Availability Approach to Optimizing Railway Track Renewal Operations

Luis Filipe Caetano[1] and Paulo Fonseca Teixeira[2]

Abstract: This paper proposes a multiobjective optimization approach for planning railway ballast, rail, and sleeper renewal operations. The objective is to support an informed decision that considers not only the railway track life-cycle cost (LCC) but also the track occupation times required to perform interventions. Two objective functions were minimized, as follows: (1) railway track unavailability caused by railway track maintenance and renewal operations, and (2) railway track components' LCC. Furthermore, to rationalize the renewal strategy, the model considers a multicomponent formulation that assesses, in time and space, the opportunistic combination of railway track renewal activities. A numerical application of the model on a real case study (Lisbon-Oporto line) is developed and discussed. The results show the interest of using this simple multiobjective optimization approach to obtain a decision-making process to support the scheduling of major railway track renewal works with an informed LCC-unavailability trade-off. **DOI: 10.1061/(ASCE)TE.1943-5436.0000575.**

CE Database subject headings: Railroad tracks; Maintenance; Life cycles; Optimization.

Author keywords: Railway track; Track maintenance; Track renewal; Life-cycle cost; Availability; Multiobjective optimization; Nonsorting genetic algorithm 2 (NSGA2).

Introduction

The European reform model launched in the 1990s, starting with the splitting of railway infrastructure management from the operation business brought new challenges to infrastructure managers (IM). Nowadays, to meet the needs of operators more effectively, new requirements for infrastructure availability and reliability are being stipulated. Apart from guaranteeing that the infrastructure will perform well (by keeping a good record availability parameters such as the mean down time), the IM must be a reliable service provider. Among others, reliability and availability are particularly dependent and related with track maintenance and renewal (M and R) activities. When the preventive M and R works are not planned sufficiently far in advance, contingency measures such as speed reductions might be needed and the impact on track availability can be quite high. However, network M and R operations themselves cause temporary reduction of supply and also directly affect infrastructure availability. To support the decision maker, an optimization process that considers not only the investment but also performance parameters is particularly important in this context. However, common railway renewal optimization approaches generally tend to focus only on total investment (Budai et al. 2005; Gorman and Kanet 2010).

[1]Ph.D. Student, CESUR, Dept. of Civil Engineering, Instituto Superior Técnico (IST), Technical Univ. of Lisbon, Ave. Rovisco Pais, 1049-001 Lisbon, Portugal (corresponding author). E-mail: luis.caetano@ist.utl.pt

[2]Assistant Professor, CESUR, Dept. of Civil Engineering, Instituto Superior Técnico (IST), Technical Univ. of Lisbon, Ave. Rovisco Pais, 1049-001 Lisbon, Portugal. E-mail: pft@ist.utl.pt

Note. This manuscript was submitted on February 19, 2013; approved on May 13, 2013; published online on May 15, 2013. Discussion period open until February 1, 2014; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, Vol. 139, No. 9, September 1, 2013. © ASCE, ISSN 0733-947X/2013/9-941-948/$25.00.

Multiobjective formulation for optimization problems in the decision context of railway infrastructure management systems gives the opportunity to measure the trade-off between the various evaluation criteria that are important for an effective infrastructure investment. The primary advantage of this approach is the hypothesis that the decision-making choice either accepts or rejects a solution by comparing different feasible possibilities. Contrary to single-objective optimization problems, this formulation considers two objectives and analyzes them simultaneously (with the same significance). The objectives to be considered in an optimization formulation should reflect the true concerns of IMs in managing the railway network. A recent review on the use of multiobjective optimization models in the transport infrastructure context has been discussed in Wu et al. (2012). Despite the importance of such decision models, in railway track M and R planning this kind of optimization approach is still rare. Podofillini et al. (2006) developed a multiobjective optimization model to optimize the scheduling of rail inspections. To construct the Pareto frontier, this model considers the trade-off between the risks related to safety factors and the costs related to the inspection strategy. Additionally, Andrade and Teixeira (2011a) presented a multiobjective approach to assess the influence of M and R investment on train delays caused by speed restrictions required for corrective maintenance purposes with an origin in track geometry quality defects.

This paper addresses the development of a multiobjective optimization model for the railway track M and R management problem, focusing on the availability requirements (to perform preventive M and R operations) and the costs of such operations. The first objective of the model aims to minimize the life-cycle cost (LCC, i.e., costs during the life span of the analysis) of the track components subjected to constraints such as annual budget restrictions, whereas the second aims to minimize the track unavailability. Evolutionary algorithms are recently gaining significant attention in the field of transportation asset management because of their flexibility, efficiency, and robustness in finding the good solutions (Bai et al. 2012; Meneses and Ferreira 2013; Zhang et al. 2013). The model

J. Transp. Eng. 2013.139:941-948.

presented in this paper will use a nonsorting genetic algorithm 2 (NSGA2) approach to solve the optimization problem and it will be applied in the Lisbon-Oporto line using real degradation data.

Multiobjective Optimization Model

Multiobjective Optimization Using Genetic Algorithms

In a multiobjective optimization problem, if f_i and f_j are the objective functions that must be maximized and y and z are two feasible solutions, a solution z is said to dominate y if

$$\forall\ i{:}f_i(y) \le f_i(z) \quad \text{and} \quad \exists\ j{:}f_j(y) < f_j(z) \tag{1}$$

When y is not dominated by any other feasible solutions, y is known as a Pareto-optimal solution of the k-objective maximization problem (Coello et al. 2007). The objective of multiobjective optimization is to search for solutions in the global Pareto-optimal region (i.e., nondominated solutions for all the objective functions) and to achieve solutions that are separated from one another to the maximum possible extent on the nondominant front. The set of all Pareto optimal solutions forms the trade-off surface in the objective space. This trade-off surface is termed the Pareto front.

