty considerations in the track geometry degradation model, included in a ballast LCC model, in order to give two steps forward to understand the origin of uncertainty. Next chapter will put these ideas into practice.
Chapter 4

Practical example: Rail and Ballast renewals from a life-cycle cost perspective

This chapter discusses and exemplifies an integrated approach to support decision-making process on rail and ballast renewals from a life-cycle cost perspective. Therefore, an existing Rail LCC model is discussed and involving costs are discounted, whereas a Ballast LCC model is developed using a probabilistic approach with the Monte Carlo simulation involved. Integrating both components LCC models, optimal life-cycles are discussed for a 100-km plain track section example.

4.1 - Rail LCC model

The aim of this practical chapter is to develop a probabilistic approach to maintenance strategy from a life-cycle cost perspective. Therefore, a 100-km plain track section will be studied, analysing the costs associated with rail component and ballast bed component. Later on, these component costs will be integrated, so that renewal decisions are made based on a life-cycle perspective tool. In the next paragraphs, we will present the intricacies and the outputs of the Birmingham model to assess economic life of rail for the example proposed.

The Railway Research Centre at the University of Birmingham developed a model to evaluate the economic life of rail using a stochastic analysis of rail failures (Zhao et al, 2006). In the model, two types of defects are distinguished: the ATW (alumino-thermic welds) defects and type B defects (defects other than ATW defects). Typical type B defects are surface-initiated defects and internal defects. These defects are usually removed by replacing the short piece of rail where they occur, and then using alumino-thermic welding. Thus, two additional welds are introduced, increasing the probability of further weld defects. That is mainly the reason why ATW defects are predicted using a birth process with immigration. As to type B defects, they are assumed to follow a Weibull’s law, in which the hazard rate of a defect appearing in a section of rail at time \( t \) is given by the following expression:

\[
\text{hazard rate of defect at time } t = \left( \frac{\alpha}{\eta} \right) \left( \frac{t}{\eta} \right)^{\alpha - 1}
\]

In which: \( \alpha \) is the shape parameter and \( \eta \) is the scale parameter of the Weibull distribution.
For ATW defects, the hazard rate of a defect occurring within each ATW at time \( t \) \( (\lambda_a(t)) \) also follow a Weibull's law. Considering the birth process mentioned above, it can be shown that the hazard rate of an ATW defect within a section of rail at time \( (v_a(t)) \), is given by the expression:

\[
v_a(t) = \lambda_a(t) e^{\int_0^t \lambda_a(x)dx} \left[ n_0 + 2 \int_0^t v_b(r) e^{-\int_0^r \lambda_a(x)dx} dr \right]
\]

In which: \( n_0 \) is the number of ATWs in a section of rail at time \( t = 0 \) and \( v_b \) is the hazard rate of a type B defect. Note that \( v_b \) is the hazard rate of any type B defect, meaning that is given by the sum of the hazard rates \( v_j(t) \) of type B defects.

An important assumption of the model is that grinding is conducted periodically, and the consequent decreasing of the hazard rate of a rolling contact fatigue (RCF) defect. As grinding can significantly remove surface-initiated defects, especially RCF defects, the hazard rate should be rectified as the following expression shows:

\[
v_{RCF}(t) = \gamma(q) \cdot \lambda_{RCF}(t)
\]

In which: \( v_{RCF}(t) \) is the hazard rate of an RCF defect within a rail section at time \( t \) when grinding is being undertaken; \( \gamma(q) \) is the probability that a surface-initiated defect cannot be removed by grinding (which depends on the grinding rate \( q \), i.e. the removal depth of rail through grinding per 10 million gross tones) and \( \lambda_{RCF}(t) \) is the hazard rate of an RCF defect when no grinding is being undertaken.

A defect may evolve to a failure if no maintenance is performed or if it is not detected through inspection within a time period. The time duration from the initiation of a defect to a functional failure is referred to as P-F interval. In the model, inspections are assumed to be imperfect, meaning that not all defects are detected, but only a percentage of them. For a type defect \( j \), a fixed detection rate \( \beta_j \) (probability that a defect can be detected through inspection) is settled. Imperfect inspections are conducted at scheduled times \((t_1, t_2, ..., t_m)\).

Therefore, the failures arising within the rails follow a filtered non-homogeneous Poisson Process. As Zhao et al (2006) demonstrated the expected number of failures during the period \((0, T)\) in a section of rail \( N_f(T) \) is given by the expression:

\[
N_f(T) = \sum_{j=1}^b \sum_{i=1}^{m+1} \sum_{k=1}^i \left\{ (1 + \beta_j)^{j-k} \int_{t_{k-1}}^{t_k} v_j(r) \times [G_j(t_i - r) - G_j(t_{i-1} - r)] dr \right\}
\]
In which: \( h \) is the number of different defects (all type B defects plus ATW defects); \( \beta_j \) is the detection rate of \( j \)th type of defect; \( m \) is the number of inspections conducted in the period \((0, T)\) so that \( t_m + 1 = T \); \( \nu_j(\tau) \) is the hazard rate of \( j \)th type of defect and \( G_j \) is the cumulative distribution function of the P-F interval (typically a exponential distribution).

It is important to note that failures in the interval \((t_{i-1}, t_i)\) may be induced by defects occurring in any interval \((t_{k-1}, t_k)\), where \( 1 \leq k \leq i \). Moreover, note that for ATW failures, \( \nu_j(\tau) = \nu_a(\tau) \), calculated by the formula presented earlier.

Although the model allows analysis of both periodic and non-periodic inspection policies, only periodic inspection will be analyzed, where the instance of the \( i \)th inspection is given by \( t_i = i \cdot s_I \), with \( s_I \) being the interval of inspection.

The expected number of defects that have been detected by inspections during the period \((0, T)\) in a rail section \( N_d(T) \) is given by the expression:

\[
N_d(T) = \sum_{i=1}^{n} \left[ \int_0^T \nu_i(t) \, dt \right] - N_f(T)
\]

The economic life of a rail is analyzed considering major costs over a life-cycle of a rail section. Rail's life-cycle comprises the amount of traffic passing over the rail between two consecutive renewals. Therefore, life-cycle costs include inspection costs, repair costs, accident costs and renewal costs. There are mainly two types of rail inspection: ultrasonic inspection and geometric inspection. Ultrasonic inspection cost over the life time \( T \) of the rail is \( c_I T / s_I \), where \( c_I \) is the cost of inspection. In the model, it is assumed that geometric inspection is conducted periodically and that the related cost has no influence on the determination of rail economic life, and therefore, it is not included\(^8\). If defects are identified through inspections, they can be removed by planned maintenance, whose cost in the interval \((0, T)\) is \( c_d N_d(T) \), where \( c_d \) is the cost of repairing a defect. Concerning accident costs, it is assumed that the probability of occurring an accident (derailment) is very small. In fact, accidents may occur in the case of failures, but failures may be repaired through unplanned maintenance. The corresponding cost of these types of events is \((1 - \varepsilon) c_f + \varepsilon c_x \) \( N_f(T) \), where \( c_f \) is the cost of repairing a failure, \( c_x \) is the cost of a derailment and \( \varepsilon \) is the probability of an accident being caused by a rail failure. Another cost included is grinding cost. Grinding is a corrective maintenance activity, whose cost under a periodic grinding policy is \( c_g T / s_g \), where \( c_g \) is the cost of a grinding operation and \( s_g \) is the interval between grinding operations. Finally, rail is only renewed once during its life-cycle.

