Andrade, António Ramos and Teixeira, P. Fonseca

Statistical modelling of railway track geometry degradation using hierarchical Bayesian models

Statistical modelling of railway track geometry degradation using hierarchical Bayesian models

Andrade, A. R.[1] and Teixeira, P. F.[2]

Abstract

Railway maintenance planners require a predictive model that can assess the railway track geometry degradation. The present paper uses a hierarchical Bayesian model as a tool to model the main two quality indicators related to railway track geometry degradation: the standard deviation of longitudinal level defects and the standard deviation of horizontal alignment defects. Hierarchical Bayesian Models (HBM) are flexible statistical models that allow specifying different spatially correlated components between consecutive track sections, namely for the deterioration rates and the initial qualities parameters. HBM are developed for both quality indicators, conducting an extensive comparison between candidate models and a sensitivity analysis on prior distributions. HBM is applied to provide an overall assessment of the degradation of railway track geometry, for the main Portuguese railway line Lisbon-Oporto.

Keyworkds: Statistical model; railway track geometry; hierarchical Bayesian model; railway infrastructure

1 – Introduction

For railway Infrastructure Managers, predicting railway track geometry degradation is crucial to plan maintenance and renewal actions associated with it, and as become more and more relevant within their decision support systems. The use of more complex predictive models that tackle important aspects of railway track geometry degradation, namely the spatial correlations between degradation models' parameters, may enhance decision-making processes related to maintenance and renewal decisions, while preserving parsimonious in statistical modelling.

[1] Research Fellow, Institute of Railway Research, University of Huddersfield, UK.
Queensgate, Huddersfield, Yorkshire, HD1 3DH, UK. E-mail: a.r.andrade@hud.ac.uk
[2] Assistant Professor, CESUR, CEris, Instituto Superior Técnico, Universidade de Lisboa, Portugal.
Av. Rovisco Pais 1049-001 Lisboa, Portugal.

Bayesian statistical models provide a flexible framework to combine prior information from past samples or from expert judgment with the new data. Moreover, they allow specifying hierarchical probability structures that can capture uncertainty associated with a model parameter. They also provide a learning mechanism that can update information through time. Therefore, Bayesian statistical models seem a promising tool in transport infrastructure management, namely in railway infrastructure. The present paper intends to explore Hierarchical Bayesian Models (HBM) as a predictive tool to assess maintenance needs given a maintenance and renewal startegy.

This paper is structured in the following way: this first section briefly introduces the motivation and idea behind it, section 2 provides some background on railway track geometry degradation, focusing on current states of the Art and of the Practice, while identifying some current limitations in the statistical modelling of railway track geometry degradation. Section 3 introduces in a brief way Hierarchical Bayesian Models (HBM). Then, section 4 explores the statistical modelling of railway track geometry degradation with HBM, focusing namely on the model assumptions, on the definition of prior distributions and the derivation of the joint posterior distribution, and the description and comparison of different HBM specifications. Section 5 applies the HBM to a particular track segment of the main Portuguese line, with an extensive sensitivity analysis on the influence of prior distributions, and model comparison. Section 6 provides an overview of the railway track geometry degradation of the main Portuguese line using the best selected HBM model. Finally, section 7 discusses main conclusions and sketches directions for further research.

2 – Statistical modelling of railway track geometry degradation

2.1 – State of the Art

Research on the topic of statistical modelling of railway track geometry has benefitted from a number of contributions in the last three decades. First published works from the 80s focused on obtaining quantitative measures that described rail-vehicle performance through two components: i) ride comfort and ii) safety (probability of derailment). Corbin and Fazio [1] called them performance-based track-quality measures, and used simple linear models to obtain response modes for a given vehicle to some track geometry irregularities, and thus computed envelopes of allowed profile (longitudinal level defect) deviations for different speeds depending on the spatial frequency (cycles per meter) or on its inverse (i.e. the wavelength in meters). They used the power spectral density of rail profile (longitudinal level defect) as a statistical representation of railway track geometry, for ride comfort and safety purposes.

Nearly at the same time, Hamid and Gross [2] also discussed the need to objectively quantify track quality for maintenance planning and ride comfort purposes, analysing rail track geometry data collected over a period of one year on approximately 290 miles (467 km) of main-line, investigating statistical dependencies on track geometry defects (measured at 1-ft (0.3048-m) intervals by the T-6 vehicle) and some indicators (e.g. mean, 99th percentile, standard deviation and higher-order moments). Moreover, they developed empirical degradation models through linear autoregressive techniques that could describe the relationship between track quality indexes (defined as the standard deviation of profile) and physical parameters. For instance, simple linear regressions were put forward to relate the track quality index with the root mean square of vertical acceleration. Nevertheless, as the resulting empirical equations exhibited autoregressive terms, i.e. they included previous values as explaining variables, these expressions proved unreliable to predict track quality for the medium- or long-term. Development of rail track degradation models to predict future track quality indices for maintenance planning purposes had a major contribution with Bing and Gross [3], where they used a multiplicative form of model including as explaining variables: traffic information (equivalent train speed), track structure, maintenance (e.g. time since surfacing), or even ballast index in order to explain the rate of degradation at two consecutive time periods. Another important contribution was put forward by Hamid and Yang [4], in which they followed an approach based on analytically describing typical variations of track geometry, distinguishing random waviness, periodic behaviours at joints, and isolated variations, which occur occasionally but with regular patterns. In fact, they represented some track geometry defects as stationary random processes (modelled through the power spectral density) and others as periodic processes.

Besides these first published works, there were previous investigations carried out in the 70s by the former Office for Research and Experiments (ORE) trying to understand the fundamentals of the deterioration mechanism and to control this phenomenon, which examined data available from a number of administrations and showed that the factors governing the rate of deterioration were not obvious. They also showed that 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. Although the results were not very conclusive, track quality on relaying (i.e. the initial quality) was identified as the most important factor. More measurements on track geometry were recorded and other main qualitative conclusions were drawn [5]:

i) After the first initial settlement, both vertical quality and alignment deteriorate linearly with tonnage (or time) between maintenance operations;

ii) The rate of deterioration varies drastically from section to section even for apparently identical sections carrying the same traffic;

iii) There is no proved effect on the quality and on the rate of deterioration by the type of traffic or track construction;

iv) The rate of deterioration appears to be a constant parameter for a section regardless of the quality achieved by the maintenance machine;

v) Tamping machines improve the quality of a section of track to a more or less constant value.

Therefore, note that the second conclusion above emphasizes the importance of modelling degradation at the track section level, whereas the fourth conclusion (together with the second conclusion) makes clear that the inherent uncertainty of the degradation model parameters should be assessed with a learning mechanism associated with it.

Later on in the 90s, Iyengar and Jaiswal [6] proposed modelling railway track irregularities as a non-Gaussian model and concluded that in terms of level crossing and peak statistics, the proposed non-Gaussian model was consistently better than the Gaussian model in predicting the number of upward level crossings and peaks. Their non-Gaussian model to statistically represent railway track irregularities used a finite series of uncorrelated terms: the first term was a Gaussian process and the remaining terms were derived using a Gram-Schmidt procedure. Iyengar and Jaiswal [7] eventually modelled track irregularities (both Absolute Vertical Profile (longitudinal level defects) and unevenness (standard deviations)) as a stationary Gaussian random field so that the classical level-crossing and peak-statistics theory could be exploited to relate the sample deviation to the highest peak value in a simpler way than in their first non-Gaussian model. In a way, Iyengar and Jaiswal opted to simplify their initial approach so that they could make probabilistic approaches appealing to practical engineers. Nevertheless, as it was discussed by Kumar and Stathopoulos [8], unevenness data (standard deviations) indicated a non-Gaussian character with skewness and kurtoses significantly different from the typical values for Gaussian processes.

More recently, other approaches have also tried to capture the nonlinear characteristics of track quality deterioration [9] [10] [11], though they typically would not consider all track geometry defects, and would model instead a quality index. An important contribution was given in [12] for the case of a Spanish high speed railway line, identifying the embankment height as a dependent variable on the density of maintenance works, though the study analyses the past tamping actions and vertical accelerations rather than the track geometry records themselves. In that sense, that work is analysing the 'outputs' of maintenance decisions, rather than their inputs (i.e. degradation) so that one may assess whether or not that decision was a good one. Moreover, another relevant issue in statistical modelling of railway track geometry is how to model the tamping/maintenance recuperation, i.e. the impact of levelling and tamping operations in the railway track geometry defects and their evolution. In fact, specialized literature usually does not cover this improvement in great detail and only a few references have been found, such as [11] and [13]. A recent work by Quiroga and Schnieder [14] has used accumulated tamping interventions as an explaining variable, showing higher variances for higher number of accumulated tamping interventions. Furthermore, a recent work by Vale and Lurdes [16] also discussed a stochastic model for the geometrical railway track degradation process, focusing on the standard deviation of longitudinal level defects and not on the standard deviation of horizontal alignment. Other contributions focusing on different prediction techniques used to assess future railway track geometry condition have been proposed, namely using artificial neural networks [17], stochastic state space methods [18] or even Petri net models [19]. A Bayesian approach is explored in [20], but this time following a nonparametric specification with a Dirichlet Process Mixture Model, focusing on the failure of different railway components rather than specifically on railway track geometry degradation. Finally, a very recent work by Gong et al. [21] has put more focus on the deterioration of lateral alignment using vehicle dynamic simulation and considering an elastic lateral model for the railway track, while assessing the effect of different factors like the vehicles, running speed, traffic mixes and different wheel/rail contacts.

