A heuristic approach to railway track maintenance scheduling

L. M. Quiroga & E. Schnieder
Institute for Traffic Safety and Automation Technologies, TU Braunschweig, Germany

Abstract

Travelling safely and comfortably on high speed railway lines requires excellent conditions of the whole railway infrastructure in general and of the railway track geometry in particular. The maintenance process required to achieve such excellent conditions is largely complex and expensive, demanding an increased amount of both human and technical resources. In this framework, an optimal scheduling of maintenance interventions is an issue of increased relevance. In this work a method for optimization of the tamping scheduling is presented. It is based on a heuristic algorithm, which finds a very detailed tamping schedule where each planned intervention is fully specified. The algorithm tries to maximize an objective function, which is a quantitative expression of the maintenance process's objectives defined by the railway company. It first finds an upper bound for the objective function value, and then returns the best feasible solution found. The method is validated by means of a case study based on real data of the 240 km track of a French high speed TGV line. The results presented show that the value of the best solution found is very near to the upper bound (the difference is smaller than 1%), with a calculation time of under 1 second using a standard computer, so we think the heuristic has a great performance potential.
Keywords: track maintenance, heuristics, tamping, scheduling.

1 Introduction

Measuring and keeping railway geometry under control are fundamental tasks of a railway infrastructure maintenance process. Railway geometry is representative of the travelling comfort and the derailment risk, so if its deviation exceeds a certain limit value, the travelling speed on that sector must be reduced. Therefore, railway

WIT Transactions on The Built Environment, Vol 114, © 2010 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/CR100631

geometry is both a measure of travelling quality and safety. For these reasons the French railway operator SNCF has been measuring periodically the geometrical characteristics of its high speed network since its commissioning, i.e. for more than 20 years now. Figure 1 shows the measurements of the longitudinal levelling (in French Nivellement Longitudinal, NL) for a 1 km track sector for the last 20 years. The NL parameter is the longitudinal mean deviation of rails in respect to the ideal position, and it is considered representative of the general railway geometry deterioration [1]. By default the deterioration grade increases with time, reflecting the track geometry deterioration. Due to confidentiality reasons, the measurement units are not shown. Degradation decrements take place only when some maintenance intervention is performed. In figure 1 the maintenance activities most relevant for track geometry are included: tamping interventions. In the figure, the bar heights represent the fraction of the railway sector affected by the maintenance activity. Tamping yields a visually obvious effect, yielding a sudden drop in NL.

The figure shows some very interesting behaviour: in the autumn of 2001 a tamping action has taken place. However, afterwards, an extremely fast degradation of the NL has set in. Some possible reasons for such counterproductive interventions are water under ballast, adverse weather conditions, or poor intervention quality (operator incompetence). This is a good demonstration of the stochastic characteristic of the ageing and restoration process. The effect of these characteristics on the proposed process model is that both the NL value after a tamping intervention (equation 1) and the degradation speed coefficient (equation 3) are modelled as normally distributed variables. Furthermore, in 2005 the NL improved several times (measurement line with negative gradient) but no tamping action was registered. Possible reasons for negative increments on NL without interventions are measurement errors, mainly offset errors, i.e. NL is not always measured on exactly the same 200 mts. Eventually, it could also

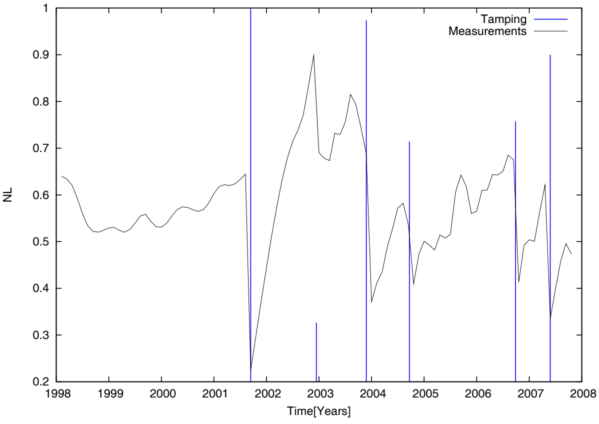

Figure 1: Course of longitudinal levelling degradation for a railway sector.

be that interventions are not registered in the database. The details of how this influences the model can be found in [2]. It can be shortly described as the addition of stochastic noise to represent the measurement errors, and the assumption of interventions when the negative increment is greater than a certain threshold.