Genetic algorithms are a popular method to solve multiobjective problems. A considerable number of methods have been published that are widely used to perform multiobjective optimization using genetic algorithms, i.e., strength Pareto evolutionary algorithms (SPEAs), Pareto archived evolution strategy (PAES), vector-evaluated genetic algorithms (VEGAs), NSGA2, and so on. Some are better than others, and the most prevalent in the literature are NSGA2 and SPEA2. In this paper an NSGA2 is applied to solve the multiobjective problem. As demonstrated by Konak et al. (2006), this method achieves good results for most problems.

Regarding the performance of evolutionary multiobjective optimization techniques, the NSGA2 algorithm usually performs very well on two-objective problems, and tends to give a better spread of the solutions and better convergence than SPEA2 (Deb 1999, 2001; Deb et al. 2002). NSGA2 is designed to search for a set of well-distributed nondominated solutions that approximates the entire Pareto front very well. The pseudocode of NSGA2 is presented in the appendix. Further details on NSGA2 are given by Deb et al. (2002).

Development of an NSGA2 for Optimal Railway Track Renewal Planning: Model Formulation

This model was formulated considering the possibility of performing different combinations of renewal operations of the railway track components. More specifically, the ballast, rail, and sleepers' renewal operations can be combined among them to achieve a better solution for the optimization problem.

Consider a railway track composed of a set of $K = \{1,\ldots,k\ldots K\}$ components and a set of track segments, $N = \{1,\ldots,n\ldots N\}$. Let $T = \{1,\ldots,t\ldots T\}$ be the set of time steps in the planning timespan and $W = \{1,\ldots,w\ldots W\}$ the set of renewal works that can be executed. The decision variable z_{kn} is the time in which the renewal of the track component $k \in K$ in the track segment $n \in N$ (Fig. 1) is performed. This variable defines the model as a discrete multiobjective optimization problem.

The multiobjective optimization model introduced previously can be formulated as

$$\min\ LCC = \sum_{w\in W}\sum_{t\in T}\sum_{n\in N} C_w^r L_{wtn} + \sum_{k\in K}\sum_{t\in T}\sum_{n\in N} C_k^m X_{ktn} - \sum_{w\in W}\sum_{t\in T}\sum_{n\in N} C_w^m L_{wtn} P_{wt} - \sum_{k\in K}\sum_{n\in N} R_{kn} \tag{2}$$

$$\min\ U = \sum_{w\in W}\sum_{t\in T}\sum_{n\in N} U_{wtn}^r + \sum_{k\in K}\sum_{t\in T}\sum_{n\in N} U_{ktn}^m \tag{3}$$

$$R_{kn} = \left[\frac{X_k^{\max} - (X_{ktn}/L_n)}{X_k^{\max}}\right] RC_k L_n \qquad \forall\ t = T, \quad k \in K, \quad n \in N \tag{4}$$

$$P_{wt} = 1 - \left(\sum_{n\in N} L_{wtn}\right)^{-0.1} \qquad \forall\ w \in W, \quad t \in T \tag{5}$$

$$U_{wtn}^r = \frac{\mathrm{MDT}_w^r L_{wtn}}{M_t} \qquad \forall\ w \in W, \quad t \in T, \quad n \in N \tag{6}$$

$$U_{ktn}^m = \frac{\mathrm{MDT}_k^m X_{ktn}}{M_t} \qquad \forall\ k \in K, \quad t \in T, \quad n \in N \tag{7}$$

$$\sum_{n\in N} L_{wtn} \le L_w^{\max} \qquad \forall\ w \in W, \quad t \in T \tag{8}$$

$$\sum_{w\in W}\sum_{n\in N} C_w^r L_{wtn} + \sum_{k\in K}\sum_{n\in N} C_k^m X_{ktn} - \sum_{w\in W} C_w^r P_{wt} \le B_t \quad \forall\ t \in T \tag{9}$$

where C_w^r = unitary cost of the renewal work $w \in W$; C_k^m = maintenance cost for track component $k \in K$; X_{ktn} = number of maintenance operations for each track component $k \in K$ performed at

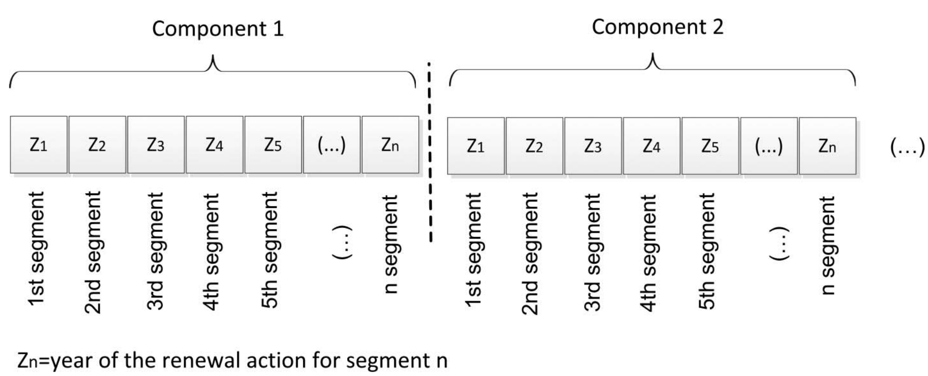

Fig. 1. Codification of decision variables

time step $t \in T$ in each track segment $n \in N$, computed based on the degradation model of each track component (to be presented in a subsequent section); R_{kn} = residual value of track component $k \in K$ in track section $n \in N$ obtained at the end of the planning time horizon T (ensure that the renewal scheduling does not depend on the value of the planning time horizon); $X_k^{\max}$ = threshold that identifies the maximum yearly number of cumulative maintenance operations (maximum serviceability) of the track component $k \in K$; RC_k = renewal cost of component $k \in K$ for large-scale major renewal operations (minimum renewal cost for each track component); L_n = length of each segment $n \in N$; P_{wt} = percentage saved in the unit renewal cost $w \in W$ initiated in time period $t \in T$ (scale economy factor) by performing clusters of segments $n \in N$; L_{wtn} = length of the segment $n \in N$ in which a renewal work $w \in W$ was performed in the time step $t \in T$; U_{wtn}^r = unavailability caused by the renewal work $w \in W$ in the time step $t \in T$ on the track segment $n \in N$; M_t = total amount of time in the time step $t \in T$; MDT_w^r = mean downtime (per renewed length) required to perform renewal work $w \in W$; U_{ktn}^m = unavailability caused by systematic maintenance of the track components $k \in K$ measured in time step $t \in T$ on the track segment $n \in N$; MDT_k^m = mean downtime needed to perform maintenance on track component $k \in K$ (localized interventions); $L_w^{\max}$ = maximum extent of renewal works of type $w \in W$ in each time step $t \in T$; and B_t = maximum budget in the time step $t \in T$.