---

\(^8\) Later on, in the ballast LCC model, the geometric inspection cost will be included, so that an integrative approach remains valid.
and renewal cost is $c_R$. Therefore, the life-cycle cost (LCC) per unit traffic (MGT) is given by the following expression:

$$C(T) = \left( c_R + \frac{c_f T}{s_I} + \frac{c_g T}{s_g} + \left[ (1 + \varepsilon)c_f + \varepsilon c_g \right] N_f(T) + c_d N_d(T) \right) \frac{1}{T}$$

Note that economic life is assessed based on efficiency criteria, as the equation above balances inputs (LCC) and outputs (rail life-cycle - $T$). In fact, optimizing LCC should be understood as optimizing LCC per unit of traffic, and thus searching for a minimum of the function $C(T)$. However, the model does not take into account the depreciation of the value of money through time, meaning that no discount rate is considered in the evaluation of costs. Nevertheless, the costs considered above can be discounted to the base year, representing the present value of LCC per unit traffic (MGT) according to a constant discount rate ($r$) as the following expression shows:

$$C_o(T) = \frac{1}{T} \left[ \frac{c_R}{(1 + r)^{T / T_{year}}} + \sum_{n=0}^{N} \frac{T_n \cdot C(T_n) - T_{n-1} \cdot C(T_{n-1})}{(1 + r)^{T_n / T_{year}}} \right]$$

In which: $C_o(T)$ is the present value of $C(T)$; $T_{year}$ is the annual accumulated tonnage; $T_n$ is the cumulative tonnage at year $n$, thus $T_{-1} = 0$ and $T_N = 0$, and it may be calculated $T_n = n \cdot T_{year}$. In fact, $N$ should be given by $N = \left\lfloor \frac{T}{T_{year}} \right\rfloor$, where $\lfloor \cdot \rfloor$ rounds down the number inside.

A demonstrative example is considered in the paper for a 1-km section of rail with an annual tonnage of 10 MGT, in which parameter values for rail defects and coefficients of the model are set. For the studied example of 100-km plain track section, the model was implemented using macros programmed with Visual Basic in Excel. Note that some residual error of numerical integration may arise from the model implementation, and therefore a reasonable trade-off between computational time and error significance must be pursued. In Appendix 1, the functions programmed to run the model are presented. Having said that, we will assume the same values for the parameter values for rail defects as in the example of the paper, presented in Table 4.1 below:

<table>
<thead>
<tr>
<th>Defects</th>
<th>$\alpha$</th>
<th>$\eta$ (MGT)</th>
<th>P-F interval (MGT)</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATW defects</td>
<td>1.01</td>
<td>315.8</td>
<td>10</td>
<td>0.7</td>
</tr>
<tr>
<td>Flash butt weld defects</td>
<td>2.00</td>
<td>286.6</td>
<td>10</td>
<td>0.7</td>
</tr>
<tr>
<td>Squats defects</td>
<td>2.50</td>
<td>191.8</td>
<td>5</td>
<td>0.6</td>
</tr>
<tr>
<td>Tache ovale defects</td>
<td>2.17</td>
<td>182.3</td>
<td>7</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Table 4.1 - Parameter values for rail defects used in the example analyzed, Zhao et al (2006).
The coefficients of the model are also based on the values used by Zhao et al (2006) in the example referred, but the costs presented are in Euros (€), instead of British pounds (£).

Table 4.2 presents the values used for the example of a 100-km plain track section:

<table>
<thead>
<tr>
<th>$c_R$ (€/km)</th>
<th>$c_f$ (€/failure)</th>
<th>$c_d$ (€/defect)</th>
<th>$c_l$ (€/km)</th>
<th>$c_x$ (€/defect)</th>
<th>$c_g$ (€/km)</th>
<th>$n_0$</th>
<th>$\varepsilon$</th>
<th>$\gamma(q)$</th>
<th>$s_g$ (MGT)</th>
<th>$s_l$ (MGT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>160000</td>
<td>7000</td>
<td>1000</td>
<td>145</td>
<td>3944000</td>
<td>2700</td>
<td>22</td>
<td>0.000556</td>
<td>0.6</td>
<td>10</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Table 4.2 - Coefficients of the model used in the example analyzed, adapted from Zhao et al (2006).

Running the model, assuming the values presented above for parameter values for rail defects and coefficients, the life-cycle cost of rail for a 1-km section per MGT for different renewal tonnages is assessed as the Figure 4.1 below shows:

Although life-cycle costs have not been discounted yet, a minimum for life-cycle cost per MGT can be identified. In fact, it can be seen that the cost curve is steady at its minima, for $T = 340\ MGT$, with $C(340) = 1249\ €$ and $C(300) = C(400) = 1263\ €$, which only deviates around 1% from the minimal value. However, as mentioned before, costs should be discounted depending on the discount rate used and on the annual accumulated tonnage. The next two Figures (4.2 and 4.3) intend to give a perception of the sensibility of life-cycle costs respectively discounted to discount rate and annual accumulated tonnage:
Figure 4.2 - Sensibility of the present value of Life-cycle Cost of rail per MGT for a 1-km section to different discount rates (1%, 2%, 4% and 6%), for an annual accumulated tonnage of 12 MGT.

Figure 4.3 - Sensibility of the present value of Life-cycle Cost of rail per MGT for a 1-km section to different annual accumulated tonnages (12 MGT, 24 MGT, 36 MGT and 48 MGT), for a discount rate of 4%.

As illustrated by the figures above, life-cycle cost of rail per MGT ($C(T)$) is influenced by the discount rate and the annual accumulated tonnage. In fact, the present value of $C(T)$, defined earlier as $C_0(T)$, decreases as discount rate increases for the same values of accumulated tonnage (Figure 4.2); whereas $C_0(T)$ increases as annual accumulated tonnage increases for the same values of accumulated tonnage. Note that $C_0(T)$ coincides with $C(T)$ if
we assume a value for discount rate equals to zero, though this is not realistic as it implies that money has the same value through time.