2.2 – State of the Practice

According to a best practice guide for optimum track geometry durability [22], European Infrastructure Managers tend to trigger their preventive tamping actions based on a single indicator: the standard deviation of longitudinal level defects. Nevertheless, the European standard EN 13848-5 puts forward recommended Alert Limits for preventive maintenance actions based on two indicators: i) the standard deviation of longitudinal level defects and also ii) the standard deviation of horizontal alignment defects.

In fact, recent research has discussed the use of these two indicators as predictors of other impacts associated with planning maintenance of railway track geometry, namely the corrective/unplanned maintenance needs and the delays imposed due to temporary speed restrictions. Both works [23] [24] have found that not only the standard deviation of longitudinal level defects was a statistically significant predictor, but also the standard deviation of horizontal alignment defects. Therefore, the present paper intends to statistically model the evolution of these two quality indicators relative to railway track geometry using a HBM.

The degradation of railway track geometry is usually quantified by seven track geometry defects: the left and right longitudinal level defects (LLL and RLL), the left and right horizontal alignment defects (LHA and RHA), the cant defects (C), the gauge deviations (GD) and the track twist (T). These defects are measured through automated measuring systems, typically integrated in inspection vehicles, and saved as signal data. Signal digital processing techniques are then used to align signals and derive indicators (for each type of defect) that can support maintenance and renewal decisions as part of planned maintenance or eventually as unplanned maintenance actions. Many Infrastructure Managers tend to combine all these indicators into a track quality index, which is typically a function of the standard deviations of each defect and/or train permissible speed (as reported in [25] or [26]). Nevertheless, the standard deviation for the short wavelength (3m - 25m) of longitudinal level defects is still perceived as the crucial indicator for planned maintenance decisions for many European Infrastructure Managers.

The use of the standard deviation of the short wavelength (3m - 25m) of longitudinal level defects (SD_{LL}) as the crucial indicator for planned maintenance on railway track geometry degradation may be attributed in our perspective to two reasons: i) due to the simplicity in the empirical expressions describing its evolution (which depends on the accumulated tonnage, usually in Million Gross Tons (MGT)), and ii) due to the fact that it correlates well with the vertical force [13] [27], which is a proxy or vertical acceleration felt by the passenger and thus, of ride quality.

Railway track geometry defects should be within certain limits according to a given safety standard. The European Standard EN 13848-5 [28] provides limits for several indicators for each type of defect depending on the maximum permissible speed and for three main levels:

- IAL – Immediate Action Limit: refers to the value which, if exceeded, requires imposing speed restrictions or immediate correction of track geometry;
- IL – Intervention Limit: refers to the value which, if exceeded, requires corrective maintenance before the immediate action limit is reached;
- AL – Alert Limit: refers to the value which, if exceeded, requires that track geometry condition is analysed and considered in the regularly planned maintenance operations.

Although the IAL limits are considered normative, providing the highest admissible limits to ensure safety and ride comfort; the IL and the AL limits are purely indicative, reflecting common practice among most European Infrastructure Managers. They are even expressed as a range rather than a single value. In fact, the EN 13848-5 also directs each Infrastructure Manager to select their own IL and AL limits according to their inspection and maintenance systems, which in turn relate to different targets for safety, ride quality, lower life-cycle costs and track access availability.

2.3 – Current limitations in statistical modelling of railway track geometry degradation

Current statistical approaches tend to focus on track quality indexes rather than on the standard deviations of longitudinal level defects (SD_{LL}) and of horizontal alignment defects (SD_{HA}). These track quality indexes are sometimes dependent on the maximum permissible speed so they do not refer only to railway track physical degradation but also to its use. The statistical approaches that model the standard deviation of longitudinal level defect (SD_{LL}) do not consider the standard deviation of horizontal alignment defects (SD_{HA}), arguing that current practice of maintenance decisions rely solely on the SD_{LL} [22]. Nevertheless, the SD_{HA} indicator seems to play an important role as a predictor of localized defects and corrective maintenance needs.

Moreover, current statistical models have overlooked the spatial correlations of the deterioration rates and the initial qualities for consecutive track sections. In fact, this idea followed from an initial exploratory work previously conducted in [15], showing that spatial correlation between deterioration rates and initial quality were statistically significant. In that sense, none of the previous statistical models takes advantage of these spatial correlations between deterioration rates and the initial qualities for consecutive track sections, in order to improve the current predictive models. In the present paper, these spatial correlations are handled using HBM, so that the parameters for the deterioration rates and the initial qualities (i.e. the slope β and the y-intercept α in simple linear regression models $y = \alpha + \beta . x + \varepsilon$) can be considered random quantities, and thus, Conditional Autoregressive (CAR) probability structures can be assigned to these parameters associated with consecutive track sections . Note that in classical statistical approaches, the spatial correlations would have to be tackled through CAR structures assigned to the random error ε, as the slope β and the y-intercept α are not random, and in fact they are assumed to be fixed but unknown, and estimated from the data. And, therefore as a result, HBM is the mathematical statistical method that let us model directly the spatial correlation between the deterioration rates and initial qualities and not on the random error.

Our HBM approach models separately the two main indicators (SD_{LL} and the SD_{HA}) and adds CAR probability structures to the deterioration rates and the initial qualities parameters to handle the previously overlooked spatial correlations between consecutive tracks. As it will be seen later on, this proved to provide a better fit to the data according to the Deviance Information Criterion (DIC).

3 – A brief note on Hierarchical Bayesian Models (HBM)

This section briefly explores Hierarchical Bayesian Models (HBM) to predict the evolution of railway track geometry degradation. Bayesian models are different than classical statistical models in the fact that they assume parameters as random variables, whose uncertainty can be quantified by a prior distribution. This prior distribution

$p(\theta)$ is then combined with the traditional likelihood $p(y|\theta)$ to obtain the posterior distribution of the parameters of interest. The posterior distribution $p(\theta|y)$ of the parameters θ given the observed data y can be computed according to Bayes' rule as:

$$p(\theta|y) = \frac{p(y|\theta).p(\theta)}{\int p(y|\theta').p(\theta')\,d\theta'} \propto p(y|\theta).p(\theta)$$

The specification of the prior distribution constitutes a very important step in any Bayesian model, using for instance a non-informative (or vague prior), or incorporating preceding known information using old samples (hopefully under the same boundary conditions) or from expert judgment techniques. Further details on Bayesian statistics can be found in [29] and [30], or for a more practical approach [31]. However, in almost every case in real applications, one finds that that the joint posterior distribution $p(\theta|y)$ has a reasonably high dimension, and integration through numerical methods must rely on Markov Chain Monte Carlo (MCMC) methods, which are built in such a way that their stationary distribution is the desired posterior distribution.

HBM are a special case of Bayesian models. They benefit from a major property: they can be factorized through Directed Acyclic Graphs (DAG) in a convenient way. This not only enables arbitrarily complex full probability models to be specified based on the simple local components, but it also makes the identification of full conditional distributions straightforward. Moreover, this hierarchical construction is particularly useful, because once the full conditional distributions are identified/available, one can sequentially sample from them using the Gibbs sampler, and sample from the posterior distribution of interest, or using any more general sampling method (e.g. adaptive rejection sampling, Metropolis-Hastings algorithm).

A DAG is a graph that is directed because each node is linked through an arrow (i.e. with a defined direction) and it is acyclic because no cycles are formed with the arrows (i.e. it is not possible to return to a given node once you left it). Parent nodes from a given node (i.e. parents[v]) are all nodes that are connected through arrows from them to that node 'v'; whereas children nodes from a given node (i.e. children[v]) are all nodes that are connected through arrows from that node 'v' to them.