Eq. (2) is the first objective function and consists of the minimization of the total railway track LCC. Eq. (3) is the second objective function and intends to minimize the track unavailability events during the planning time span. Eq. (4) defines the track residual at the end of the planning timespan (this value depends on the remaining serviceability of the different components). Eq. (5) defines the percentage of saving related to economy-of-scale benefits and results from preliminary research work conducted under a research project with the Portuguese IM Rede Ferroviária Nacional (REFER). This being the case, the expression coefficients should be regarded as a first approximation of the real coefficients and it is expected that further research will clarify whether or not this relationship is appropriate. Eqs. (6) and (7) determine the unavailability attributable to M and R works. The writers are considering the railway track as a series system. This indicates that a component k in a track segment n will be out of operation during the maintenance or renewal of any other component (the railway track will be in an unavailable state).

Eq. (8) indicates the yearly renewal length limit. Eq. (9) represents the yearly budget restriction. The model constraints [i.e., Eqs. (8) and (9)] were imposed indirectly by penalizing the objective functions with a quantity proportional to the degree of violation (i.e., distance to feasibility) given by each constraint.

Development of an NSGA2 for Optimal Railway Track Renewal Planning: NSGA2 Codification

The solution coding represents the translation of the problem into chromosomes. The total number of chromosomes is termed the population size. In this case, the gene position represents the spatial location of a track segment such that the total number of genes in a chromosome represents the length of the line that is being analyzed. The integer value in each gene is the decision variable (z_{kn}) and represents the year in which the renewal should be done in each track component $k \in K$ in segment $n \in N$. A zero value in a gene indicates that no renewal action is assigned in the planning timespan. This codification allows the assessment of the possible renewal combination of segments [Eq. (5)] and of components (by considering benefits of performing renewal works groupings) in the same time step.

Model Implementation

The proposed model was applied to a real case study to demonstrate the robustness and usefulness of the methodology. The Portuguese Lisbon-Oporto line has been partially renewed and its track components are at different stages of their life cycles. The track structure is composed of a ballasted track, concrete monoblock or wood (in some nonrenewed segments), and UIC60 and UIC54 rails (in renewed and nonrenewed segments, respectively). In the renewed segments the substructure was reinforced and a subballast layer was constructed that resulted in a modulus of elasticity at the subballast level of 120 MPa. This is interesting in terms of M and R planning because of the temporal dispersion of M and R requirements. Furthermore, the Lisbon-Oporto line is a double-track line, 336 km in length. It is subject to a considerable level of mixed freight and passenger traffic amounting to approximately 15 million gross tonnes (MGT) per track per year at maximum train running speeds of 220 km/h. The annual traffic load considered in the model was assumed to be constant over the planning timespan. The most important service in this line is the long-distance connection between Lisbon and Oporto.

Railway Track Availability

Unplanned maintenance has serious impacts on infrastructure availability, leading to train cancelations and delays (primary and secondary). Unavailability directly affects the infrastructure performance and it is measurable through indices such as mean downtime. In the context of rail infrastructure, mean down time consists of the following:

- MDT^m, time when railway track is occupied because of systematic maintenance actions; and
- MDT^r, time when rail traffic is occupied because of planned heavy renewal works.

This paper distinguishes systematic maintenance tasks from major renewal works. Systematic maintenance includes partial rail replacement (approximately 3–6 m) attributable to failure, local replacement of broken sleepers, and ballast tamping on track segments of 200 m. Major renewal works, however, corresponds with total replacement of track components (combined or not) in a track segment. For an LCC optimization approach the writers are just considering MDT^m and MDT^r indices when assessing the track availability. The state assessment performed in this model was divided by k components in the track segment position n.

In other industry sectors, some M and R strategies have been developed that take into account several operating units and the production loss that occurs when each component is shut down. The problem of multicomponent systems is tackled by several opportunistic replacement models reported in the literature (Haque et al. 2003; Laggoune et al. 2010). When the system is down, either correctively or preventively, the opportunity to replace nonfailed components preventively is considered to increase availability and reduce the operating costs of the system. This is also important in the railway infrastructure system. Therefore, an opportunistic replacement approach was developed and integrated into the multiobjective optimization problem model to optimize major renewal works presented in this paper.

Railway Track Components' Degradation Modeling: Ballast

Railway track geometry is degraded under traffic loads and must undergo periodic maintenance. Tamping is the operation that restores the track geometry. The requirements for this intervention are usually related to levels of quality that have a direct influence

J. Transp. Eng. 2013.139:941-948.

on user comfort and safety. In the most primary lines in Europe, the vertical and horizontal track alignments are the primary parameters that trigger preventive tamping operations. These are commonly quantified as a SD for the short wavelength (3–25 m) of the vertical and horizontal alignment defects in a 200-m track segment. As discussed in Andrade and Teixeira (2013), the SD of the vertical and horizontal alignment, when used to trigger preventive maintenance actions, are good predictors of corrective maintenance needs regarding all the remaining indicators of track geometry.