Some functions presented above do not exhibit a minimum value in the interval 0-600 MGT analyzed. In fact, for our example we had assumed an annual accumulated tonnage of 12 MGT and a discount rate of 4%, and as seen it does not exhibit a minimum value in the interval 0-600 MGT (Figures 4.2 and 4.3 in red). Therefore, later on, when both rail and ballast cost models are integrated, we will face the problem of choosing when to renew rails. Later on, new assumptions will be drawn in order to overcome this situation.

4.2 - Ballast LCC model

As mentioned before, after the initial phase of large vertical settlements, especially marked by the consolidation of the ballast bed, track geometry degradation follows a linear trend. In fact, geometry degradation should be here understood as the increase of the standard deviation of longitudinal defects. Note that longitudinal defects are the track defects that reach respective alarm levels earlier after maintenance, and therefore, tamping operations are mainly predicted based on the evolution of the standard deviation of longitudinal defects. To quantify the economic life of ballast, as it is done for the rail, we should assume two types of maintenance operations: tamping and renewal. Note that periodic geometric inspection is performed so that track defects can be monitored and predicted maintenance activities scheduled. Although semi-logarithmic expressions may model the evolution of standard deviation of longitudinal defects in a very reliable way, some simplification will be introduced by ignoring the initial phase of large settlements (2 million tons), when estimating regression lines based on the recorded data given by geometric inspection.

The figure below illustrates recorded data and respective estimated regression lines for 10 successive maintenance sections of 200 meter each. Tamping operations’ periods are signalled and statistical measures (average and standard deviation) for the coefficients of regression lines are presented (initial quality measured and calculated rate of deterioration). Regression lines are estimated using the following linear relationship:

\[ \sigma = c_1 + c_0 T \]

In which: \( \sigma \) is the standard deviation of longitudinal defects; \( c_1 \) is the initial quality measured after renewal or tamping operations; \( c_0 \) is the rate of deterioration and \( T \) is the accumulated tonnage between tamping operations.
As presented in Figure 4.4 above, the average and the standard deviation of coefficients $c_1$ and $c_0$ vary depending on the number of tamping operation. In fact, the average and the standard deviation of $c_1$ increases through tamping operations, the average and the standard deviation of $c_0$ decreases through tamping operations. Although the former behaviour may be quite expectable as successive tamping operations become more inefficient, the latter is quite surprising, as many references point out that the rate of deterioration increases after consecutive tamping operations. Therefore, in the practical model developed, only the values from the first regression line will be used:

$c_1$: average = 0.50 mm, SD = 0.05 mm
$c_0$: average = 2.7 mm/100 MGT, SD = 0.95 mm/100 MGT

To perform the Monte Carlo simulation, these two coefficients should be defined as random variables and a joint probability distribution must be found. In fact, if a data set was available, fitting tests would be performed in order to find a suitable distribution. Moreover, other independence tests would be taken, in order to assume (as it will be later done) that these two random variables are independent, and therefore, their joint density function factorizes into the
product of their marginal density functions. This simplifies the problem, to a very large extent, and makes simulation easier as the random numbers for the coefficient $c_1$ do not have any connection with the random numbers for the coefficient $c_0$.

Although the sample dimension might be too small\(^9\) and no reference on fitting probability distributions of these coefficients has been found, we will assume that they are independent and follow a normal distribution with parameters estimated by the given statistical measures listed above. Note that the real distribution of these coefficients might not be normal and the statistical measures might not be suitable estimations of the distribution parameters. Moreover, the assumption that these two coefficients are independent might not be true. Nevertheless, some assumptions must be made and no reference to refute these assumptions has been found. Furthermore, as many factors influence track degradation and assuming that track degradation is a sum of many independent factors (intended as independent random variables), the Central Limit Theorem might be a strong inspiration to assume that track degradation (meaning the standard deviation of longitudinal defects) follows approximately a normal distribution.

In conclusion, we will assume that:

\[
\begin{align*}
  c_1 & \sim N(\mu = 0.50 \text{ mm}; \sigma = 0.05 \text{ mm}) \\
  c_0 & \sim N(\mu = 2.7 \text{ mm}/100 \text{ MGT}; \sigma = 0.95 \text{ mm}/100 \text{ MGT})
\end{align*}
\]

$c_1, c_0$ are independent

Tamping operations are scheduled when the standard deviation of longitudinal defects reach a specified target limit. This limit standard deviation ($\sigma_{lim}$) will be assumed to be equal to 1.0 mm. Therefore, a prediction of accumulated tonnage between tamping operations can be calculated based on the following expression:

\[
T = \frac{\sigma_{lim} - c_1}{c_0}
\]

Moreover, in order to include the loss of effectiveness of consecutive tamping operations, we will assume that $c_1$ will increase at a fixed rate $r_1$. Concerning the behaviour of the deterioration rate ($c_0$) of consecutive tamping operations, we will assume that $c_0$ will also increase at a fixed rate $r_0$. Note that both assumptions contribute to diminish the accumulated tonnage between tamping operations over time. Nevertheless, the evolution over successive tamping operations of the rate of deterioration ($c_0$) may vary from a maintenance section to another. Therefore, the rate $r_0$ will be modelled as a random variable, following a normal distribution.

\(^9\) Note that, in fact, no data sample is literally available, but instead only statistical values of a data sample are available (average and standard deviation). Therefore, no statistical tests can be performed.
distribution and independent from \( c_0 \) and \( c_1 \), representing the fixed rate increase of the rate of deterioration \( c_0 \) for a given maintenance section. Note that \( r_0 \) is considered a fixed rate because it does not change over time for the same maintenance section, whereas \( r_1 \), as a fixed rate, assumes the same value for all maintenance sections over time. Therefore, we will assume that:

\[
\begin{align*}
  r_0 &\sim N(\mu = 0.04; \sigma = 0.02) \\
  r_1 & = 0.01
\end{align*}
\]

In order to quantify the tamping cost, tamping scheduling must be known. Note that the assumptions mentioned above must be included in the group of assumptions made before, so that a series of accumulated tonnage for each maintenance section must be found. In order to quantify the uncertainty related with tamping scheduling, we will perform the Monte Carlo Simulation for 100 km of plain track (500 maintenance sections of 200 m). Note that IMs benefit from the fact that they own a high number of maintenance sections, as the probability of every maintenance section having a high deterioration rate is very small. Therefore, the series of accumulated tonnage for each maintenance section between tamping operations will be given by the following formula:

\[
T_i = \frac{\sigma_{lim}}{c_{0i}} - c_{1i} + 2\text{MGT}; \quad i = 1, 2, ...
\]

\[
\begin{align*}
  c_{1i} &= c_1 \cdot (1 + r_1)^{i-1} \\
  c_{0i} &= c_0 \cdot (1 + r_0)^{i-1}
\end{align*}
\]