According to Lunn et al. [32] on the WinBUGS use of DAG factorization, let V denote the set of all nodes (v) in a DAG for a given HBM, it can be shown that:

$$p(\theta|y) \propto p(\theta, y) = p(V) = \prod_{v \in V} p(v|\text{parents}[v])$$

Let $V \setminus v$ denote 'all elements of V except v'. The full conditional $p(v \mid V \setminus v)$ is then proportional to the product of terms in $p(V)$ which contain v:

$$p(v \mid V \setminus v) \propto p(v|\text{parents}[v]) \times \prod_{w \in \text{children}[v]} p(w|\text{parents}[w])$$

Note that OpenBUGS – the most recent development of WinBUGS, allows user-friendly inference for these HBM, especially if the adaptive rejection sampling algorithm from Gilks and Wild [33] is needed.

4 – Modelling railway track geometry degradation with HBM

Having briefly introduced the main concept of HMB in section 3, let us explore the intricacies of the proposed modelling of railway track geometry degradation using HBM. The associated DAG for the more complicated models is provided at the end of this section. We start by discussing the main assumptions on statistical modeling of railway track geometry degradation that our modelling approach will rely on, without forgetting the past research findings previously discussed in subsections 2.1 and 2.2. The models developed in this section are specified in the same way for the main two quality indicators related to railway track geometry degradation, i.e. for the standard deviation of longitudinal level defects (SD_{LL}) and for the standard deviation of horizontal alignment defects (SD_{HA}). In the following, when we refer to the dependent variable as a quality indicator, the reader should have in mind that the dependent variable can be either the SD_{LL} or the SD_{HA}.

4.1 – Model assumptions

Let us start by assuming that a quality indicator $(\mathrm{y_{svkl}})$ at inspection l for track section k from track segment v from area s is normally distributed with mean $\mathrm{m_{svkl}}$ and variance σ_s^2, i.e. $\mathrm{y_{svkl}} \sim \mathrm{N}(\mathrm{m_{svkl}}, \sigma_s^2)$. Figure 1 should support the reader to understand better the meaning of indices s, v, and k for a typical double track line.

Then, some assumptions on its mean value result from a combination of factors:

1) A constant linear evolution with accumulated tonnage (T_{svkl} – accumulated tonnage since last tamping or renewal operation), given by the deterioration rate - β_{svk} (slope of a linear regression), which assumes different values for each track section k in track segment v for area s.
2) An initial value for the quality indicator, given by the initial quality - α_{svk} (y-intercept of a linear regression), which assumes different values for each track section k in track segment v for area s.

These first two assumptions presume a linear evolution of the quality indicator with accumulated tonnage since last maintenance or renewal action and an initial quality, allowing different values for each track section.

3) A disturbance effect (δ_{sv}) of the quality indicator after each tamping operation, i.e. the quality indicator does not recover to its initial value α_{svk}, but it is affected by a rate $1+\delta_{sv}$, given by δ_{sv}, which assumes different values for each track segment v for area s. Therefore, note that at each new tamping cycle the initial quality would be $\alpha_{svk}(1+\delta_{sv})^{N_{svkl}}$, in which N_{svkl} is the number of tamping operations conducted since last renewal.
4) Distinction between renewed track sections ($R_{svkl}=1$) and non-renewed track sections ($R_{svkl}=0$) is assured through the separation of the initial quality and the deterioration rates for a non-renewed track section k from segment v for area s – α'_{svk} and β'_{svk} respectively; whereas the disturbance effect (δ_{sv}) is considered the same for renewed and non-renewed track sections.

These next two assumptions provide an exponential effect of the number of tamping operations conducted since last renewal and different parameters for renewed and non-renewed track sections.

These four assumptions can be compiled in a single mathematical expression for the mean m_{svkl} of the quality indicator $\mathrm{y}_{\mathrm{svkl}}$ as:

$$m_{svkl} = [\alpha_{svk}(1+\delta_{sv})^{N_{svkl}} + \beta_{svk}T_{svkl}] \cdot R_{svkl} + [\alpha'_{svk}(1+\delta_{sv})^{N_{svkl}} + \beta'_{svk}T_{svkl}] \cdot (1-R_{svkl})$$

In this expression, we should regard $\alpha, \beta, \alpha', \beta'$ and δ as parameters, to which there should be assigned a hierarchical probability structure, whereas N, T and R should be regarded as known/explaining variables. Figure 2 provides a graphical representation of the intended behaviour of the mean of the quality indicator expressed by the equation above (m_{svkl}).

4.2 – Introduction of spatial correlations through CAR probability structures

In this subsection, we discussed two CAR probability structures to handle the spatial correlation between the deterioration rates and the initial qualities of consecutive track sections.

In order to be parsimonious in modelling, i.e. trying to keep the number of parameters to a minimum, a good strategy is considering (Gaussian) Conditional Autoregressive (CAR) terms. Besag [34] showed in 1974 that conditional probability structures could deal with spatial interactions in the more complex structures. The CAR model is used to model the spatial dependencies in deterioration rates and initial qualities for consecutive track sections. For example, take the initial quality $\mathrm{\alpha_{svk}}$ for a given track section k in segment v in area s, we will consider that $\mathrm{\alpha_{svk}}$ is a combination of two additive components: an average value $\mathrm{\alpha_{sv}}$ and a spatially correlated term $\varepsilon_{\mathrm{\alpha_{svk}}}$ so that $\mathrm{\alpha_{svk}} = \mathrm{\alpha_{sv}} + \varepsilon_{\mathrm{\alpha_{svk}}}$. For $\varepsilon_{\mathrm{\alpha_{svk}}}$, we then assign a CAR probability structure such as $\varepsilon_{\mathrm{\alpha_{svk}}} \middle| \boldsymbol{\varepsilon}_{\boldsymbol{\alpha}_{\mathbf{sv(-k)}}}, \sigma_{\alpha_{sv}}^2 \sim N\left(\bar{\varepsilon}_{\mathrm{\alpha_{svk}}}, \sigma_{\alpha_{sv}}^2 / n_{svk}\right)$, in which $\bar{\varepsilon}_{\mathrm{\alpha_{svk}}} = \sum_{j \in \mathcal{N}_{svk}} \varepsilon_{\mathrm{\alpha_{svj}}} / n_{svk}$, $\mathcal{N}_{svk}$ denotes the set of track sections which are considered neighbors to track section k (in segment v in area s), and n_{svk} is the number of neighbors of track section k (in segment v in area s), and finally $\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}_{\mathbf{sv(-k)}}}$ is the vector with all components $\varepsilon_{\mathrm{\alpha_{svk}}}$ from segment v in area s except the component related to track section k.

Two well-known CAR structures were tested: the first-order random walk $\left(RW(1)\right)$ and the second-order random walk $\left(RW(2)\right)$ as hierarchical structures for parameters $\alpha, \beta, \alpha', \beta'$.

The first-order random walk $\left(RW(1)\right)$ is defined by considering as neighbour structure $n_{svk} = 1$ for $k = 1$ and $k = K_s$, and $n_{svk} = 2$ for $k = 2, \dots, K_s - 1$, and $\bar{\varepsilon}_{\mathrm{\alpha_{svk}}} = \varepsilon_{\mathrm{\alpha_{sv(k+1)}}}$ for $k = 1$, $\bar{\varepsilon}_{\mathrm{\alpha_{svk}}} = \left(\varepsilon_{\mathrm{\alpha_{sv(k-1)}}} + \varepsilon_{\mathrm{\alpha_{sv(k+1)}}}\right)/2$ for $k = 2, \dots, K_s - 1$, and $\bar{\varepsilon}_{\mathrm{\alpha_{svk}}} = \varepsilon_{\mathrm{\alpha_{sv(k-1)}}}$ for $k = K_s$.

For the second-order random walk $\left(RW(2)\right)$ these expressions complicate a little more but they can be derived from the following equivalence, where the symbol * denotes the different possible parameters $\alpha, \beta, \alpha', \beta'$:

$$\varepsilon_{*_{svk}} \sim \mathrm{RW}(2) \;\Leftrightarrow\; \varepsilon_{*_{svk}} \middle| \varepsilon_{*_{sv_{-k}}}, \sigma_*^2 \;\sim \begin{cases} \mathrm{N}\left(2\,\varepsilon_{*_{sv(k+1)}} - \varepsilon_{*_{sv(k+2)}}, \sigma_*^2\right), \mathrm{k} = 1 \\ \mathrm{N}\left(\left[2\,\varepsilon_{*_{sv(k-1)}} + 4\,\varepsilon_{*_{sv(k+1)}} - \varepsilon_{*_{sv(k+2)}}\right]/5\,, \sigma_*^2/5\right), \mathrm{k} = 2 \\ \mathrm{N}\left(\left[-\,\varepsilon_{*_{sv(k-2)}} + 4\,\varepsilon_{*_{sv(k-1)}} + 4\,\varepsilon_{*_{sv(k+1)}} - \varepsilon_{*_{sv(k+2)}}\right]/6\,, \sigma_*^2/6\right), \mathrm{k} = 3, \ldots, \mathrm{K_s} - 2 \\ \mathrm{N}\left(\left[-\,\varepsilon_{*_{sv(k-2)}} + 4\,\varepsilon_{*_{sv(k-1)}} + \varepsilon_{*_{sv(k+1)}}\right]/5\,, \sigma_*^2/5\right), \mathrm{k} = \mathrm{K_s} - 1 \\ \mathrm{N}\left(-\,\varepsilon_{*_{sv(k-2)}} + 2\,\varepsilon_{*_{sv(k-1)}}, \sigma_*^2\right), \mathrm{k} = \mathrm{K_s} \end{cases}$$