A linear trend of the SD of the vertical and horizontal alignment measure in a 200-m track segment is a widely-used approximation (Esveld 2001). This degradation law is given by the expression

$$\sigma_{VA}(l) = c_{0,VA} + c_{1,VA}l \qquad \sigma_{HA}(l) = c_{0,HA} + c_{1,HA}l \tag{10}$$

where l = accumulated axle tonnage/loads in MGT in a 200-m track segment subsequent to the last renewal or tamping; $c_{0,VA}$ = initial value of the SD of the vertical alignment defects (mm); $c_{0,HA}$ = initial value of the SD of the horizontal alignment defects (mm) after tamping or renewal; $c_{1,VA}$ = degradation rate of the vertical alignment defects (mm/MGT); and $c_{1,HA}$ = degradation rate of the horizontal alignment defects (mm/MGT).

The effectiveness of ballast tamping operations tends to decrease with the number of tamping cycles. The declining ability of tamping to restore the geometrical parameters is explained by the degradation of ballast with traffic loads and successive tamping operations (crushing of the ballast particles). This behavior leads to a track degradation model that exhibits an increasing need for maintenance to preserve riding comfort levels. Introducing the efficiency parameter related to tamping in Eq. (10) results in Eq. (11)

$$\sigma_{VA}(g;l) = c_{0,VA}\beta_{VA}^{g} + c_{1,VA}l \qquad \sigma_{HA}(g;l) = c_{0,HA}\beta_{HA}^{g} + c_{1,HA}l \tag{11}$$

where g = cumulative number of tamping operations subsequent to the last renewal; and β = loss of efficiency of tamping as more of this work is carried out (from the beginning of the ballast life cycle) for the vertical and horizontal alignments.

Railway Track Components' Degradation Modeling: Rail and Sleeper Failure

Rail service life is largely determined by three primary deterioration indices, as follows: (1) rate of fatigue, (2) level of rail wear, and (3) corrugation level. Rail fatigue failures can generally be divided into surface-initiated, weld, and internal failure, which can lead to the need of rail renewal.

Published studies on modeling sleeper degradation generally present estimations that consider the failure of the entire sleeper in most cases without specification of individual modes. Such estimations were based on periodic visual inspections. Most of the IMs establish their own standards for systematic maintenance operations of these components. In this paper, the writers assumed that a minimum number of one or two consecutive failed sleepers is required to trigger a maintenance operation (localized sleeper replacement).

Given that the prediction of the occurrence and progression of rail and sleeper defects is complex and depends on the condition and features of the track section, the writers adopted models that have already been developed in previous studies. Most cases showed that rail fatigue defects and sleeper failure follow a Weibull law (Shyr and Ben-Akiva 1996; Nilsson and Olofsson 2002; Yun and Ferreira 2003; Stirling et al. 2006; Zhao et al. 2007). In track maintenance planning, stochastic modeling of component degradation is nearly always represented by hazard rate (Lyngby et al. 2008). Assuming that the probability distribution that describes failure occurrence is approximated by a two-parameter Weibull distribution, the hazard rate of a defect at time t is given by

$$\lambda(t) = \left(\frac{\alpha}{\eta}\right)\left(\frac{t}{\eta}\right)^{\alpha-1} \tag{12}$$

where α is a shape parameter; η = scale parameter of the distribution; t = accumulated tonnage borne by the track component over its life cycle; and $\lambda(t) \geq 0$, $\alpha \geq 0$, $\eta \geq 0$.

Parameters of the Model for the Lisbon-Oporto Line

Each railway track component has a different degradation process and requires different maintenance and renewal operations to restore its performance. The inputs for the degradation models explained previously were obtained through an exhaustive analysis of track component behavior. Ten years (from 2001–2010) of historical data from track geometry, ultrasonic, and visual inspections of the Lisbon-Oporto line were used to define the degradation models.

As shown in previous track degradation models, the track geometry degradation rates for the SD of vertical and horizontal alignment defects have an important dispersion for different 200-m segments, even if they are contiguous (Esveld 2001; Andrade and Teixeira 2011b). In this paper, the deterministic values of the degradation parameters that describe the current track behavior were computed for each 200-m track segment [Eq. (11)]. The tamping operation for each 200-m track segment is triggered when the value of σ_{VA} or σ_{HA} reaches the tolerances described in Table 1 and was established in the Portuguese standard *IT.VIA.018* (REFER 2009).

The uncertainty of track behavior after major renewal operations was modeled using a random selection of these values ($\sigma_{0,VA}$, $\sigma_{0,HA}$, $c_{0,VA}$, and $c_{0,HA}$) in accordance with the distributions obtained through goodness-of-fit tests using the real data on renewed track segments. Moreover, the writers also calibrated the parameters that represent the declining efficiency in recovering the track geometry quality parameters (SD of vertical and horizontal profile, i.e., β_{VA} and β_{HA}) with an increasing number of tamping operations. These values represent the dispersion of the degradation parameters obtained from the geometry inspection data of renewed track segments.

The parameters of rail and sleeper Weibull distribution were calibrated to consider the characteristics of the components (age, type of rail and sleeper, and so on) provided by inspection data from the case study line (Table 2). With this, and considering the cumulative traffic for each track segment, the number of systematic maintenance operations was determined for the rail and sleeper components. The simulation performed for rails follows the model proposed by Zhao et al. (2006), and that for sleepers follows the models proposed by Yun and Ferreira (2003).