Note that, as the first expression shows, 2 MGT is summed in each accumulated tonnage because the first phase of large settlements had been ignored, and the estimations of \( c_0 \) and \( c_1 \) are for the regression lines, also ignoring the first 2 MGT of accumulated tonnage. Having the series \( \{T_i; i = 1, 2, \ldots\} \) for each maintenance section, we can assess the tamping costs:

\[
LCC_{\text{tamping}} = \sum_{i=1}^{N} \frac{c_{\text{tamping}}}{(1 + r)(T_i^{\text{acum}}/T_{\text{year}})}
\]

In which: \( LCC_{\text{tamping}} \) is the life-cycle cost for tamping activities; \( c_{\text{tamping}} \) is the cost of tamping a 200-meter section; \( r \) is the discount rate; \( T_{\text{year}} \) is the annual accumulated tonnage and \( T_i^{\text{acum}} \) is the accumulated tonnage till the \( i \)th tamping operation. In fact, \( T_i^{\text{acum}} \) may be calculated by \( T_i^{\text{acum}} = \sum_{j=1}^{i} T_j \). Note that \( \lfloor \rfloor \) rounds down the number inside, and therefore, costs are assumed to be discounted at the beginning of each year. \( N \) is the last tamping operation before ballast renewal.
Ballast renewal will be necessary when tamping operations become too demanding, meaning that they interfere to a great extent with the availability of the infrastructure, becoming traffic disruptive. Later on, we will include unavailability costs to search for an optimal renewal strategy. Nevertheless, at this stage we will determine $N$ indirectly, by considering that above a certain percentage ($\alpha$) of maintenance sections with an interval between tamping operations smaller than a certain amount of accumulated tonnage ($\beta$), tamping operations are no longer suitable, and the ballast bed must be renewed. Note that $N$ will vary on each section, representing the number of tamping operations performed, till $T_{\text{cum}}^k$ equals $T_{\text{renewal}}$ (accumulated tonnage till renewal). Moreover, an estimation of $c_{\text{tamping}}$ is needed. Table 4.3 shows examples of average costs for tamping operations:

<table>
<thead>
<tr>
<th></th>
<th>Average tamping cost (€/km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High speed line</td>
<td></td>
</tr>
<tr>
<td>Paris – Lyon</td>
<td>6142.14</td>
</tr>
<tr>
<td>Madrid – Seville</td>
<td>2699.47</td>
</tr>
<tr>
<td>French conventional lines</td>
<td>3549.04</td>
</tr>
</tbody>
</table>

Table 1.3 - Example of tamping costs for different lines, CENIT (2008).

As the costs presented are per kilometer, we must divide them by five to have an estimation of the tamping cost for a maintenance section (200 meter-length). Therefore, $c_{\text{tamping}}$ will be assumed to be equal to 600 € (3000 €/5).

Different values for $\alpha$ and $\beta$ were tried, resulting in different values for $T_{\text{renewal}}$ and $\text{LCC}_{\text{tamping}}$ for each maintenance section. It is important to mention that this technique to gather a collection of values for $T_{\text{renewal}}$ and $\text{LCC}_{\text{tamping}}$ intends to simulate in a realistic way the IM’s decision-making process of ballast renewal based on parameters like $\alpha$ and $\beta$. Therefore, summing all different $\text{LCC}_{\text{tamping}}$ for the distinctive 500 maintenance sections, we will reach an estimation of the total tamping cost for the 100 km analyzed. Moreover, by simulating a great number of times, we will reach a collection of values for total tamping cost and respective $T_{\text{renewal}}$ and we may quantify the uncertainty associated with this cost. However, as argued for the rail model, a trade-off between life-cycle cost (LCC) and accumulated tonnage must be searched, and therefore we must look to the efficiency of LCC and a new function $T_{\text{C}_{\text{tamping}}}$ should be defined:

$$T_{\text{LCC}_{\text{tamping}}} = \frac{\sum_{k=1}^{500} \text{LCC}_{\text{tamping}} \cdot k}{T_{\text{renewal}}}$$

In which: $T_{\text{LCC}_{\text{tamping}}}$ is the total life-cycle cost of tamping operations for the 100 km of plain track per MGT, $\text{LCC}_{\text{tamping}} \cdot k$ is the life-cycle cost of tamping operations for the $k$th maintenance section and $T_{\text{renewal}}$ is the respective accumulated tonnage till ballast renewal.
Figure 4.5 below illustrates the life-cycle cost of tamping operations per MGT for different renewal tonnages. In fact, it shows average values and extreme values for the 5\textsuperscript{th}, 10\textsuperscript{th}, 90\textsuperscript{th} and 95\textsuperscript{th} percentiles. Moreover, the respective quadratic trend lines are adjusted to presented values. Note that a collection of more than 100,000 pairs of values for $T_{\text{tamping}}$ and $T_{\text{renewal}}$ was gathered as the Monte Carlo simulation requires, in order to quantify the uncertainty related to life-cycle cost of tamping operations per MGT.

![Figure 4.5 - Total Life-Cycle Cost of Tamping operations per MGT of a 100 km plain track section ($TLCC_{\text{tamping}}$).](image)

Note that the costs presented above do not include renewal cost neither inspection cost. In fact, ultrasonic inspection cost was included in the model for assessing rail economic life explained earlier. Nevertheless, geometric inspection cost was excluded at that point, as rail defects and failures are mostly detected by ultrasonic inspection, and it would later on be included when assessing the economic life of ballast. Therefore, ‘the time has come’. We will assume that geometric inspection is performed periodically in every trimester. The following expression quantifies the life-cycle cost of geometric inspection per maintenance section ($LCC_{\text{geom,insp}}$):

$$
LCC_{\text{geom,insp}} = \sum_{j=1}^{N} \frac{C_{\text{geom,insp}} \cdot n_{\text{insp}}}{(1 + r)^j}
$$
In which: $c_{geom\_insp}$ is the cost of geometric inspection per inspection and maintenance section; $n_{insp}$ is the number of geometric inspections per year ($n_{insp} = 4$); $r$ is the discount rate and $N$ is the year when ballast renewal takes place. Note that $N$ may be calculated by $N = \frac{T_{renewal}}{T_{year}}$.

To quantify this cost for the analyzed example of 100 km of plain track, we will assume that the cost of each geometric inspection per kilometer is 100 €. Therefore, $c_{geom\_insp}$ is assumed to be equal to 20 €. Note that this assumption might be controversial as inspection cost may be difficult to estimate, involving equipment amortizations (if IM uses its own equipment), labor cost associated and therefore, $c_{geom\_insp}$ may vary with $n_{insp}$.