4.3 – Definition of prior distributions

After assigning a CAR probability structure to tackle spatial correlations in a parsimonious way, we define the prior distributions for all other parameters. Therefore, for the parameter δ_{sv}, i.e. the disturbance effect of the initial quality after each tamping operation, we define a typical probability structure expressing vague information on that parameter, i.e. $\delta_{sv} \sim N(0,\ \sigma_\delta^2)$. Moreover, for each variance component in each hierarchical structure, we finalize by assigning inverse gamma distributions to each component, i.e. $\sigma_s^2 \sim IG(c_0, d_0)$, $\sigma_{\delta_s}^2 \sim IG(c_1, d_1)$, $\sigma_{\alpha_{sv}}^2 \sim IG(c_2, d_2)$, $\sigma_{\beta_{sv}}^2 \sim IG(c_3, d_3)$, $\sigma_{\alpha'_{sv}}^2 \sim IG(c_4, d_4)$ and $\sigma_{\beta'_{sv}}^2 \sim IG(c_5, d_5)$, where $IG(c, d)$ denotes an inverse gamma distribution with shape parameter c and scale parameter d, whose density is proportional to $x^{-(c+1)} \exp\left(-\frac{d}{x}\right)$, $x > 0$. For all models (M1, M2, M3 and M4), inverse gamma priors $IG(0.5, 0.0005)$ were assigned for each variance parameter (σ_s^2, $\sigma_{\alpha_{sv}}^2$, $\sigma_{\beta_{sv}}^2$, $\sigma_{\alpha'_{sv}}^2$, $\sigma_{\beta'_{sv}}^2$ and $\sigma_{\delta_s}^2$). Note that flat priors are improper distributions, i.e. do not integrate to one, but assume a constant value everywhere, attempting to describe vague or no prior information on that parameter.

As a side note, it is important to mention that the use of inverse gamma prior distributions for the variance terms is not free of criticism and has been discussed by Gelman [35], who recommended starting with a non-informative uniform prior density, or when more prior information is desired, working within the half-t family of prior distributions, which are more flexible and have better behaviour near 0, compared to the inverse-gamma family, which are very sensitive in data sets where the variance terms may assume low values of σ^2. However, the choice of assigning inverse gamma distributions 'is an attempt at non-informativeness within the conditional conjugate family' [35], which mainly translates into full conditional posterior distributions for each variance component within the same distributional family, i.e. also inverse gamma distributions, as later seen in the Appendix for i), iii), v), vii), ix) and xi). This choice is not only attractive for pedagogical purposes, but it is also a common choice in many BUGS software applications.

Figure 3 provides a DAG representing the proposed HBM, focusing in 3 models explored later on: models M2, M3 and M4. Model M1 has a simple representation without the eight upper nodes (σ_*^2 and $\varepsilon_{*_{svk}}$ for $* = \{\alpha, \beta, \alpha', \beta'\}$). Stochastic nodes are represented in ovals with solid line, whereas constants are in rectangular boxes. Hyperparameters (c, d) of the variance components are not represented for clarity and ovals with dashed contours are logically dependent on their parent nodes and are not stochastic.

Note that the definition of the proposed HBM with completely independent DAG for each segment s facilitates simulation in a separate manner for each segment s. This is guaranteed by defining hyperparameters - σ_s^2, $\sigma_{\delta_s}^2$, $\sigma_{\alpha_{sv}}^2$, $\sigma_{\beta_{sv}}^2$, $\sigma_{\alpha'_{sv}}^2$ and $\sigma_{\beta'_{sv}}^2$ that are each different for each area s. Of course, a more parsimonious modelling option could have been chosen, but then the simulation could not be conducted separately for each track segment s (which favours for instance the possibility of applying parallel processing techniques in the future).

4.4 – Joint posterior distribution

To derive the joint posterior distribution, prior independence is assumed amongst the model parameters so that the joint posterior density is then proportional to:

$$\prod_{s=1}^{S}\prod_{v=1}^{V_s}\prod_{k=1}^{K_s}\prod_{l=1}^{L_s}\left\{\frac{1}{\sigma_s}\exp\left(-\frac{1}{2}\left(\frac{y_{svkl}-m_{svkl}}{\sigma_s}\right)^2\right)\right\}\cdot\prod_{s=1}^{S}\left\{\left(\frac{1}{\sigma_s^2}\right)^{c_0+1}\exp\left(-\frac{1}{\sigma_s^2}d_0\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\left\{\frac{1}{\sigma_{\delta_s}}\exp\left(-\frac{1}{2}\left(\frac{\delta_{sv}}{\sigma_{\delta_s}}\right)^2\right)\right\}\cdot\prod_{s=1}^{S}\left\{\left(\frac{1}{\sigma_{\delta_s}^2}\right)^{c_1+1}\exp\left(-\frac{1}{\sigma_{\delta_s}^2}d_1\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\prod_{k=1}^{K_s}\left\{\frac{\sqrt{n_{svk}}}{\sigma_{\alpha_{sv}}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{\alpha_{svk}}-\bar{\varepsilon}_{\alpha_{svk}}}{\sigma_{\alpha_{sv}}}\sqrt{n_{svk}}\right)^2\right)\right\}\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\left\{\left(\frac{1}{\sigma_{\alpha_{sv}}^2}\right)^{c_2+1}\exp\left(-\frac{1}{\sigma_{\alpha_{sv}}^2}d_2\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\prod_{k=1}^{K_s}\left\{\frac{\sqrt{n_{svk}}}{\sigma_{\beta_{sv}}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{\beta_{svk}}-\bar{\varepsilon}_{\beta_{svk}}}{\sigma_{\beta_{sv}}}\sqrt{n_{svk}}\right)^2\right)\right\}\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\left\{\left(\frac{1}{\sigma_{\beta_{sv}}^2}\right)^{c_3+1}\exp\left(-\frac{1}{\sigma_{\beta_{sv}}^2}d_3\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\prod_{k=1}^{K_s}\left\{\frac{\sqrt{n_{svk}}}{\sigma_{\alpha'_{sv}}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{\alpha'_{svk}}-\bar{\varepsilon}_{\alpha'_{svk}}}{\sigma_{\alpha'_{sv}}}\sqrt{n_{svk}}\right)^2\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\left\{\left(\frac{1}{\sigma_{\alpha'_{sv}}^2}\right)^{c_4+1}\exp\left(-\frac{1}{\sigma_{\alpha'_{sv}}^2}d_4\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\prod_{k=1}^{K_s}\left\{\frac{\sqrt{n_{svk}}}{\sigma_{\beta'_{sv}}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{\beta'_{svk}}-\bar{\varepsilon}_{\beta'_{svk}}}{\sigma_{\beta'_{sv}}}\sqrt{n_{svk}}\right)^2\right)\right\}$$

$$\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\left\{\left(\frac{1}{\sigma_{\beta'_{sv}}^2}\right)^{c_5+1}\exp\left(-\frac{1}{\sigma_{\beta'_{sv}}^2}d_5\right)\right\}\cdot\prod_{s=1}^{S}\prod_{v=1}^{V_s}\{P[\alpha_{sv}]\cdot P[\beta_{sv}]\cdot P[\alpha'_{sv}]\cdot P[\beta'_{sv}]\}$$

In which:

$$\mathrm{m_{svkl}} = [\alpha_{\mathrm{svk}}(1+\delta_{\mathrm{sv}})^{\mathrm{N_{svkl}}}+\beta_{\mathrm{svk}}\mathrm{T_{svkl}}]\cdot\mathrm{R_{svkl}}+\left[\alpha'_{\mathrm{svk}}(1+\delta_{\mathrm{sv}})^{\mathrm{N_{svkl}}}+\beta'_{\mathrm{svk}}\mathrm{T_{svkl}}\right]\cdot(1-\mathrm{R_{svkl}})$$

$$\alpha_{\mathrm{svk}} = \alpha_{\mathrm{sv}}+\varepsilon_{\alpha_{\mathrm{svk}}};\ \alpha'_{\mathrm{svk}} = \alpha'_{\mathrm{sv}}+\varepsilon_{\alpha'_{\mathrm{svk}}};\ \beta_{\mathrm{svk}} = \beta_{\mathrm{sv}}+\varepsilon_{\beta_{\mathrm{svk}}} \text{ and } \beta'_{\mathrm{svk}} = \beta'_{\mathrm{sv}}+\varepsilon_{\beta'_{\mathrm{svk}}}$$

In order to ensure that the CAR model structures are identifiable, we follow the typical constraint suggested by Besag and Kooperberg [36] that is to impose that the $\sum_k \varepsilon_{\alpha_{\mathrm{svk}}} = 0$, and use a flat prior for the constant α_{sv} on the whole real line. Note that these CAR probability structures are also adopted for β, α' and β' parameters.