Technical and human resources required for performing tamping interventions are a major cost factor in high speed railway systems [3]. Furthermore, due to high logistic costs constraints, most track geometry maintenance activities need to be planned up to one year in advance.

In this context, a crucial question to be answered is the following: with the available human and technical resources, and considering the current track railway geometry deviation, when and where should tamping interventions be performed? This paper presents a method for answering this question. It consists of two main components: a track geometry deterioration forecasting method, and a heuristic for interventions scheduling. Additionally, it needs a series of input data, such us a database with the available track geometry measurements, some characteristics of the tamping machines available, and the topology of the railway network to be maintained.

Section 2 presents the proposed forecasting method, section 3 describes the heuristic algorithm used for schedule generation, and in section 4 the method is validated be means of a case study with real data of a French high speed line. Finally 5 presents some concluding remarks.

2 Railway track geometry forecasting

2.1 Row data preprocessing

Railway geometry is measured periodically by means of special measuring coaches equipped with mechanical and/or electrical sensors. As it can be observed in figure 1, the periodicity of the measuring runs has been irregular since line commissioning, so the first problem for forecasting railway geometry deviation is the irregular sampling rate. To overcome this, we interpolate the measured points using splines, and then resample with the sampling rate of the last years. This is a compromise solution minimizing information loss in the last measurement years and keeping the addition of artificial measurements in the first years at an acceptable level. The resampled data is then used to tune the forecasting algorithm.

2.2 Process model

The process model used for forecasting is the one presented in [2]. It relies on 2 assumptions, namely:

1. The degradation value $NLinit_n$ achieved after the nth tamping intervention can be described as a normally distributed stochastic variable, i.e.

$$NLinit_n \sim \mathcal{N}(\mu_{NLinit}(n), \sigma_{NLinit}(n)). \tag{1}$$

2. The evolution of the degradation value between two tamping activities can be described by an exponential function of the form

$$NLinit_n \cdot e^{b_n \; t - t_n} \tag{2}$$

where n is the number of cumulated tamping interventions since track renewal, t_n is the time at which the nth tamping activity has taken place, and b_n is a is a normally distributed stochastic variable, i.e.

$$b_n \sim \mathcal{N}(\mu_b(n), \sigma_b(n)) \tag{3}$$

The first assumption relies on the study on the effects of tamping interventions on high speed railway lines presented in [1]. The second assumption is based on the model presented in [4], which postulates that geometry degradation grows exponentially between tamping interventions. According to these assumptions, for the model to be applied we would need to find expressions for $\mu_{NLinit}(n)$, $\sigma_{NLinit}(n)$, $\mu_b(n)$, and $\sigma_b(n)$. To obtain these functions, we need a database with track geometry measurements on many railway sectors for many years, including tamping activities performed. For each of the sectors recorded in the database, the curve $NLinit_n \cdot e^{b_n \; t - t_n}$ that best fits, i.e. minimizes the quadratic error for the measurements between the nth and the $(n+1)th$ tamping interventions, for $n \in 1, \ldots, N_{max}$, where N_{max} is the number of tamping interventions performed in the lapse of time recorded in the database for that sector. In doing so it must be taken into account that it is known that track geometry exhibits a transient behaviour in the first months after a tamping intervention, so we do not consider measurements taken in the three first months after an intervention. Doing this at each sector available in the database, the mean value and variance of $NLinit_n$ and b_n can be estimated. Furthermore, it is common knowledge that the degradation of NL depends on the annual track load rather than on time. In case that the track load had changed within the time period registered in the database, a transformation could be used to standardize the data, i.e. to unmake the effects of the track load modification, basing on the results presented in [5]. The next step is to find the functions $\mu_{NLinit}(n)$, $\sigma_{NLinit}(n)$, $\mu_b(n)$, and $\sigma_b(n)$ which best fit the estimated values.