Concerning renewal cost, it represents the large expenditure throughout the life-cycle that only takes place at the end of it, and should therefore be discounted based on $T_{renewal}$. The following expression gives the life-cycle cost of renewal per maintenance section ($LCC_{renewal}$):

$$LCC_{renewal} = \frac{c_{renewal}}{(1 + r)^{\frac{T_{renewal}}{T_{year}}}}$$

In which: $c_{renewal}$ is the renewal cost per maintenance section. $T_{renewal}$, $T_{year}$ and $r$ have the same meaning from the equations listed above.

The ballast renewal cost is assumed to be equal to 200 000 €/km. Therefore, $c_{renewal}$ will be taken as equal to 40 000 €. It is important to refer that this estimation for renewal cost takes implicitly into account the traffic interference provoked by ballast renewal operation which, in fact, depends largely on the track layout and its level of redundancy.

Note that to quantify renewal and geometric inspection cost for the analyzed example of 100 kilometers (500 maintenance sections of 200 meters), we have to multiply $LCC_{geom\_insp}$ and $LCC_{renewal}$ by 500. Note that we assume that geometric inspection is performed at the same time for each maintenance section and that renewal is also performed for all the maintenance sections when the ballast life-cycle ends.

Another life-cycle cost that we will include in the model is the cost of unavailability of the infrastructure. As mentioned before, most temporary maintenance operations (tamping operations) are performed during non-operative time (maintenance night shifts), and therefore considered non-disruptive. However, as tamping operations become more demanding they will interfere with the normal operation of infrastructure, causing delays and even train cancellations. In fact, performance payment regimes contracted between the regulator entity and the IM determine the penalties for bad performance on availability. These regimes have an exponential behaviour, meaning that as unavailability increases, penalties become more
severe. Therefore, quantifying unavailability costs may be quite complex and demand more information on the robustness of the operation schedule. However, in order to quantify total life-cycle cost we will assume a function of unavailability costs with exponential behaviour depending on $T_{\text{renewal}}$. Note that this assumption might be quite controversial, but no reference has been found to quantify these costs. Moreover, it is important to mention that more studying is needed on these aspects and therefore, they represent a challenging point for further research.

Having said that, life-cycle unavailability cost of infrastructure per MGT for the example of 100-km plain track may be estimated by the following expression:

$$TLCC_{\text{unavailability}} = \begin{cases} 
0, & T_{\text{renewal}} < T_{\text{disruptive}} \\
300 \times e^{0.02T_{\text{renewal}}}, & T_{\text{renewal}} \geq T_{\text{disruptive}}
\end{cases}$$

In which: $TLCC_{\text{unavailability}}$ is the total life-cycle unavailability cost per MGT for 100-km of plain track (€); $T_{\text{renewal}}$ is the accumulated tonnage till renewal (MGT) and $T_{\text{disruptive}}$ is the accumulated tonnage limit above which tamping operations become traffic disruptive. Note that this limit depends on multiple factors, such as: maintenance equipment, human resources teams, IM internal organization, the number of maintenance bases, the extent of track analyzed and of course capacity (number of slots). Nevertheless, we will assume that $T_{\text{disruptive}} = 135 \ MGT$.

The expression above does not take into account the discount rate. Therefore, we will assume that this expression only remains valid for this example for a constant discount rate $r = 4\%$. Note that this discount rate takes the same value used to discount the different types of costs presented above.

Having presented the costs associated with economic life of the ballast bed, we may systematize the total life-cycle cost per MGT of the ballast bed for the example of 100-km plain track by the following expression:

$$TLCC_{\text{ballast}} = TLCC_{\text{tamping}} + TLCC_{\text{renewal}} + TLCC_{\text{geom.insp}} + TLCC_{\text{unavailability}} =$$

$$= TLCC_{\text{tamping}} + 500 \times \frac{LCC_{\text{renewal}}}{T_{\text{renewal}}} + 500 \times \frac{LCC_{\text{geom.insp}}}{T_{\text{renewal}}} + TLCC_{\text{unavailability}}$$

Note that once again, as we are trying to optimize the efficiency of total life-cycle cost, we want to minimize total life-cycle cost per MGT, and therefore total life-cycle costs presented are referred to costs per MGT. Moreover, it is important to refer that each component of presented cost is calculated by the expressions shown above, for a constant discount rate...
\( r = 4 \% \). Therefore, the figure below shows the total life-cycle cost per MGT of the analyzed example:

![Graph showing total life-cycle cost per MGT](image)

**Figure 4.6 - Total Life-cycle Cost of ballast per MGT of a 100-km plain track section (TLCC<sub>ballast</sub>).**

As Figure 4.6 above illustrates the major components that contribute to total life-cycle cost of the ballast bed per MGT are the renewal cost and the unavailability cost. This may compromise the present research as unavailability costs need further investigation, though total life-cycle cost of ballast per MGT shows a minimum between 150 to 250 MGT, which is the usual interval referred to ballast renewal. In fact, the minimum value for total life-cycle cost of ballast per MGT is 85 573 € for an accumulated tonnage of 215 MGT. Note that total life-cycle cost of ballast per MGT without unavailability cost is also presented.

Moreover, it is important to refer that the cost of tamping operations and geometric inspection do not have a major impact on total life-cycle cost. That is mainly the reason why Monte Carlo simulation results to quantify the uncertainty associated with the tamping cost is not presented here, as it does not have any major impact on total life-cycle cost, and therefore, not affecting the maintenance strategy of ballast renewal. Nevertheless, it should be highlighted that Monte Carlo simulation performed to calculate tamping operations costs would give relevant data to estimate unavailability costs, but again this is left out of analysis for further research. Note that the scheduling of tamping activities, given by \( \{T_i; i = 1, 2, \ldots \} \) for each maintenance section would provide data to investigate on how demanding infra availability...
tamping operations become as the ballast renewal is postponed. Therefore, the argument above gives evidence that uncertainty in ballast LCC estimations may be due to unavailability costs, rather than tamping costs, and therefore further analysis should focus on the variability of RAM parameters rather than on the variability of degradation model parameters. Nevertheless, they are related and further research is needed to link them.

4.3 - Integrating rail and ballast LCC models

Having described both models to assess economic life of rail and ballast separately, an integrative maintenance strategy approach should comprehend the outputs of both models and determine best renewal times (quantified by accumulated tonnage) for each component or both components in case of simultaneous renewal. As no model was developed to assess economic life of fastenings and sleepers, only the life-cycle cost models of these two components are integrated. Nevertheless, that might not compromise the efficiency of the maintenance decisions as sleepers and fastenings have an expected life-cycle larger than rail and ballast, though a degradation model of one component may influence the other. In fact, note that ballast renewal may reduce sleepers’ life-cycle, producing cracks or even breaking them, whereas in fastenings, irreversible defects may be developed.