As the joint posterior is rather complex, we present the full conditional posterior distribution in the appendix, so that a Gibbs sampling strategy can iteratively draw for each parameter and use them as current values for each conditional posterior distribution.

4.5 – Different model specifications and comparison

For each quality indicator (SD_{LL} and SD_{HA}), mainly 4 models (M1, M2, M3 and M4) are explored whose differences are related with hierarchical structures for α, β, α′ and β′:

- Model M1 assigns simple flat priors to each parameter $\alpha_{sv}, \beta_{sv}, \alpha'_{sv}, \beta'_{sv}$;
- Model M2 assigns a flat prior to each parameter $\alpha_{sv}, \beta_{sv}, \alpha'_{sv}, \beta'_{sv}$ and a normal prior with mean equal to zero and variance $\sigma^2_{\alpha_{sv}}, \sigma^2_{\beta_{sv}}, \sigma^2_{\alpha'_{sv}}, \sigma^2_{\beta'_{sv}}$ for each parameter $\varepsilon_{\alpha_{svk}}, \varepsilon_{\beta_{svk}}, \varepsilon_{\alpha'_{svk}}, \varepsilon_{\beta'_{svk}}$, respectively.
- Model M3 assigns a flat prior to each parameter $\alpha_{sv}, \beta_{sv}, \alpha'_{sv}, \beta'_{sv}$ and a first-order random walk for each parameter $\varepsilon_{\alpha_{svk}}, \varepsilon_{\beta_{svk}}, \varepsilon_{\alpha'_{svk}}, \varepsilon_{\beta'_{svk}}$;
- Model M4 assigns a flat prior to each parameter $\alpha_{sv}, \beta_{sv}, \alpha'_{sv}, \beta'_{sv}$ and a second-order random walk to each parameter $\varepsilon_{\alpha_{svk}}, \varepsilon_{\beta_{svk}}, \varepsilon_{\alpha'_{svk}}, \varepsilon_{\beta'_{svk}}$.

In terms of model comparison, HBM tend to rely on a standard measure proposed by Spiegelhalter et al. [38] for comparing models of any degree of complexity. In fact, the Deviance Information Criterion (DIC), whose assessment is straightforward using MCMC inference methods, balances two components: the expected posterior deviance (i.e. the goodness of fit) and the effective number of parameters (i.e. the complexity of the model). It is defined as: $\mathrm{DIC} = 2\overline{\mathrm{D}(\theta)} - \mathrm{D}(\bar{\theta})$, where $\overline{\mathrm{D}(\theta)}$ is the posterior mean of deviance and $\bar{\theta}$ is the posterior mean of the model parameters. In practical terms, a lower DIC value means a better model in terms of relative comparison.

5 – HBM application to a particular area of the main Portuguese line

This section will analyse some detailed results from the application of the proposed HBM on a particular area/sample from the historical data from the main Portuguese line (Lisbon-Oporto).

5.1 – A brief description of the sample

This historical data mainly refers to: i) the inspection records from the EM 120 vehicle to get the standard deviation of longitudinal level defects (SD_{LL}) and the standard deviation of horizontal alignment defects (SD_{HA}) relative to 200-m long track sections, ii) the operation records to get the accumulated tonnage (T_{svkl}), and finally to iii) the maintenance records to get the past maintenance and renewal actions (N_{svkl}, R_{svkl}). As the main contribution of this paper is the HBM itself rather than a comprehensive analysis of the experience from the Lisbon-Oporto line regarding inspection, operation and maintenance, we decided to focus on a particular area in this section, though in section 6 an overview of the all areas of the Lisbon-Oporto line is provided as an additional application of the present HBM. This first tested area mainly involves a double-track $(V_s = 2)$ area of about 15 km $(K_s = 74)$, from a total of 36 inspections $(L_s = 36)$ from February 2001 up to October 2009.

5.2 – MCMC simulation details

MCMC samples were of size 20,000, taking every tenth iteration (thin=10) of the simulated sequence, after 10,000 iterations of burn-in period, for the four models. Initial values were set: for the variance terms σ_s^2, $\sigma_{\alpha_{sv}}^2$, $\sigma_{\beta_{sv}}^2$, $\sigma_{\alpha'_{sv}}^2$, $\sigma_{\beta'_{sv}}^2$ and $\sigma_{\delta_s}^2$ equal to 10 (which is equivalent to precision $(1/\sigma^2)$ equal to 0.1), for the spatially correlated terms $\varepsilon_{\alpha_{svk}}$, $\varepsilon_{\beta_{svk}}$, $\varepsilon_{\alpha'_{svk}}$ and $\varepsilon_{\beta'_{svk}}$ equal to 0, and finally, for each parameter α_{sv}, β_{sv}, α'_{sv}, β'_{sv} and δ_{sv} equal to 0. To have a reasonable confidence on MCMC convergence, we ran an alternative MCMC simulation for every model using different initial values, and found very similar results for posterior estimates. We also used the diagnostic convergence tests from BOA program [37], which did not reveal any evidence against convergence for all MCMC outputs.

5.3 – Discussion of model results

Table 1 provides a comparison of the four models based on the DIC for each quality indicator (SD_{LL} and SD_{HA}). We can spot that model M3 with first-order random walk CAR structure has the lowest DIC value among the four models explored for the SD_{LL} and for the SD_{HA}. However, note that for each quality indicator (SD_{LL} and SD_{HA}), the DIC for model M4, which has second-order random walk CAR structures, is not so distant from the DIC value for model M3.

Based on the MCMC simulations for each quality indicator and for each model, table 2 provides estimates of the posterior parameters for the example track segment, for both dependent variables (SD_{LL} and SD_{HA}).

Regarding Table 2 for the dependent variable SD_{LL}, note that both the parameters related to renewal track sections, i.e. initial quality α_{sv} and deterioration rate β_{sv}, present very close values, especially for models M2, M3 and M4. This is also true for the parameters related to non-renewed track sections (α'_{sv} and β'_{sv}), though for model M1 values differ very much particularly for the deterioration rate. Regarding the disturbance effect (δ_{sv}), the standard deviations (s.d.) associated with its estimates are high compared to its mean values for models M2, M3 and M4. Moreover, note that all values for the disturbance effect seem to be very close to zero. Besides, note that the deterioration rates for non-renewed track sections are (on average) at least four times higher than the deterioration rates for renewed track sections for models M2, M3 and M4, and only 40% higher for model M1; whereas the initial quality for non-renewed track sections seem to be at least four times higher than for the renewed track sections (on average) for models M2, M3 and M4, and five times higher for model M1.

Regarding table 2 for the dependent variable SD_{HA}, note that both the parameters related to renewed track sections, i.e. initial quality α_{sv} and deterioration rate β_{sv}, also present very close values, especially for models M2, M3 and M4. Again, this is also true for the parameters related to non-renewed track sections (α'_{sv} and β'_{sv}), though for model M1 values differ a little bit more, particularly for the initial quality. Regarding the disturbance effect (δ_{sv}), the standard deviations (s.d.) associated with its estimates are relatively high compared to its mean values for models M2, M3 and M4. Moreover, all values for the disturbance effect seem to be close to zero. Besides, note that the deterioration rates for non-renewed track sections are (on average) at least 60% higher than the deterioration rates for renewed track sections for models M1, M2, M3 and M4; whereas the initial quality for non-renewed track sections seems to be 3 times higher than for the renewed track sections (on average) for models M1, M2, M3 and M4.

Moreover, comparing between SD_{LL} and SD_{HA}, one finds that, on average, deterioration rates for renewed and non-renewed rail track sections are lower for the horizontal alignment than for the longitudinal level defects, whereas the initial qualities are higher for renewed rail track sections and lower for non-renewed rail track sections. Regarding the disturbance effect (δ_{sv}), it assumes higher values for the model relative to the standard deviation of horizontal alignment defects (SD_{HA}) than for the model relative to the standard deviation of longitudinal level defects (SD_{LL}).