2.3 Forecasting algorithm

To describe the forecasting procedure, a few definitions are necessary:

Defining $\widehat{NL}(t+h)$ as the forecast of NL at $t+h$ with the information available at time t the algorithm can be described as follows:

1. If the time elapsed since the last tamping intervention is longer than `TIME_MIN` and there is no intervention planned before time $t+h$, then find the function of the form of equation 2 which best fits the degradation curve since the last tamping intervention, and obtain $\widehat{NL}(t+h)$ by extrapolation.
2. If the time elapsed since the last tamping intervention is shorter than `TIME_MIN` and there is no intervention planned before time $t+h$, then

consider equation 2 with $b = b_n$, where n is the current cumulated number of tamping interventions, i.e. the mean value of b after n tamping interventions.
3. If a tamping intervention is planned before time $t+h$, then consider equation 2 with $b = b_n$, where n is the current cumulated number of tamping interventions, i.e. the mean value of b after $n + 1$ tamping interventions.

Summing up, the algorithm looks for the curve best fitting the geometry degradation course since the last tamping intervention, but if this was too recent it just takes the *mean* curve for the current number of accumulated tamping, according to the model of section 2.2. The same happens if a tamping intervention is planned within the forecasting horizon. For the parameter `TIME_MIN` a value of one year seems to be reasonable. For a more detailed description of the forecasting method see [6].

3 Interventions scheduling method

3.1 Problem definition

In order to formalize the problem definition, we model the railway net as a graph. The edges are the railway tracks and the nodes are the railway switches. The edges are in turn divided into sectors of 200 m. Then a criteria has to be established to assess the benefit of performing a tamping intervention at each of the sectors. In 3.2 some possible objective functions are presented. Furthermore, the following constrains has to be taken into account for scheduling tamping activities:

1. Tamping interventions take place in the night service interruptions, i.e. approx. 4 to 5 hours per night are available.
2. The number of tamping machines for the whole railway net is limited, so at each line the tamping machines are available for a limited number of nights per year, i.e. N nights. These nights are in general consecutive, so we call a tamping campaign the set of N consecutive nights at which a tamping machine is available for a given line or net.
3. Each tamping machine has limited travelling speed S_{Trav} and tamping speed $S_{Tamping}$.
4. T_{SetUp} is the preparation time needed between arrival at the starting sector and the intervention start time and $T_{TakeDown}$ the time needed between finalization of the intervention and departure to the end depot.
5. For tamping on and near switches special machines are needed. Furthermore the first and the last 200 m of a tamping intervention are transient sectors used to smooth the transit from a probably deteriorated sector to a fleshly tamped one. According to expert opinion, the number of transient sectors should be kept low. This leads to a further constraint: to minimize the number of transient sectors, each night tamping interventions can only take place in contiguous sectors, and they all must belong to the same edge (switches can not be tamped).

6. In order not to disturb normal train operations, the tamping machine must be allocated at a side track, the so called depots, before the first scheduled train runs. Additionally, on the fist intervention night the machine has to be picked up from a specific depot, and after the last intervention it has to be given back also at a specific depot.

The problem the heuristic scheduling algorithm solves is to find a feasible solution consisting a set of N interventions (one per night) which maximizes the defined objective function, which should be a mathematical representation of the railway operator's objectives. An intervention consists of the following elements:

- An intervention number $i, i \in 1, .., N$
- Start depot D_i and end depot D_e
- Start tamping sector S_i and end tamping sector S_e

Furthermore, for an intervention to be feasible, the inequality

$$T_{SI} \geq (Dist(D_i, S_i) + Dist(S_e, D_e))/S_{Travel} + T_{SetUp} + T_{TakeDown} + Dist(S_i, S_e)/S_{Tamping} \quad (4)$$

must hold, where T_{SI} is the night service interruption time, $Dist(D_i, S_i)$ is the distance between initial and end depot, $Dist(S_i, S_e)$ is the distance between initial and end intervention sectors. What inequality 4 expresses is that the blocking time must be enough for the maintenance team to travel with the machine to the intervention start sector, get ready to start working (duration of the procedure to block the track, T_{SetUp}), perform the intervention, get ready to leave the track (duration of the procedure to unblock the track, $T_{TakeDown}$) and travel to the end depot.