Table 1. Degradation and Tolerances Parameters for Railway Track Geometry Considered in the Case Study

Parameter	SD vertical alignment	SD horizontal alignment
Tolerances in the Portuguese standard (mm) 160 km/h < speed ≤ 230 km/h	2.3	1.3
Initial quality σ_0 (mm)	$\sigma_{0,VA}$ $\sim \log N(0.55; 0.17)$	$\sigma_{0,HA}$ $\sim \log N(0.57; 0.13)$
Initial degradation rate c_0 (mm/MGT)	$c_{0,VA}$ $\sim \log N(0.037; 0.041)$	$c_{0,HA}$ $\sim \log N(0.020; 0.016)$

J. Transp. Eng. 2013.139:941-948.

Table 2. Parameters of the Weibull Distribution for Rail and Sleeper Components Considered in the Case Study

Component	Shape parameter β	Scale parameter η (MGT)
Sleeper, wood	3	300
Sleeper, concrete	3	500
Rail, UIC56	1.01	315
Rail, UIC60	1.01	400

A simulation procedure was used to determine the time and location of maintenance works during the planning timespan. Fig. 2 summarizes the process of simulating the degradation for each track component (ballast, sleepers, and rail). The first step was to define the probability density function for each random variable that describes the degradation process of each track component. A Monte Carlo simulation (MCS) was then run for each time step of the model to set random numbers in accordance with the defined distributions [for further details on the Monte Carlo simulation, see Kalos and Whitlock (2008)]. The loop continues until the planning timespan is reached. This simulation is performed for each component and for the different track segments considered in the model. Before the renewal work (scheduled in the planning timespan), the track geometry degradation simulation process is different because the writers have an accurate deterministic model to estimate ballast tamping operations. Therefore, to represent the uncertainty in the track geometry quality achieved by the renewal operation, the MCS results for track geometry are just used in the optimization model after performing this operation.

The data analysis of ultrasonic inspections of the Lisbon-Oporto line found that 52% of rail fatigue defects result from aluminothermic weld defects, 20% from transversal cracks, 8% from vertical/longitudinal cracks, and 20% are attributable to other fatigue defects. The hazard rate for aluminothermic weld defects depends on the number of welds along a determined length of rail. The simulation model takes this into account by considering that aluminothermic welds increase with the number of local rail replacements.

For each track component in each track segment a function was adjusted that describes the increase of systematic maintenance works given by the simulation procedure. The resulting function of the accrued number of maintenance operations X_{ktn} for each track component $k \in K$ performed during the time step $t \in T$ in each track segment $n \in N$ is given by

$$X_{ktn} = H_{nk}(1+\mu_{nk})^{t+1} - H_{nk}(1+\mu_{nk})^{t} \qquad \forall\ t \in T,\quad k \in K,\quad w \in W \tag{13}$$

where H_{nk} is a scale factor; and μ_{nk} = maintenance power factor in track segment $n \in N$ of component $k \in K$.

Fig. 3 gives a graphical representation of this expression for all components. As the life cycle increases the maintenance requirements grows with a nonlinear trend.

If component $k \in K$ in a track section $n \in N$ is replaced, the function of the number of systematic maintenance operations is updated and returns to a state of good-as-new. Given that the Lisbon-Oporto line has components of varying ages the values of the failure and maintenance requirements in each track segment $n \in N$ in the same time step $t \in T$ can differ significantly. For this paper it was assumed that wear and corrugation do not have a significant influence on the number of systematic rail maintenance activities (e.g., localized rail replacement and maintenance) and therefore they were not considered in the model.

Most of the results from heuristic techniques are sensitive to the input parameters. Population size (i.e., the number of solutions in each generation N_p) must be carefully chosen; if the population is too small the risk of premature convergence of the algorithm to a poor local optimum solution will be serious, and if it is too large the effort needed to run the algorithm will be excessive. Crossover and mutation are used to create new populations and control the diversity of the solution to be tested in each iteration. Hence, extensive simulations are required when setting these parameters to find suitable values for the various parameters. The best parameters for the NSGA2 described in the appendix were selected by running the simulation for a 20-year planning time span and 21 track segments with heterogeneous characteristics (traffic level and track age) constituted by three components (ballast, rail, and sleepers). Table 3 shows the chosen values.

M and R costs and work times were estimated by analyzing the data provided by the Portuguese IM. Table 4 presents the parameters considered for the different M and R works considered in the model. These values can vary depending on the contractors involved in railway track renewal. The writers are considering the opportunistic renewal of track components by establishing a combination of renewal operations. The unitary costs and down times can be adapted in accordance with the other factors that influence the difficulty of performing the maintenance or renewal operations in specific track segments.

The possession time for systematic maintenance operations considers an additional period for the time spent traveling to the work site (ACEM-Rail 2011). In this paper the writers adopted an average period of 1 h for traveling to the workplace.

Table 5 presents the remaining parameters of the model.

Given that the writers are considering a large planning timespan, the discounted costs should be considered. However, because this

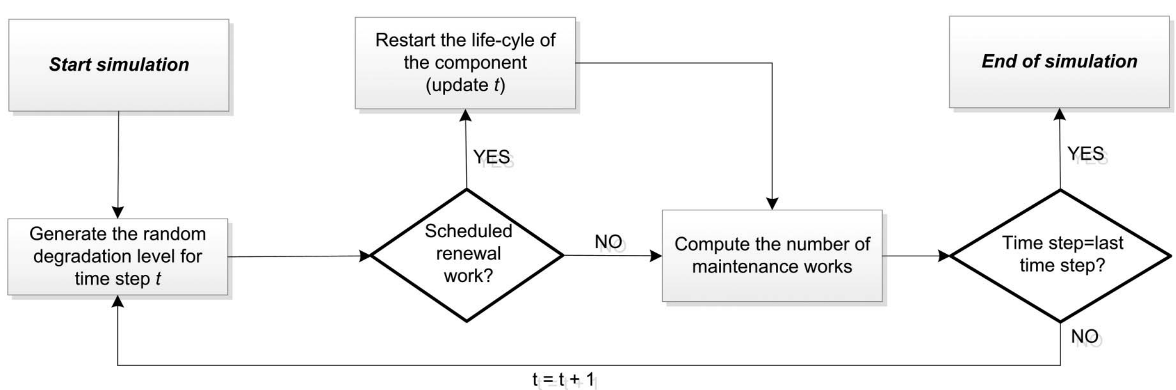

Fig. 2. Simulation procedure to estimate the maintenance requirements for each track component in each track segment