A straight-forward approach to integrate both cost models may be based on the construction of hypothetical scenarios for distinct renewal times for each component, while quantifying total life-cycle costs per MGT for each scenario, as total life-cycle cost functions for rail and ballast are given by separate LCC models.

As seen in the ballast model, the 100-km plain track section has an optimal life-cycle of 215 MGT with a corresponding total life-cycle cost per MGT of 85 573 €, whereas in the rail model no optimal life-cycle was identified for the present value of life-cycle cost with a discount rate set at 4% and an annual accumulated tonnage set at 12 MGT. Note that in the ballast cost model, the same values for discount rate and annual accumulated tonnage were assumed, as the integrative cost model should sum values discounted at the same discount rate. Nevertheless, in order to search for scenarios to exemplify the integrative approach, we will assume that the discount rate is lower for the rail model. Obviously, this is wrong as both separate models should contemplate costs discounted at the same discount rate, though for practical reasons, to exemplify how to do this integrative approach, we will assume that the discount rate is equal to 1 % and the annual accumulated tonnage is equal to 12 MGT. Therefore, the total life-cycle cost function of rail per MGT is given by figure 11. Note that total life-cycle cost function of rail per MGT is for a 1-km plain track, and therefore it should be multiplied by 100 for the analyzed example of a 100-km plain track. This function has a
minimum of 420 MGT with a corresponding total life-cycle cost per MGT of 992.1 €. As this value is per km, for the 100-km plain track section, minimum total life-cycle cost of rail per MGT is 99 210 €. Figure 4.7 shows total life-cycle costs per MGT functions for rail and ballast components for the 100-km plain track section example.

As the minimal values for total life-cycle cost per MGT of rail and ballast are very similar, it is not obvious which renewal strategy optimizes life-cycle of components, while minimizing total life-cycle cost per MGT of infrastructure. Once again, note that no costs associated with fastenings and sleepers are considered, though it is believed that this fact would not compromise the efficiency of maintenance strategy proposed. Therefore, twenty-five distinct scenarios were built with different renewal times for ballast and rail. As the rail optimal life-cycle fits with approximately two optimal ballast life-cycles, only scenarios with simultaneous rail renewal and ballast second renewal are investigated \( T_{ballast\ 1} + T_{ballast\ 2} = T_{rail} \). Therefore, the best scenario is that which presents the minimum total life-cycle cost per MGT, calculated by the following expression:

\[
TLCC_{ballast\ +\ rail} = \frac{T_{ballast\ 1} \times TLCC_{ballast\ 1}(T_{ballast\ 1}) + T_{ballast\ 2} \times TLCC_{ballast\ 2}(T_{ballast\ 2})}{T_{rail}} + T_{rail} \times TLCC_{rail}(T_{rail})
\]

In which: \( TLCC_{ballast\ +\ rail} \) is the total life-cycle cost per MGT of ballast and rail; \( T_{ballast\ 1} \) and \( T_{ballast\ 2} \) are respectively the accumulated tonnages at first and second ballast renewals.

\(^{10}\) Note that accumulated tonnages are counted since the start of infrastructure operation or the last time the component was renewed. Therefore, the accumulated tonnage at second ballast renewal corresponds to the accumulated tonnage since first ballast renewal.
$T_{rail}$ is the accumulated tonnage at rail renewal; $r$ is the discount rate and $T_{year}$ is the annual accumulated tonnage. For further calculations they will assume the values: $r = 4\%$ and $T_{year} = 12\,MGT$. $TLCC_{ballast}$ and $TLCC_{rail}$ are given by the functions in figure 4.7.

Note that in the equation above, the total life-cycle cost per MGT of the second ballast should be discounted to the base year (beginning of operation) at a constant discount rate ($r$), given an annual accumulated tonnage ($T_{year}$). Once again, optimal life-cycle is obtained by the efficiency of total life-cycle cost, and therefore, total life-cycle costs per MGT for the 100-km plain track section example are calculated.

In the next page, Table 4.4 illustrates the twenty-five scenarios that were analyzed with the respective total life-cycle costs per MGT. The best scenario is marked in green, scenario 9, corresponding to a total life-cycle cost per MGT (ballast + rail) of 160 466 €, which consists of a first ballast renewal at an accumulated tonnage of 195 MGT, and a simultaneous rail renewal and second ballast renewal at a rail accumulated tonnage of 430 MGT$^{11}$. Therefore, note that though ballast optimal life-cycle was set at 215 MGT in the separate model analysis, none of ballast renewals is set at optimal life-cycle (195 and 235 MGT). This happens because of the additional discount that total life-cycle cost per MGT function of ballast (for the second ballast renewal) suffers, related to the given shape of that function, presenting a steeper slope on the right of the minimum value.

Nevertheless, note that this simplistic approach to integrate rail and ballast LCC models, through the construction of hypothetical scenarios, may be improved by considering gains of efficiency for simultaneous renewal, obtained as renewal activities are clustered. In fact, future research should contemplate these efficiency gains by including them as negative costs in an integrative LCC model.

---

$^{11}$ Note that a rail accumulated tonnage of 430 MGT corresponds in this case to a second ballast accumulated tonnage of 235 MGT (≈ 430-195).
<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Renewal accumulated tonnages (MGT)</th>
<th>TLCC per MGT (€/MGT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{ballast\ 1}$</td>
<td>$T_{ballast\ 2}$</td>
</tr>
<tr>
<td>1</td>
<td>215</td>
<td>205</td>
</tr>
<tr>
<td>2</td>
<td>205</td>
<td>215</td>
</tr>
<tr>
<td>3</td>
<td>195</td>
<td>225</td>
</tr>
<tr>
<td>4</td>
<td>185</td>
<td>235</td>
</tr>
<tr>
<td>5</td>
<td>175</td>
<td>245</td>
</tr>
<tr>
<td>6</td>
<td>225</td>
<td>205</td>
</tr>
<tr>
<td>7</td>
<td>215</td>
<td>215</td>
</tr>
<tr>
<td>8</td>
<td>205</td>
<td>225</td>
</tr>
<tr>
<td>9</td>
<td>195</td>
<td>235</td>
</tr>
<tr>
<td>10</td>
<td>185</td>
<td>245</td>
</tr>
<tr>
<td>11</td>
<td>225</td>
<td>215</td>
</tr>
<tr>
<td>12</td>
<td>215</td>
<td>225</td>
</tr>
<tr>
<td>13</td>
<td>205</td>
<td>235</td>
</tr>
<tr>
<td>14</td>
<td>195</td>
<td>245</td>
</tr>
<tr>
<td>15</td>
<td>185</td>
<td>255</td>
</tr>
<tr>
<td>16</td>
<td>225</td>
<td>225</td>
</tr>
<tr>
<td>17</td>
<td>215</td>
<td>235</td>
</tr>
<tr>
<td>18</td>
<td>205</td>
<td>245</td>
</tr>
<tr>
<td>19</td>
<td>195</td>
<td>255</td>
</tr>
<tr>
<td>20</td>
<td>185</td>
<td>265</td>
</tr>
<tr>
<td>21</td>
<td>235</td>
<td>225</td>
</tr>
<tr>
<td>22</td>
<td>225</td>
<td>235</td>
</tr>
<tr>
<td>23</td>
<td>215</td>
<td>245</td>
</tr>
<tr>
<td>24</td>
<td>205</td>
<td>255</td>
</tr>
<tr>
<td>25</td>
<td>195</td>
<td>265</td>
</tr>
</tbody>
</table>

Table 4.4 – Renewal scenarios for ballast and rail with respective total life-cycle cost per MGT (ballast + rail).
Chapter 5

Conclusions and Future Research

This chapter discusses the findings of the present research and some contributions for future research.