For the interventions to be unambiguously defined, an arbitrary sense is assigned to each edge, and the sectors are numbered in the sense of the edge. According to this sector enumeration, a further constraint can be set for an intervention to be valid: $S_i \geq S_e$.

3.2 Objective function

The objective function is a key part of the whole scheduling method. It should express the objectives of the railway track maintenance process, which may vary significantly from one company to another. Next three possible implementations are presented.

Total reduction track geometry deviation. The benefit of a tamping intervention is directly proportional to the current geometry degradation NL. This means that the degradation speed, i.e. $\frac{\partial NL}{\partial t}$, is not taken into account. This is the approach used in [7].

Expected time to failure. The benefit of a tamping intervention is inversely proportional to the time it is going to take to reach the maximal allowable geometry degradation value NL_{max}. This time to failure can be estimated using the forecasting method presented in section 2.

Expected *NL* at next campaign. Let h be the time between campaigns (in general one year, eventually six months). Suppose we are interested in finding out the optimal schedule for a campaign starting next week, i.e. at time $t = t$. Then an estimate of the value of NL at time $t = t + h$, i.e. $\widehat{NL}(t + h)$, could also be a measure of the benefit of tamping it. The more $\widehat{NL}(t + h)$ exceeds NL_{max}, the more value it would have to perform tamping next week. Likewise, the more NL_{max} exceeds $\widehat{NL}(t + h)$, the lesser it is worth to perform an intervention on that sector next week. The punctual forecasting $\widehat{NL}(t + h)$ represents the expected value. However, being the process model stochastic (see section 2.2) a confidence interval could also be included in the objective function.

In the case study presented in section 4 we use the expected NL at next campaign to enunciate the objective function.

3.3 Heuristic as an optimization method

In general, to find a solution for an optimization process a process model is used. When the model is highly complex and there is no standard optimization method, like in this case, there are two possibilities: to adapt the model for it to fit to a standard optimization method, or to adapt or create a new method to fit to the model.

In the literature some approaches to the railway track maintenance scheduling problem can be found, e.g. [7–9]. What [7] and [8] do is to adapt the process model by relaxing some constraints and then apply commercial linear programming optimization packages, as illustrated by approach *B* in figure 2. Our approach is more similar to [9]. We take the model as described in 2.2 and apply a heuristic algorithm, i.e. approach *C* in figure 2. The heuristic returns two results: an upper bound for the total solution value, and a feasible solution, namely the best one it has been able to find. The upper bound is a value which is guaranteed to be equal or greater than the optimal feasible solution. The results presented in section 4 show that the value of the best solution found is very near the upper bound (the difference is smaller than 1%), which gives us a hint of the heuristic's great performance potential.

3.4 Heuristic

According to the problem definition in 3.1, the heuristic can be described as follows:

1. Let an intervention be *Maximal* if it is feasible, i.e. equation 4 holds, and changing its S_e for the next sector, i.e. S_e would turn the intervention into infeasible, i.e. inequality 4 would no longer hold. The first step of the heuristic is to find for each edge i the set of all maximal interventions and calculate for each of them its value according to the objective function

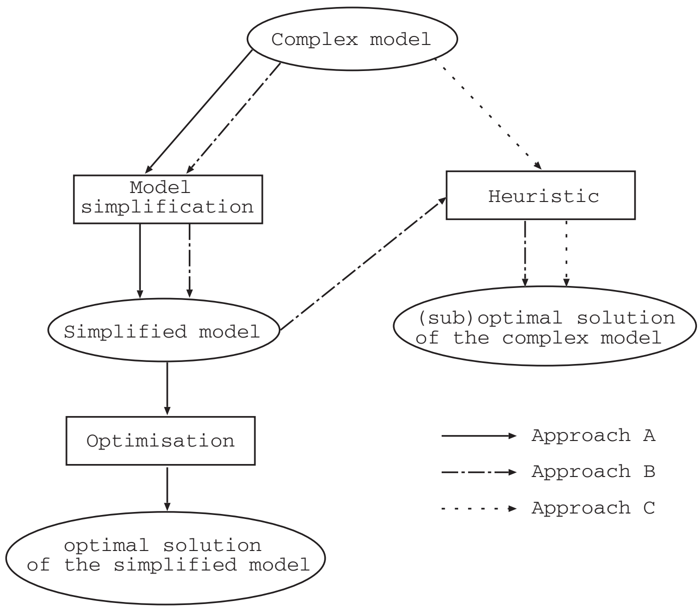

Figure 2: Some possible approaches to complex optimization problems.

described in 3.2. Remember that as stated in 3.1 all interventions must start and end in the same edge, i.e. in the same track, so the correspondence is unambiguous.