5.4 – A Sensitivity Analysis on the influence of prior distributions

To analyse the influence of priors' specifications, a sensitivity analysis is conducted in this subsection, assuming different inverse gamma priors $IG(c, d)$. For this purpose, we followed the same experimental design from a similar investigation contained in Silva et al. [39] on Bayesian Hierarchical models to analyse revascularization odds, using only inverse gamma distributions but different combinations: $(c, d) = (0.5, 0.0005)$, $(0.001, 0.001)$, $(0.01, 0.01)$, $(0.1, 0.1)$, $(2, 0.001)$, $(0.2, 0.0004)$ and $(10, 0.25)$, which are denoted by A, B, C, D, E, F and G, respectively. As it is noted in [39], priors C and D (with $c = d$) are variants of prior B with a larger associated dispersion (compared to B) in increasing order, whereas prior E and F correspond to distributions with the same mode $(d/(c + 1))$ as prior A, but with lower (E) and larger (F) dispersion. Prior G is a less vague prior from this experimental design set.

Table 3 provides for model M3 and for both quality indicators SD_{LL} and SD_{HA} some estimates of the variances σ_s^2, $\sigma_{\alpha_{sv}}^2$, $\sigma_{\beta_{sv}}^2$, $\sigma_{\alpha'_{sv}}^2$, $\sigma_{\beta'_{sv}}^2$ and $\sigma_{\delta_s}^2$ for different parameters and the sensitivity of these estimates using different priors. Note that for both the SD_{LL} and the SD_{HA} cases, the mean values for σ_s^2, σ_α^2 and $\sigma_{\alpha'}^2$, are not very sensitive to different inverse gamma priors; whereas σ_β^2 and $\sigma_{\beta'}^2$ seem to be very sensitive for priors D and G and σ_δ^2 seems to vary very much for all priors with associated standard deviation (s.d.) higher than its mean values, which indicates that priors are conveying a lot of information to the variance of the disturbance effect - σ_δ^2.

We also conducted a sensitivity analysis on model selection, i.e. we analysed if different priors would affect the selection of model 3 as the best model according to the Deviance Information Criterion (DIC) of both quality indicators. Table 4 explores the sensitivity of DIC for different models using different priors. For this analysis, we decided to leave out model M1 because its DIC value for Prior A was high compared to the others. The last two columns of the table present the selected model according to the lowest DIC for each prior.

Regarding Table 4 on the SD_{LL}, note that model M3 seems to be the best model for priors A, B, D, F and G; whereas model M4 seems to present a lowest DIC value for priors C and E. Therefore, the selection between models M3 and M4 is sensitive to the prior used for the variance components for the quality indicator SD_{LL}. Apart from prior C, for which DIC values are almost the same, it seems that model M4 has a lowest DIC for more vague priors in the experimental set used, and model M3 has a lowest DIC for less vague (more precise or more informative) priors. Regarding table 4 on the SD_{HA}, model M3 seems to be the best model for priors A, B, E, F and G; whereas model M4 seems to present a lowest DIC value for priors C and D. Similar to what happen for the standard deviation of the longitudinal level defects, the selection between models M3 and M4 is sensitive to the prior used for the variance components for the horizontal alignment case. Again, it seems that model M4 has a lowest DIC for more vague priors in the experimental set used, and model M3 has a lowest DIC for less vague (more precise or more informative) priors.

6 – Applying HBM: an overview of the Portuguese main line degradation

This section explores the application of the HBM model to all the track segments of the Lisbon-Oporto line. In section 5, the HBM were validated for a particular area with a double-track line, and a detailed sensitivity analysis was conducted. In the current section, the HBM for the SD_{LL} and for the SD_{HA} will be explored for all the track segments without the detailed sensitivity analysis previously conducted. Model M3, with a first-order random walk CAR structure for the deterioration rates and the initial quality parameters, was the selected model explored in the following analysis, as it exhibited the lowest values in the Deviance Information Criterion (DIC) for that particular example.

For the application of the HBM model to all track segments, some adaptations were conducted: i) in non-renewed track segments the parameters α_{sv} and β_{sv}, associated with renewed track segments were excluded from the equation to express the mean value (m_{svkl}) in the HBM, as well as the corresponding CAR structures and ii) in renewed rail track segments (i.e. renewed before 2001), the parameters α'_{sv} and β'_{sv} were excluded.

The Lisbon-Oporto railway line has a total length of 336.2 km and links the most populated cities in Portugal by using passenger trains running at a maximum speed of 220 km/h and freight trains running at 80 km/h. The maximum permissible load of trains is equal to 22.5 t and the renewal works have been conducted since 1996. Approximately 2/3 of the line was renewed, whereas the remaining 1/3 will be renewed in the next years. These renewal operations included a thorough improvement of the track-bed bearing capacity and a complete renewal of the track superstructure incorporating monoblock concrete sleepers (with a length of 2.60 m and a seating surface of 3,125 cm^2 per rail seat) spaced at 600 mm, rail UIC 60 and Vossloh fastening system with plastic railpads ZW 687 (vertical stiffness 450 kN/mm). The Lisbon-Oporto line is a double-track line in its majority, and between *Braço de Prata* (4.0 km) and *Alverca* (21.8 km), it is a multiple track segment, more specifically a four-track segment: two tracks (Rapid and Slow Lines – RL and SL) in the Oporto direction (Northwards) and two tracks (RL and SL) in the Lisbon direction (Southwards). This multiple track configuration is particularly useful as it allows faster trains to overtake slower trains. For clarity, only the two Rapid tracks (RL) of the multiple track segments between *Braço de Prata* (4.0 km) and *Alverca* (21.8 km) are represented in the following figures.

In terms of maintenance and renewal actions performed in the analysed period, i.e. between 2001 and 2009, Figure 4 provides an overview of the annual number of track sections which benefited from planned maintenance/tamping actions (N) and the ratio of renewed tracks sections (R) from 2001 to 2009. This ratio quantifies the number of renewed track sections over the total number of track sections; or in other terms, the number of renewed kilometers over the line total length. Note that the ratio of renewed track sections has increased from 0.33 in 2001 (i.e. approximately 1/3 of total length) to 0.64 in 2009 (i.e. approximately 2/3 of total length) that has been renewed since 1996.

In terms of MCMC simulation details, the initial values, burn-in period, and number of samples were set equal to the MCMC simulation runs for the exemplifying area in section 5, i.e. MCMC samples were of size 20,000, taking every tenth iteration (thin=10) of the simulated sequence, after 10,000 iterations of burn-in period, with the initial values set for the variance terms σ_s^2, $\sigma_{\alpha_{sv}}^2$, $\sigma_{\alpha'_{sv}}^2$, $\sigma_{\beta_{sv}}^2$, $\sigma_{\beta'_{sv}}^2$ and $\sigma_{\delta_s}^2$ equal to 10 (which is equivalent to precision ($1/\sigma^2$) equal to 0.1), for the spatially correlated terms $\varepsilon_{\alpha_{svk}}$, $\varepsilon_{\beta_{svk}}$, $\varepsilon_{\alpha'_{svk}}$ and $\varepsilon_{\beta'_{svk}}$ equal to 0, and finally, for each parameter α_{sv}, β_{sv}, α'_{sv}, β'_{sv} and δ_{sv} equal to 0.

Figures 5, 6 and 7 provide the posterior mean for all the degradation parameters for the HBM for the SD_{LL}: for the initial standard deviations for renewed track sections α and for non-renewed track sections α' for both directions (in Figure 5), for the deterioration rates for renewed track sections β and for non-renewed track sections β' for both directions (in Figure 6), and for the disturbance effect due to tamping δ for both directions (in Figure 7).

Both directions are presented in the following figures: the Oporto direction (towards Oporto) is presented in solid lines for renewed (in blue) and non-renewed track sections (in red) and the Lisbon direction (towards Lisbon) is presented in dashed lines for renewed (in blue) and non-renewed track sections (in red). Moreover, some track segments do not exhibit values for the renewed or non-renewed track segments. This happens because some track segments, for the analysed period between 2001 and 2009, were already renewed or still need renewal. Track sections which present both parameters for renewed and non-renewed track sections are the segment that benefitted from a renewal action in that analysed period.

Figure 5 provides the posterior mean for the initial quality parameters of the hierarchical Bayesian model for the standard deviation of longitudinal level defects, i.e. for the renewed track segments (α_{sv}) and for the non-renewed track segments (α'_{sv}). Non-renewed track segments represented in red tend to exhibit larger values than renewed track segments represented in blue. For the first track segment between *Lisbon – Sta. Apolónia* (0.0 km) and *Bifurcação Xabregas* (1.6 km), non-renewed track sections exhibit extremely high values compared to the other track segments, whereas renewed track segments also exhibit the higher values compared to the other track segments. This might indicate that the data associated with this particular track segment is not reliable. Figure 6 exhibits the posterior means for the degradation parameters β_{sv} and β'_{sv}, respectively the deterioration rates for renewed and non-renewed track segments. Again the two directions of the Lisbon-Oporto line are incorporated in the same figure: the Oporto direction in solid lines, and the Lisbon direction in dashed lines. Non-renewed track segments (marked in red) exhibit higher deterioration rates than renewed track segments (marked in blue) for both directions. Note that some segments do not exhibit values for the renewed or non-renewed track segments. For instance, the track segments comprehended between locations *Pombal* (169.6 km) and *Pampilhosa* (231.2 km), only the parameter β'_{sv} for non-renewed track sections is represented for the Oporto and the Lisbon directions, because no renewed track sections are within those locations for the analyzed time period and thus, no value for β_{sv} is represented.