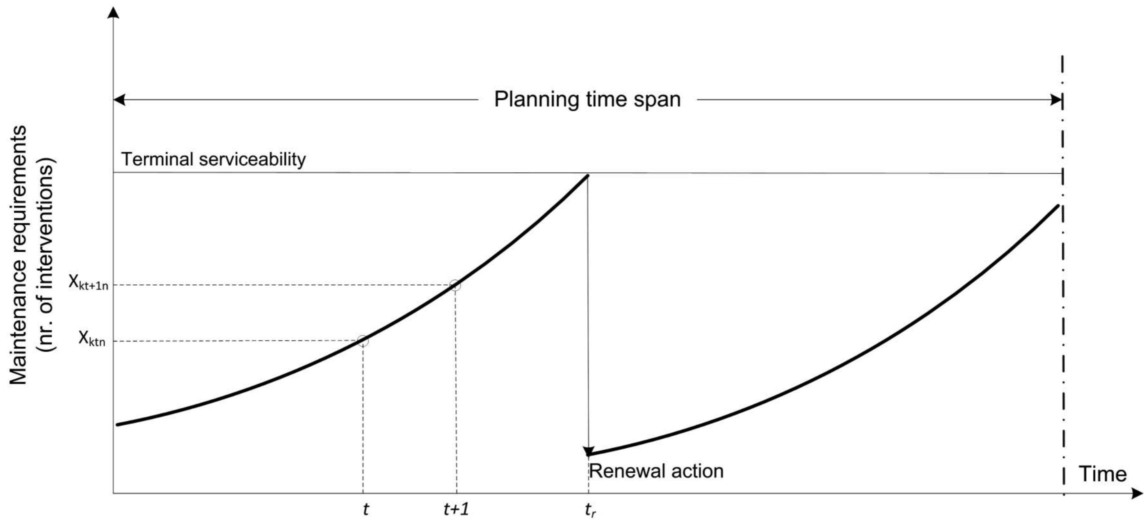

Fig. 3. Function of maintenance works requirements

Table 3. Best Parameters for 20-Year Planning Horizon

NSGA2 parameters	Type of parameter values
Population size, N_p	30
Number of iterations	500
Crossover probability, P_c	0.4
Mutation probability, P_m	0.33, i.e., 1/3
Crossover index, η_c	SBX
Mutation index, η_m	Polynomial mutation
Controlled elitism, r	Geometric distribution

Note: SBX = simulated binary crossover.

Table 4. Railway Track M and R Reference Costs and Reference Mean Downtimes Considered

Type of work	Cost (€)	Mean down time (h)
Ballast tamping and stabilization: c_1^m (€/200-m), MDT_1^m (h/200-m)	2,500	1.0
Ballast renewal: c_1^r (€/km), MDT_1^r (h/km)	130,000	6.5
Systematic rail maintenance, approximately 3 m segment of new rail: c_2^m (€/failure), MDT_2^m (h/failure)	850	1.0
Rail renewal: c_2^r (€/km), MDT_2^r (h/km)	75,000	3.5
Systematic sleeper maintenance: c_3^m (€/failure), MDT_3^r (h/failure)	300	2.0
Sleepers renewal: c_3^r (€/km), MDT_3^r (h/km)	60,000	5.0
Combination of ballast and rail renewal: c_4^r (€/km), MDT_4^r (h/km)	190,000	8.0
Combination of ballast and sleeper renewal: c_5^r (€/km), MDT_5^r (h/km)	180,000	8.0
Combination of rail and sleeper renewal: c_6^r (€/km), MDT_6^r (h/km)	113,000	6.5
Combination of rail, ballast and sleeper renewal: c_7^r (€/km), MDT_7^r (h/km)	228,000	10.0

Table 5. Limit Parameters Considered in the Model

Parameter	Value
Maximum serviceability of track components $X_k^{\max}$, number of systematic maintenance operations per year	15
Maximum extent of major renewal works in 1 year $L_w^{\max}$ (km)	70
Maximum yearly budget B_t (M€)	20

affects the renewal schedule (advantage of performing renewal operations as late as possible), the writers decided not to incorporate this in the model. Whether the costs are discounted or not, the developed model and ideas are still valid, without loss of generality.

Optimization Results

The proposed algorithm was coded in *MATLAB* with a Windows 7 operating system (OS), Intel Core Duo 2.70-GHz [4-GB random-access memory (RAM)] personal computer. Fig. 4 shows the optimization results after the convergence of the Pareto frontier with a total number of 500 generations. Even though the optimization model considers the minimization of the LCC, the costs presented in Fig. 4 are average values of the M and R costs incurred during the planning timespan (to better assess the trade-off between the required investment and availability achieved by the solutions).

Fig. 4 indicates that greater investment in railway track M and R works guarantees a lower level of unavailability. In practical terms, if there is an average unavailability of 10%, it corresponds with approximately 2.4 h/day that the railway line is occupied by M and R operations.

If the decision maker is cost-sensitive, solution 1 (S_1) should be chosen because it optimizes track availability for the least M and R investment. However, if high track availability is needed on the line regardless of the M and R cost, solution 3 (S_3) should be selected because it gives the highest track availability. Furthermore, an M and R investment or an unavailability level less than preestablished values may be required. For instance, if the IM must have more than 92% track availability, solution 2 (S_2) should be selected because it guarantees the required availability level at the minimum M and R cost. This requirement can occur when, for instance, trains are running during most of the day (freight operations during night) and the daily maintenance time window is very restricted.

Budget constraints can have a considerable influence on track availability. For example, if the expected deterioration level in a given time step is very high, it is necessary to maintain the segment with spot repair and frequent inspection. However, because the budget for these interventions is limited, in extreme situations (when the intervention works are not planned with sufficient notice) this type of work increases in such a manner that the impact on track availability is significantly high. To determine this effect, different budget constraint values were tested. Fig. 5 shows that the Pareto front for a railway track LCC-unavailability trade-off.