2. Let i^m be the number of edges of the graph representing the network. The second step is to find, for each edge $i \in 1, \ldots, i^m$ and each $n \in 1, \ldots, n^{m_i}$, where n^{m_i} is the number of interventions needed to tamp the whole edge i, the set $M_{i,n}$ consisting of n maximal interventions in edge i which maximizes the objective function. To put it in a nutshell, $M_{i,n}$ is the optimal solution if we only consider edge i and exactly n interventions are to be scheduled. This is the part of the heuristic requiring the most computational power, because at each edge i the set M_{i,n_0} may not be the set M_{i,n_0} plus some other intervention, but a completely different set, so for each $n \in 1, \ldots, n^{m_i}$ all possible combinations have to be explored. However, the fact that interventions are not allowed to have common sectors (that would mean performing an intervention twice in the same sector) keeps the number of combinations within an acceptable bound, even for edges with 300 sectors, as shown in the case study in 4.
3. The third step is to find the set of sets of maximal interventions $L = M_{i_1,n_1}$, $M_{i_2,n_2}, \ldots, M_{i_m,n_m}$ contained in the sets $M_{i,n}$ found in step 2, that maximizes the objective function, under the constrains that each set belongs to a different edge, i.e. $i_j \neq i_l \;\; \forall j, l \in 1, \ldots, m$, and the total number of interventions is equal to the number of interventions to be scheduled N, i.e.

$\sum_k^m n_k = N$. This step is quite straightforward, because in this case it holds that the solution for n days is contained in the solution for $n+1$ days. The set L is the set of disjointed (i.e. with no common sectors) feasible interventions which maximizes the objective function. The only additional requisites it has to fulfil to be a feasible solution is that the start depot of the first intervention and the end depot of the last intervention coincide with the specified ones (see 3.1), and that the end depot of each day equals the start depot of the next day, i.e. $D_i(j) = D_i(j+1) \ \forall j \in 1, \ldots, N-1$.

4. Let us define *bridge interventions* as interventions for which the start depot is not equal to the end depot, i.e. $D_i \neq D_e$. Because of the procedure used to calculate it, the solution L does not include any *bridge interventions*. Then the conversion would consist in finding a set of *bridge interventions* such that the start depot of the first intervention and the end depot of the last intervention are as specified, and that all depots included in solution L are visited at least once. This is nothing but the well-known travelling salesman problem. But a necessary condition to solve this problem it to know the cost of going from one node to another. To calculate this cost in this case is very difficult, because the number of possible combinations is enormous, so we choose to perform a local search. To assess the cost of introducing a *bridge intervention* from edge j to edge k, we do the following: for each $M_{i,n} \in L$, consider the n different sets which result of subtracting one single intervention to $M_{i,n}$. Then add the best possible *bridge intervention* from j to k to each of them. This will result in n different sets, each of them containing a *bridge intervention* from j to k. After doing this for all edges, the best solution, i.e. the one that maximizes the objective function, is chosen and the cost of going from depot j to depot k is the decrement of the objective function generated by the introduction of the *bridge intervention*. Being the described process merely a local search, we can not guarantee that the costs calculated are the minimum possible, but in practice this drawback is minimal, as illustrated in 4.
5. The fifth and last step of the heuristic is to convert the solution L into a feasible solution by solving the travelling salesman problem posed in step 4. This is done by means of the bench and bound method. This method has the advantage of finding the optimal solution without necessarily exploring the whole search tree. However, as the costs calculation described in step 4 may not be optimal, the solution achieved may as well not be optimal. But we can easily calculate how much better the solution could potentially be, because **the objective function value of the solution L is an upper bound for all feasible solutions**.