5.1 - Conclusions

The aim of this research work was to develop an approach that could support decision-making on maintenance and renewal in a Railway infrastructure, especially evaluating optimal ballast and rail renewals from a life-cycle cost (LCC) perspective. Some results obtained by the analysis of the state of the art drawn from the previous chapters and by the exploratory approach in chapter 4 give the evidence needed to come to important conclusions and suggest fruitful ideas for further research.

Therefore, the present research begins with a brief description of the Railway sector new challenges, faced as vertical separation is effectively implemented in Europe, where LCC plays a primordial role towards a more conscious and transparent management of Railway infrastructure. In fact, LCC may contribute decisively not only to support the management of existing infrastructures, but also to support investment decisions for new high-speed tracks in Europe and around the world. Nevertheless, note that this work does not intend to give an answer on how to evaluate the economic feasibility of a project/investment, as only costs are evaluated, and not revenues nor their respective uncertainty; instead this study intends to contribute with some ideas to improve analysis processes inside LCC for Railway infrastructure inside IM organizations.

Furthermore, Life-cycle Costing (LCC) analysis integrated with Life-cycle Assessment (LCA) may contribute to a more sustainable decision-making in Railways, taking into account long-term costs and environmental consequences. Although European transport strategy emphasizes the reduction of GHG emissions on the increase of energy efficiency in road transport, through technological improvements in vehicle and fuels, revitalizing Railways is also a way towards a more sustainable European transport sector, contributing to a fairer modal distribution.

Zoeteman model serves in this work to exemplify the complexity and the different variables that need to be taken into account to build an LCC decision support system. In fact, it permitted to focus this research on degradation models, trying to assessing the uncertainty related to LCC by assessing the uncertainty of degradation model parameters. Therefore,
Monte Carlo simulation was performed to assess life-cycle tamping cost uncertainty. It showed that compared to other costs, tamping uncertainty does not contribute decisively to ballast LCC uncertainty, and renewal decisions are mainly conducted by unavailability costs. Therefore, more emphasis on RAM (Reliability, Availability and Maintainability) parameters and assessment of unavailability costs is needed. In fact, a main conclusion of this thesis is the fact that ballast LCC uncertainty in Railway infrastructure is potentially due to unavailability costs and consequently due to RAM parameters uncertainty.

Furthermore, track geometry degradation models showed that a linear trend of quality deterioration over the period between tamping operations is a good estimation of real track geometry deterioration, validating the formula to estimate the series of accumulated tonnages for tamping operations used in the ballast LCC model.

Concerning rail renewal, the rail LCC model showed that there is a wider flexibility on rail renewal than on ballast renewal, as the rail cost curve is steadier at its minima. Moreover, as no uncertainty considerations were made in rail LCC ballast, the Monte Carlo simulation was not used there. In fact, the existing rail model used (Zhao et al, 2006) needed some adaptations to perform Monte Carlo simulation that were not developed. Concerning ballast renewal, the ballast LCC model assessed only the uncertainty related to tamping costs, which proved to be negligible compared to the potential uncertainty related to unavailability costs. The ballast LCC model developed includes four component costs: tamping cost, renewal cost, geometric inspection cost and unavailability cost. Tamping and geometric inspection costs had lesser impact on the optimal life-cycle of ballast than renewal and unavailability costs.

Integrating rail and ballast LCC models was conducted through a simplistic approach based on the construction of hypothetical scenarios. It showed that the optimal renewal strategy, for the analyzed example for a 100-km plain track, had a first ballast renewal at 195 MGT and a second ballast renewal at 235 MGT, occurring simultaneously with a rail renewal at 430 MGT.

5.2 - Limitations

One of the main limitations of this work is the fact that this research is focused on the degradation model of rail and track geometry instead of a railway track as a whole, because different components such as switches, fastenings, sleepers and subgrade will affect track degradation. The reason for this limitation is the vastness and complexity of the research area. Another important limitation is the lack of integrated data analysis, and the use of different sources of data for rail and ballast LCC models. Note that IM still lack comprehensive databases on track quality and infra data relating infra component degradation through time, so that degradation model parameters uncertainty may be estimated by assuming that they are...
random variables that follow proper probability distributions given by the historic data. Apart from that, another limitation is the fact that unavailability costs are not properly justified using a performance payment regime, and in fact an expression for unavailability costs is assumed so that total life-cycle costs of ballast per MGT has a minimum value in the interval 150-250 MGT. Moreover, it is important to remind the reader that different discount rates were also used for ballast and rail LCC models so that the integrating approach of separate LCC models could be exemplified. Nevertheless, proper integration should be done by using the same discount rate in both separate models in future research.

Another serious limitation of this work is the fact that the Monte Carlo simulation was only performed for tamping cost of LCC ballast model. In fact, potential uncertainty related to unavailability costs is considered the most relevant to renewal decisions and to assess ballast LCC model uncertainty, as tamping operations do not represent a major impact on total life-cycle cost per MGT.

Moreover, when integrating rail and ballast separate LCC models, no considerations on any efficiency gain to cluster renewal activities has been made, and in fact, clustering of maintenance and renewal activities may bring some reduction of life-cycle costs, and in the future research they may be included as negative costs in an integrative LCC model.

5.3 - Future Research

More work on how fruitful the idea of inserting uncertainty considerations in degradation model parameters, linking uncertainty generated from it to maintenance operations scheduling uncertainty and RAM parameters uncertainty, and availability costs uncertainty based on a performance payment regime is needed.

Moreover, although EU member states railway sectors may be at different stages of liberalization and even internal transport policies may differ significantly, an EU technical bulletin to serve as a guide to LCC analysis for European infrastructure managers needs to be built in the long term, similar to the US technical bulletin to perform LCC analysis in pavement design in Roadway transport. This guideline document should assess Railway infrastructure LCC taking into account relevant uncertainty considerations, and also best practices in renewal decisions increasing efficiency in Infrastructure management and also improving investment (design) decisions.