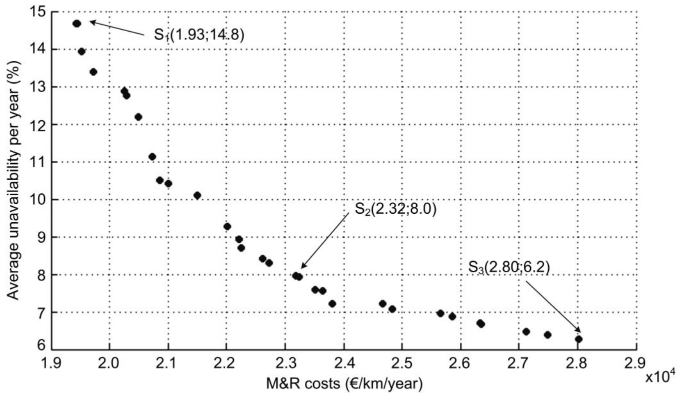

Fig. 4. Pareto optimal front

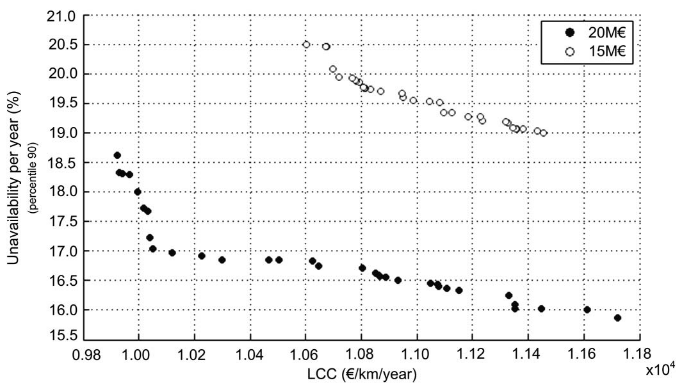

Fig. 5. Comparison of Pareto optimal front for different budget scenarios

This sensitivity analysis studied the variation of the highest values observed of track unavailability (percentile 90 of the unavailability results) when yearly budget restrictions are imposed. Contrary to Figs. 4 and 5 presents the LCC results (values for the first objective function).

Analyzing the maximum unavailability interval for each scenario (represented by the Pareto optimal frontier) illustrated in Fig. 5 shows that the level of unavailability that can be achieved worsens as the budget constraint increases. This can be described more precisely, as follows:

- Budget constraint of €15 million, maximum unavailability ranges between 19.0 and 20.5%; and
- Budget constraint of €20 million, maximum unavailability ranges between 15.8 and 18.6%.

If more M and R works can be carried out in 1 year, the availability that can be achieved is significantly higher. In a few cases, budget restrictions imply the postponement of renewal works although a high number of systematic maintenance operations are still undertaken. Given that the renewal work cannot be performed within the optimal schedule, the LCC increases. For the worst case in terms of unavailability impact, given that the writers are measuring the 90th percentile, 90 unavailability indicates that 10% of the planning timespan will require 20.5% of the downtime for M and R works. This is almost 5 h of daily downtime and exceeds by 1 h the time window that is usually available for systematic M and R works. In these circumstances a train timetable must be implemented that guarantees an additional 1 h without trains running on the line. Such implications can have a high impact on the quality of service provided to train users.

Conclusions

This paper demonstrates how with the use of a simple multiobjective optimization approach it is possible to obtain a decision-making process to plan the scheduling of major renewal works with an informed LCC-unavailability trade-off. The presented model permits the estimation of the necessary yearly budget to achieve preestablished levels of track availability, which can be a valuable tool to support strategic decisions of the IM.

Several opportunistic replacements of other components have been assessed during track occupation or downtime imposed for renewal or maintenance work on one track component. If it minimizes the total LCC, the track components' life cycles are rearranged to allow the grouping of renewal operations. The case study results indicated that more investment in M and R operations leads to higher track availability. The M and R cost-unavailability Pareto front indicates that the trade-off between these two parameters follows a nonlinear trend. Additionally, budget restriction significantly affects the LCC-unavailability trade-off. As the writers have shown, budget restrictions may require postponing renewal works from the optimal schedule and therefore increase LCC value. In extreme cases train operations can be affected by M and R operations and it is possible that timetables must be adapted.

Future steps to further improve the model will include the incorporation of other degradation factors (e.g., rail wear and rail corrugation) and consideration of more maintenance operations related to railway track components (e.g., grinding operations and manual tamping). Furthermore, other project management objectives such as track reliability maximization could also be incorporated in future developments of the model.

Appendix. Pseudocode of NSGA2

Set the parent vector $P = \varphi$, offspring vector $Q = \varphi$, collect vector $R = \varphi$, and generation number $t = 0$.
Initialize the parent vector P_0 in accordance with the decision variables' codification.
While $t <$ the terminate generation number do
  1. Combine the parent and offspring population through $R_t = P_t \cup Q_t$
  2. Sort all solutions of R_t to get all nondominated fronts F = fast-nondominated-sort (R_t) where $F = (F_1, F_2, \ldots)$
  3. Set $P_t + 1 = \varphi$ and $i = 1$
  4. While the parent population size $|P_{t+1}| + |F_i| < N$ do
    a. Calculate crowding-distance of F_i
    b. Add the ith nondominated front F_i to the parent pop P_{t+1}
    c. $i = i + 1$
  End while
  5. Sort the F_i in accordance with the crowding distance
  6. Fill the parent pop P_{t+1} with the first $N - |P_{t+1}|$ elements of F_i
  7. Generate the offspring population to Q_{t+1}
  8. Set $t = t + 1$
End while
The populations in vector P are the nondominated solutions.