4 Case study

In this section we present an example of how the proposed scheduling method can be applied in reality. The problem characteristics are next described. Furthermore, the network is depicted by figure 3.

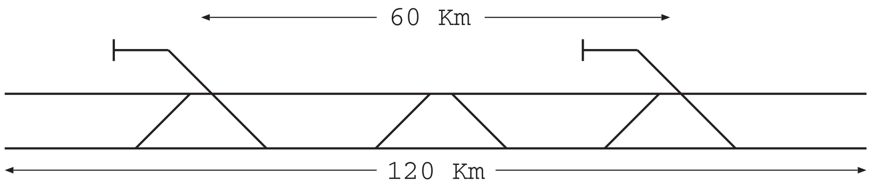

Figure 3: Railway network used for the case study.

- The modelled railways network consists of 120 km double way track, i.e. a total of 240 km track, with 3 double switches which divide the network into 8 tracks of about 30 km each.
- The network has 2 depots (secondary tracks where tamping machines are stationed during the day), being the distance between each other 60 km.
- Depot 2 must be the initial depot of the first intervention as well as the end depot of the last intervention.
- One tamping machine with a travelling speed of 80 km/h and a tamping speed of 1.4 km/h will be available for 20 nights.
- We have a database with track geometry data from the last 15 years for each track sector of 200 m, so we consider a total of 1200 sectors.

According to the problem definition in 3.1, the solution space can be calculated as

$$
\begin{aligned}
SolSpace &= (N_{Depots} \cdot N_{Sectors})^{N_{Nights}} \\
&= (2 \quad \cdot 240 \cdot 5) \quad \approx 10
\end{aligned}
\tag{5}
$$

This should clarify that exploring the whole solution space is simply out of the question.

The first step of the scheduling method is to define the objective function which best expresses the railway company interests. Therefore let S be a set of N scheduled interventions, and $TS = \{TS \ , TS \ , \ldots, TS_{max}\}$ the set of sectors included in S, i.e. the sectors for which a tamping intervention is scheduled. Also let f be the objective function. Then the objective function evaluated for S, $f(S)$ is defined as

$$
f(S) = \sum_{TS_1}^{TS_{max}} \widehat{NL}(t \ + h) \tag{6}
$$

where $\widehat{NL}(t \ + h)$ is the *expected NL at next campaign* as defined in 3.2. In our case study, tamping campaigns take place once a year, so $h = 1 \ year$.

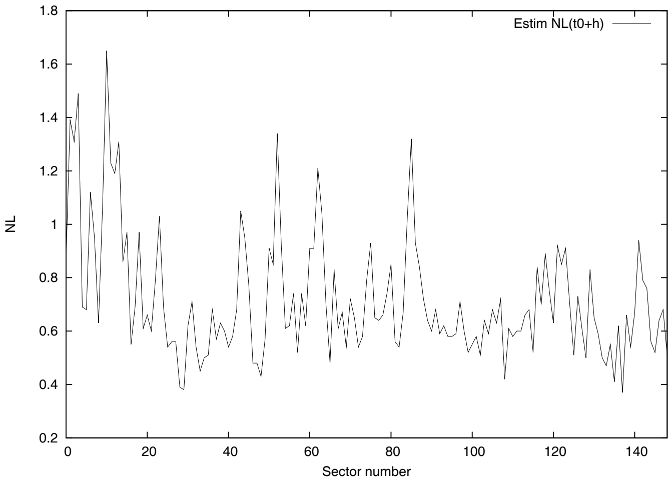

Figure 4: Expected one-year-ahead NL for one railway track.

Table 1: Obtained interventions schedule for the whole network.