Figure 7 provides values for the disturbance effect (δ) of the initial quality after each tamping operation for the SD_{LL}. It exhibits a maximum value of 0.021 for the track segment between locations 287.4 km and 290.2 km for the Lisbon direction, and a minimum value of -0.003 for the track segment between locations 106.4 km and 114.4 km.

Similar to the previous Figures 5-7, the following Figures 8, 9 and 10 provide the posterior mean for all the degradation parameters, but for the HBM for the SD_{HA}: for the initial standard deviations for renewed track sections α and for non-renewed track sections α' for both directions (in figure 8), for the deterioration rates for renewed track sections β and for non-renewed track sections β' for both directions (in figure 9), and for the disturbance effect due to tamping δ for both directions (in figure 10).

Similarly to Figure 5, Figure 8 exhibits the posterior means of the initial qualities, but for the SD_{HA}, i.e. for the renewed track segments (α_{sv}) and for the non-renewed track segments (α'_{sv}). Again, the parameter referring to non-renewed track segments exhibits a larger value than the same parameter referring to non-renewed track segments. Again the track segment near *Lisbon – Sta. Apolónia* (0.0 km) exhibit a considerable high value for α'_{sv} in the Lisbon direction, which might indicate some reliability problems in the model for that particular segment.

Figure 9 exhibits the posterior means for the deterioration rates for renewed (β_{sv}) and non-renewed (β'_{sv}) track segments in Lisbon-Oporto line for the SD_{HA}. Deterioration rates for non-renewed track segments tend to exhibit higher posterior means than for renewed track segments, though this contrast is not as visible as in figure 6 for the longitudinal level defects case. Nevertheless, the posterior means for the deterioration rates tend to exhibit higher variability among track segments for the horizontal alignment indicator (figure 9) than for the longitudinal level case (figure 6), specially for the parameters associated with non-renewed track segments.

Finally, figure 10 exhibits the posterior means for the disturbance effect (δ) of the initial quality after each tamping operation, but in this case for the SD_{HA}. For the horizontal alignment case, in contrast with the longitudinal level case, it exhibits only positive values with a considerable larger range. It exhibits a maximum value of 0.178 for the track segment between *Gaia* (332.4 km) and *Oporto-Campanhã* (336.2 km) for the Lisbon direction, and a minimum value of 0.001 for the track segment between locations 9.6 km and 13.8 km.

Note that in order to preserve some clarity in the exposition of the previous figures 5-10, the corresponding standard deviations for each degradation parameter (α, α', β, β' and δ) were excluded from these figures, neither a credible interval was included for that purpose. It is also important to mention that similarly to what happened in the exemplifying area explored in section 4.3, the standard deviations for these estimates are relatively low compared to the posterior mean of those parameters, indicating the statistically significance of those degradation parameters. Nevertheless, for the disturbance effect δ, the associated standard deviations exhibit for some segments high values compared to the posterior mean for that parameter, especially for the HBM for the SD_{LL}. Another important issue regarding the exploration of the degradation parameters is related with the reliability of the track geometry indicators for the initial segments, i.e. the stations areas near *Lisbon – Sta. Apolónia* (0.0 km) and near *Oporto-Campanhã* (336.2 km). The standard deviations associated with the posterior estimates for the degradation parameters are much higher when compared with other areas.

7 – Conclusions and further research

This paper has discussed statistical modelling of railway track geometry degradation using Hierarchical Bayesian models, in order to predict the evolution of the main quality indicators for planned maintenance, namely the standard deviation of longitudinal level defects (SD_{LL}) and the standard deviation of horizontal alignment defects (SD_{HA}). Particular attention was given to the need to insert spatial statistical dependencies between model parameters for consecutive track sections, namely make them follow random walk priors, capturing the spatial correlation between deterioration rates and initial qualities, which proved to be a better model based on the Deviance Information Criterion (DIC) comparison. Moreover, we have also provided an inference method through the derivation of the full conditional posterior distribution and the associated Markov Chain Monte Carlo (MCMC) or Gibbs sampling procedure is provided in the Appendix.

In general terms, the application of the model to a sample of railway inspection, operation and maintenance data, showed that the HBM exhibit a worse fit of the quality indicator SD_{HA} compared to the quality indicator SD_{LL}, suggesting that the horizontal alignment defects seems to be less predictable.

For further research, the present HBM can be extended by assigning a typical transportation demand model to the future tonnage usage T, relaxing the assumption that T is known quantity. Additionally, the HBM can serve as a simulation predictive tool and to compare different maintenance and renewal strategies, and ideally find a strategy that minimizes life-cycle costs and safety impacts, while improving ride comfort and track access availability. We have submitted a paper on that topic elsewhere [40].

Moreover, an important unexplored branch regards the use of more complicated correlation structures than the first- and second-order random walks. In fact, other spatial formulations, using for example power exponential functions to model the decline of correlation depending on the distance between track sections may bring more inside in the spatial correlation between degradation models. These extensions should also be comprehended within a multivariate hierarchical Bayesian model to predict SD_{LL} and SD_{HA} in a joint model instead of the separate HBM for each indicator. Finally, another future direction worth pursuing would be trying to let the model comparison be conducted for each area s, and then select as an overall predictive model the combination of all models selected at a particular s, rather than running model selection for all the areas. In that sense, we would be letting the overall prediction model select which model (i.e. M3 or M4 for example), it would use for each area s.

Acknowledgments

We would like to thank the support and collaboration of the Portuguese Railway Infrastructure Manager, REFER, E.P.E., as well as the support of the Portuguese Foundation for Science and Technology, FCT (Project Reference PTDC/SEN-TRA/112975/2009) and MIT Portugal (FCT PhD grant SFRH/BD/33785/2009). We are also grateful to Professor Carlos Paulino for letting the first author attend his classes on Bayesian Statistics and to Professor Giovani Silva for his suggestions and comments on a previous version of the manuscript. Needless to say, that any errors or lapses are totally our responsibility.

References

[1] Corbin JC, Fazio AE. Performance-Based Track-Quality Measures and Their Application to Maintenance-of-Way Planning. Transportation Research Record: Journal of the Transport Research Board 1981; 802: 19-26.

[2] Hamid A, Gross A. Track-Quality Indices and Track-Degradation Models for Maintenance-of-Way Planning. Transportation Research Record: Journal of the Transport Research Board 1981; 802: 2-8.

[3] Bing AJ, Gross A. Development of Railroad Track Degradation Models. Transportation Research Record: Journal of the Transport Research Board 1983; 939: 27-31.

[4] Hamid A, Yang T-L. Analytical Descriptions of Track-Geometry Variations. Transportation Research Record: Journal of the Transport Research Board 1981; 838: 19-26.

[5] Esveld C. Modern Railway Track. 2nd ed., 2001. ISBN 90-8004-324-3-3.

[6] Iyengar RN, Jaiswal OR. A new model for non-Gaussian random excitations. Probabilistic Engineering Mechanics 1993; 8: 281-287.

[7] Iyengar RN, Jaiswal OR. Random field modeling of railway track irregularities. Journal of Transportation Engineering 1995; 4: 303-308.

[8] Kumar KS, Stathopoulos T. Discussion of 'Random Field Modeling of Railway Track Irregularities' 1997; 123: 245.

[9] Riessberger K. Extending Maintenance Cycles and Ballast Life. In: 7th International Heavy Haul Conference, Brisbane, Australia, 2001.

[10] Ubalde L, López-Pita A, Teixeira P, Bachiller A, Gallego I. Track deterioration in high-speed railways: influence of stochastic parameters. In: 8th International Conference Railway Engineering, London, UK, 2005, pp.54-62. ISBN 0-947644-56-3.

[11] Veit P. Track Quality – Luxury or Necessity? Maintenance & Renewal. Journal of Railway Technical Review 2007; pp. 8-12. ISBN 978-3-7771-0367-9.

[12] López-Pita A, Teixeira PF, Casas C, Ubalde L, Robusté F. Evolution of track geometric quality in high-speed lines: ten years experience of the Madrid-Seville line. Proceedings of the Institution of Mechanical Engineers Part F: Journal of Rail and Rapid Transit 2007; 221: 147-155.