Acknowledgments

The writers thank the Portuguese Railway Infrastructure Manager, REFER, E.P.E., for its support and cooperation. The support of the Portuguese Foundation for Science and Technology (FCT), FCT project reference PTDC/SENTRA/112975/2009 and FCT Ph.D. grant reference SFRH/BD/51327/2010 are also gratefully acknowledged. The writers also thank the reviewers for their comments, which enabled the presentation of this paper to be improved.

J. Transp. Eng. 2013.139:941-948.

References

Andrade, A., and Teixeira, P. (2011a). "Biobjective optimization model for maintenance and renewal decisions related to rail track geometry." *Transportation Research Record 2261*, Transportation Research Board, Washington, DC, 163–170.

Andrade, A., and Teixeira, P. (2011b). "Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study." *J. Transp. Eng.*, 137(3), 193–200.

Andrade, A., and Teixeira, P. (2013). "Unplanned maintenance needs related to rail track geometry." *Proc. of the ICE – Transport*, 166, 1–11.

Automated, and Cost-Effective Railway Infrastructure Maintenance (ACEM-Rail). (2011). *D1.1 report on the state of practice of railway infrastructure maintenance*, Málaga, Spain.

Bai, Q., Labi, S., and Sinha, K. (2012). "Trade-off analysis for multiobjective optimization in transportation asset management by generating Pareto frontiers using extreme points nondominated sorting genetic algorithm II." *J. Transp. Eng.*, 138(6), 798–808.

Budai, G., Huisman, D., and Dekker, R. (2005). "Scheduling preventive railway maintenance activities." *J. Oper. Res. Soc.*, 57(9), 1035–1044.

Coello, C. A. C., Lamont, G., and Van Veldhuizen, D. (2007). *Evolutionary algorithms for solving multi-objective problems*, Springer, New York and London.

Deb, K. (1999). "Multi-objective genetic algorithms: Problem difficulties and construction of test problems." *Evol. Comput.*, 7(3), 205–230.

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*, Wiley, New York.

Deb, K., Pratap, D., Agarwal, S., and Meyarivan, T. (2002). "A fast and elitist multiobjective genetic algorithm: NSGA-II." *IEEE Trans. Evol. Comput.*, 6(2), 182–197.

Esveld, C. (2001). *Modern railway track*, MRT Productions, Zaltbommel, Netherlands.

Gorman, M., and Kanet, J. (2010). "Formulation and solution approaches to the rail maintenance production gang scheduling problem." *J. Transp. Eng.*, 136(8), 701–708.

Haque, S., Kabir, A., and Sarker, R. (2003). "Optimization model for opportunistic replacement policy using GA with fuzzy logic controller." *Proc., IEEE Congress on Evolutionary Computation*, Canberra, Australia, 2837–2843.

Kalos, M., and Whitlock, P. (2008). *Monte Carlo methods*, Wiley, Weinheim, Germany.

Konak, A., Coit, D., and Smith, A. (2006). "Multi-objective optimization using genetic algorithms: A tutorial." *Reliab. Eng. Syst. Safety*, 91(9), 992–1007.

Laggoune, R., Chateauneuf, A., and Aissani, D. (2010). "Impact of few failure data on the opportunistic replacement policy for multi-component systems." *Reliab. Eng. Syst. Safety*, 95(2), 108–119.

Lyngby, N., Hokstad, P., and Vatn, J. (2008). "RAMS management of railway tracks." *Handbook of performability engineering*, K. B. Misra, ed., Springer, London, 1123–1145.

MATLAB version R2010a [Computer software]. MathWorks, Natick, MA.

Meneses, S., and Ferreira, A. (2013). "Pavement maintenance programming considering two objectives: Maintenance costs and user costs." *Int. J. Pavement Eng.*, 14(2), 206–221.

Nilsson, R., and Olofsson, U. (2002). "Surface cracks and wear of rail: A full-scale test on a commuter train track." *Proc. Inst. Mech. Eng. Part F, J. Rail Rapid Transit*, 216(4), 249–264.

Podofillini, L., Zio, E., and Zille, V. (2006). "Risk-informed optimisation of railway tracks inspection and maintenance procedures." *Reliab. Eng. Syst. Safety*, 91(1), 20–35.

Rede Ferroviária Nacional (REFER). (2009). "Tolerâncias dos parâmetros geométricos da via." *IT.VIA.018*, Lisbon, Portugal (in Portuguese).

Shyr, F., and Ben-Akiva, M. (1996). "Modeling rail fatigue behavior with multiple hazards." *J. Infrastruct. Syst.*, 2(2), 73–82.

Stirling, A., Roberts, C., Chan, A., and Zhao, J. (2006). "Assessing the economic life of rail using a stochastic analysis of failures." *Proc. Inst. Mech. Eng. Part F, J. Rail Rapid Transit*, 220(2), 103–111.

Wu, Z., Flintsch, G., Ferreira, A., and Picado-Santos, L. (2012). "Framework for multiobjective optimization of physical highway assets investments." *J. Transp. Eng.*, 138(12), 1411–1421.

Yun, W., and Ferreira, L. (2003). "Prediction of the demand of the railway sleepers: A simulation model for replacement strategies." *Int. J. Prod. Econ.*, 81–82(1), 589–595.

Zhang, T., Andrews, J., and Wang, R. (2013). "Optimal scheduling of track maintenance on a railway network." *Qual. Reliab. Eng. Int.*, 29(2), 285–297.

Zhao, J., Chan, A., and Burrow, M. (2007). "Reliability analysis and maintenance decision for railway sleepers using track condition information." *J. Oper. Res. Soc.*, 58(8), 1047–1055.

Zhao, J., Chan, A., Stirling, A., and Madelin, K. (2006). "Optimizing policies of railway ballast maintenance and renewal." *Transportation Research Record 1943*, Transportation Research Board, Washington, DC, 50–56.

J. Transp. Eng. 2013.139:941-948.