Interv. Id	Track Id	D_i	D_e	S_i	S_e	Length	Value
1	1	1	1	1	25	25	7.8
2	1	1	1	29	54	26	8.56
3	2	1	1	196	223	28	7.79
4	2	1	1	236	262	27	9.39
5	5	1	1	626	654	29	8.14
6	5	1	1	664	690	27	7.72
7	5	1	1	700	725	26	6.54
8	6	1	1	864	892	29	8.26
9	**3**	**1**	**2**	**339**	**363**	**25**	**8.47**
10	3	2	2	366	392	27	8.49
11	3	2	2	394	421	28	11.8
12	3	2	2	422	450	29	13.43
13	4	2	2	451	480	30	7.93
14	4	2	2	493	520	28	9.99
15	4	2	2	531	557	27	10.46
16	7	2	2	919	949	29	8.99
17	8	2	2	1093	1118	26	7.01
18	8	2	2	1123	1149	27	8.66
19	8	2	2	1160	1188	29	9.77
20	**3**	**2**	**1**	**313**	**337**	**25**	**10.72**
Total value							179.92

The estimation $\widehat{NL}(t + h)$ is calculated by applying the forecasting method of section 2 to the database. Figure 4 shows the interpolated values of the estimation $\widehat{NL}(t + h)$ for the 150 sectors of one track. The curve is so irregular that significant differences can be appreciated even between contiguous sectors.

The interventions schedule obtained can be found in table 1. D_i and D_e are the start and end depots, respectively, and S_i and S_e are the initial and end intervention sectors, respectively. The length is expressed in sectors (each sector is 200 m long), and the value is calculated according to equation 6. In bold are the two *bridge* interventions, namely interventions 9 and 20. The reason why some interventions are longer than others (lengths vary between 25 and 30 sectors) is that some sectors are nearer to the depots than others, so the travelling times are shorter and then longer time is available for the tamping interventions themselves.

The best solution found has a value of 179.92, while the upper bound for the solution value was 180.18. This means that in the worst case our solution has 0.14% lesser value than the optimal solution. The heuristic has been implemented in C++ language, under the GNU/Linux operative system. **The calculation time is under 1 second** using a desktop PC with Pentium IV processor and 1 GB RAM memory.

5 Conclusions

In this work a heuristic based method for railway track tamping interventions scheduling has been presented. To our best knowledge, it is an innovative approach which goes beyond the state of the art both by incrementing the precision of the obtained interventions schedule and reducing dramatically the calculation time. This makes it possible to fine tune the maintenance strategy by evaluating the benefits or drawbacks of potential modifications in the maintenance process. Furthermore, the presented method could also be used to optimize the tamping in such a way that NL values are nowhere higher than a given NL_{max}. In fact this could be achieved by setting a non-continuous objective function, with a step at $NL = NL_{max}$. Future work includes the development of a Monte Carlo simulation environment for the railway ageing and restoration process, for integrated optimization of planning and scheduling of railway track maintenance processes.

References

[1] Meier-Hirmer, C., *Modèles et techniques probabilistes pour loptimisation des stratégies de maintenance. Application au domaine ferroviaire.* Ph.D. thesis, Université de Marne-la-Vallée, 2007.

[2] Quiroga, L. & Schnieder, E., Monte carlo simulation of railway track geometry deterioration and restoration. *Proceedings of the European Safety and Reliability Conference, ESREL 2010*, 2010.

[3] Esveld, C., *Modern railway track*. MRT-Productions, 2001.

[4] Veit, P., Track quality - luxury or necessity? *Railway Technical Review Special: Maintenance & Renewal*, 2007.

[5] Ubalde, L., López Pita, A., Teixeira, P., Bachiller, A. & Gallego, I., Track deterioration in high-speed railways: influence of stochastic parameters. *Proceedins of the Railway Engineering 2005, 8th International conference and exhibition*, 2005.

[6] Quiroga, L. & Schnieder, E., Modelling high speed railroad geometry ageing as a discrete-continuous process. *Proceedings of the Stochastic Modeling Techniques and Data Analysis International Conference, SMTDA 2010*, 2010.

[7] Oyama, T. & Miwa, M., Mathematical modeling analysis for obtaining an optimal railway track maintenance schedule. *Japan journal if industrial and applied mathematics: JJIAM*, **23(2)**, pp. 207–224, 2006.

[8] Oh, S., Lee, J., Park, B., Lee, H. & Hong, S., A study on a mathematical model of the track maintenance scheduling problem. *Computers in Railways X*, pp. 85–97, 2006.

[9] Lake, M., Ferreira, L. & Murray, M., Minimising costs in scheduling railway track maintenance. *Computers in Railways VII*, pp. 895–902, 2000.