Railway infrastructure as a complex system, in which many actors are involved, needs more knowledge on how to evaluate different interests of stakeholders in decision-making processes. In fact, LCC plays an important input on a potential multicriteria decision-making tool
that should also integrate RAMS analysis, and in which both inputs should take associated uncertainties and risk into account. Nevertheless, there is a lack of knowledge on the main obstacles to LCC implementation, envisaging competitive problems on a regulated environment in a recent vertically separated sector. In fact, note that LCC analysis may serve as the basis to an economically sustainable taxation of infrastructure to operators, and the basis for regulator entities evaluating IM decisions efficiency, towards a more competitive sector that benefits the end-users (citizens).
References

Andersson, M. (2002). Strategic Planning of Track Maintenance – State of the Art, Division of Urban Studies, Department of Infrastructure, Royal Institute of Technology, (KTH), ISSN:1651-0216, pp.41-46.


IMPROVERAIL (2003). Deliverable 10 Project Handbook. IMPROVEd tools for RAILway capacity and access management, co-ordinated by TIS.PT.


Appendix 1

Public Function lambda(a, b, x, q)
    lambda = q * (a / b) * ((x / b) ^ (a - 1))
End Function

Public Function lambdab(a1, b1, a2, b2, a3, b3, x, q)
    lambdab = lambda(a1, b1, x, 1) + lambda(a2, b2, x, q) + lambda(a3, b3, x, 1)
End Function

Public Function lambdaa(a, b, a1, b1, a2, b2, a3, b3, T, n0, q)
    lambdaa = lambda(a, b, T, 1) * Exp(intlambda(0, T, a, b, 1)) * (n0 + 2 * intlambdab(a, b, a1, b1, a2, b2, a3, b3, T, q))
End Function

Public Function intlambdaa(a, b, a1, b1, a2, b2, a3, b3, T, n0, q)
    h = 0.25
    t0 = 0
    intlambdaa = 0
    While (t0 <= T)
        intlambdaa = intlambdaa + (lambdaa(a, b, a1, b1, a2, b2, a3, b3, t0, n0, q) + lambdaa(a, b, a1, b1, a2, b2, a3, b3, t0 + h, n0, q)) * (h / 2)
        t0 = t0 + h
    Wend
End Function

Public Function intlambdab(a, b, a1, b1, a2, b2, a3, b3, x, q)
    h = 0.25
    intlambdab = 0
    t0 = 0
    While (t0 <= x)
        intlambdab = intlambdab + (lambdab(a1, b1, a2, b2, a3, b3, t0, q) * Exp((1) * intlambda(0, t0, a, b, 1)) + lambdab(a1, b1, a2, b2, a3, b3, t0 + h, q) * Exp((-1) * intlambda(0, t0 + h, a, b, 1))) * (h / 2)
        t0 = t0 + h
    Wend
End Function

Public Function intlambda(t0, t1, a, b, q)
    intlambda = q * (((t1 / b) ^ a) - ((t0 / b) ^ a))
End Function
Public Function expect(a, b, beta, taxa, ti, tk, m, q)
    i = ti / m
    k = tk / m
    expect = ((1 - beta) ^ (i - k)) * expectfailure(a, b, taxa, ti, tk, m, q)
End Function

Public Function expect_a(a, b, a1, b1, a2, b2, a3, b3, beta, taxa, ti, tk, m, n0, q)
    i = ti / m
    k = tk / m
    expect_a = ((1 - beta) ^ (i - k)) * expectfailure_a(a, b, a1, b1, a2, b2, a3, b3, taxa, ti, tk, m, n0, q)
End Function

Public Function expectfailure(a, b, taxa, ti, tk, m, q)
    h = 0.25
    expectfailure = 0
    tk_1 = tk - m
    ti_1 = ti - m
    While (tk_1 <= tk)
        expectfailure = expectfailure + (lambda(a, b, tk_1, q) * (G_PF(ti - tk_1, (1 / taxa)) - G_PF(ti_1 - tk_1, (1 / taxa))) + lambda(a, b, tk_1 + h, q) * (G_PF(ti - (tk_1 + h), (1 / taxa)) - G_PF(ti_1 - (tk_1 + h), (1 / taxa)))) * (h / 2)
        tk_1 = tk_1 + h
    Wend
End Function

Public Function expectfailure_a(a, b, a1, b1, a2, b2, a3, b3, taxa, ti, tk, m, n0, q)
    h = 0.25
    expectfailure_a = 0
    tk_1 = tk - m
    ti_1 = ti - m
    While (tk_1 <= tk)
        expectfailure_a = expectfailure_a + (lambdala(a, b, a1, b1, a2, b2, a3, b3, tk_1, n0, q) * (G_PF(ti - tk_1, (1 / taxa)) - G_PF(ti_1 - tk_1, (1 / taxa))) + lambdala(a, b, a1, b1, a2, b2, a3, b3, tk_1 + h, n0, q) * (G_PF(ti - (tk_1 + h), (1 / taxa)) - G_PF(ti_1 - (tk_1 + h), (1 / taxa)))) * (h / 2)
        tk_1 = tk_1 + h
    Wend
End Function

Public Function expectsum(a, b, beta, taxa, ti, m, q)
    expectsum = 0
    tk = m
    While (tk <= ti)
        expectsum = expectsum + expect(a, b, beta, taxa, ti, tk, m, q)
        tk = tk + m
    Wend
End Function
Public Function expectsum_a(a, b, a1, b1, a2, b2, a3, b3, beta, taxa, ti, m, n0, q)
    expectsum_a = 0
    tk = m
    While (tk <= ti)
        expectsum_a = expectsum_a + expect_a(a, b, a1, b1, a2, b2, a3, b3, beta, taxa, ti, tk, m, n0, q)
        tk = tk + m
    Wend
End Function

Public Function numberfailures(a, b, beta, taxa, T, m, q)
    numberfailures = 0
    t0 = 0
    While (t0 <= T)
        numberfailures = numberfailures + expectsum(a, b, beta, taxa, t0, m, q)
        t0 = t0 + m
    Wend
End Function

Public Function numberfailures_a(a, b, a1, b1, a2, b2, a3, b3, beta, taxa, T, m, n0, q)
    numberfailures_a = 0
    t0 = 0
    While (t0 <= T)
        numberfailures_a = numberfailures_a + expectsum_a(a, b, a1, b1, a2, b2, a3, b3, beta, taxa, t0, m, n0, q)
        t0 = t0 + m
    Wend
End Function

Public Function G_PF(x, a)
    If (x >= 0) Then G_PF = 1 - Exp((-a) * x)
    If (x < 0) Then G_PF = 0
End Function