[13] Lichtberger B. Track Compendium: Formation, Permanent Way, Maintenance, Economics. 1st ed., 2005. ISBN 3-7771-0320-9.

[14] Quiroga LM, Schnieder E. Monte Carlo simulation of railway track geometry deterioration and restoration. In: Reliability. Risk and Safety, Taylor & Francis Group, London, pp. 1077-1080. ISBN 978-0-415-60427-7.

[15] Andrade AR, Teixeira PF. Uncertainty in Rail Track Geometry Degradation: Lisbon-Oporto Line case study. Journal of Transportation Engineering 2011; 137: 193-200.

[16] Vale C, Lurdes SM. Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line. Reliability Engineering and System Safety 2013; 116: 91-98.

[17] Guler H. Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways. Structure and Infrastructure Engineering 2014; 10(5): 614-626.

[18] Prescott D, Andrews J. Stochastic State Space Methods for Railway Network Asset Management Modelling. Advances in Risk and Reliability Technology Symposium 2013.

[19] Andrews J, Prescott D, Rozieres F. A Stochastic Model for Railway Track Asset Management. Reliability Engineering and System Safety. DOI: 10.1016/j.ress.2014.04.021

[20] Mokhtarian P, Namzi-Rad MR, Ho TK, Suesse T. Bayesian nonparametric reliability analysis for a railway system at a component level. IEEE International Conference on Intelligent Rail Transportation (ICIRT) 2013; 197-202.

[21] Gong C, Iwnicki S, Bezin Y. The effect of railway vehicle dynamics on the lateral alignment of track. Proceedings of the Institution of Mechanical Engineers Part F: Journal of Rail and Rapid Transit. DOI: 10.1177/0954409714536548

[22] UIC. Best practice guide for optimum track geometry durability, 2008. ISBN 2-7461-1456-9.

[23] Andrade AR, Teixeira PF. Unplanned maintenance needs related to rail track geometry. ICE Transport - Proceedings of the Institution of Civil Engineers – Transport. DOI: 10.1680/tran.11.00060

[24] Andrade AR, Teixeira PF. A negative binomial regression model to predict temporary speed restrictions in railway infrastructure. Journal of Risk and Reliability (under review).

[25] Zhao J, Chan AHC, Roberts C, Stirling AB. Optimizing Policies of Railway Ballast Tamping and Renewal. Transportation Research Record: Journal of the Transport Research Board 2006; 1943: 50-56.

[26] El-Sibaie M, Zhang Y. Objective Track Quality Indices. Transportation Research Record: Journal of the Transport Research Board 2004; 1863: 81-87.

[27] Esveld C. Vehicle Reactions versus Track Geometry. Railroad Conference, ASME/IEEE Joint 1990; 171-174.

[28] EN 13848-5:2008. Railway applications – Track – Track geometry quality – Part 5: Geometric quality levels.

[29] Bernardo JM. Bayesian Statistics. Updated and abridged version of the Chapter 'Bayesian Statistics' published in Probability and Statistics (R. Viertl, ed) of the Encyclopedia of Life Support Systems (EOLSS), 2003, Oxford, UK: UNESCO.

[30] Paulino CD, Turkman MAA, Murteira B. Estatística Bayesiana. Fundação Calouste Gulbenkian, 2003. ISBN 972-31-1043-1.

[31] Ntzoufras I. Bayesian Modeling Using WinBUGS. Wiley Series in Computational Statistics, 2009. ISBN 978-0-470-14114-4.

[32] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – A Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 2000; 10: 325-337.

[33] Gilks WR, Wild P. Adaptive Rejection Sampling for Gibbs Sampling. Journal of the Royal Statistical Society, Series C (Applied Statistics) 1992; 41: 337-348.

[34] Besag J. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Society, Series B (Methodological) 1974; 36: 192-236.

[35] Gelman A. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). Bayesian Analysis 2006; 3: 515-534.

[36] Besag J, Kooperger CL. On conditional and intrinsic autoregressions. Biometrika 1995; 82: 733-746.

[37] Smith B. Bayesian Output Analysis Program – Version 1.1 User's Manual. Technical Report 2005, The University of Iowa.

[38] Spiegelhalter DJ, Best NG, Carlin BP, Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B (Methodological); 64: 583-639.

[39] Silva GL, Dean CB, Niyonsenga T, Vanasse A. Hierarchical Bayesian spatiotemporal analysis of revascularization odds using smoothing splines. Statistics in Medicine 2008; 27: 2381-2401.

[40] Andrade AR, Teixeira PF. Exploring different Alert Limit strategies in the maintenance of railway track geometry. Journal of Transportation Engineering (under review).

Appendix

Let $\boldsymbol{\theta}$ be the vector of the model parameters, with elements σ_s^2, δ_{sv}, σ_δ^2, $\varepsilon_{\alpha_{svk}}$, σ_α^2, $\varepsilon_{\beta_{svk}}$, σ_β^2, $\varepsilon_{\alpha'_{svk}}$, $\sigma_{\alpha'}^2$, $\varepsilon_{\beta'_{svk}}$, $\sigma_{\beta'}^2$, α_{sv}, β_{sv}, α'_{sv} and β'_{sv}, with s, v and k varying for: $s = 1, \dots, S$; $v = 1, \dots, V_s$; $k = 1, \dots, K_s$. From the joint posterior, one can derive the full conditional posterior distributions (denoted below by $\left[j|\boldsymbol{\theta}_{-j}\right]$), which are given by:

i) $\sigma_s^2|\boldsymbol{\theta}_{-\sigma_s^2} \sim \mathrm{IG}\left(c_0 + \frac{1}{2}V_s K_s L_s \,, d_0 + \frac{1}{2}\sum_{v,k,l}(y_{svkl} - m_{svkl})^2\right)$, $s = 1, \dots, S$;

ii) $\delta_{sv}|\boldsymbol{\theta}_{-\delta_{sv}} \propto \exp\left(-\frac{1}{2\sigma_\delta^2}\delta_{sv}^2 - \frac{1}{2\sigma_s^2}\sum_{k,l}(y_{svkl} - m_{svkl})^2\right)$, $s = 1, \dots, S$, $v = 1, \dots, V_s$;

iii) $\sigma_\delta^2\left|\boldsymbol{\theta}_{-\sigma_\delta^2} \sim \mathrm{IG}\left(c_1 + \frac{1}{2}V_s \,, d_1 + \frac{1}{2}\sum_v \delta_{sv}^2\right)\right.$, $s = 1, \dots, S$;

iv) $\varepsilon_{\alpha_{svk}}\left|\boldsymbol{\theta}_{-\varepsilon_{\alpha_{svk}}} \propto \exp\left(-\frac{n_{svk}}{2\sigma_\alpha^2}\left(\varepsilon_{\alpha_{svk}} - \bar{\varepsilon}_{\alpha_{svk}}\right)^2 - \frac{1}{2\sigma_s^2}\sum_l (y_{svkl} - m_{svkl})^2\right)\right.$, $s = 1, \dots, S$, $v = 1, \dots, V_s$, $k = 1, \dots, K_s$;

v) $\sigma_\alpha^2|\boldsymbol{\theta}_{-\sigma_\alpha^2} \sim \mathrm{IG}\left(c_2 + \frac{1}{2}K_s \,, d_2 + \frac{1}{2}\sum_k n_{svk}\left(\varepsilon_{\alpha_{svk}} - \bar{\varepsilon}_{\alpha_{svk}}\right)^2\right)$, $s = 1, \dots, S$, $v = 1, \dots, V_s$;

vi) $\varepsilon_{\beta_{svk}}\left|\boldsymbol{\theta}_{-\varepsilon_{\beta_{svk}}} \propto \exp\left(-\frac{n_{svk}}{2\sigma_\beta^2}\left(\varepsilon_{\beta_{svk}} - \bar{\varepsilon}_{\beta_{svk}}\right)^2 - \frac{1}{2\sigma_s^2}\sum_l (y_{svkl} - m_{svkl})^2\right)\right.$, $s = 1, \dots, S$, $v = 1, \dots, V_s$, $k = 1, \dots, K_s$;

vii) $\sigma_\beta^2\left|\boldsymbol{\theta}_{-\sigma_\beta^2} \sim \mathrm{IG}\left(c_2 + \frac{1}{2}K_s \,, d_2 + \frac{1}{2}\sum_k n_{svk}\left(\varepsilon_{\beta_{svk}} - \bar{\varepsilon}_{\beta_{svk}}\right)^2\right)\right.$, $s = 1, \dots, S$, $v = 1, \dots, V_s$;

viii) $\varepsilon_{\alpha'_{svk}}\left|\boldsymbol{\theta}_{-\varepsilon_{\alpha'_{svk}}} \propto \exp\left(-\frac{n_{svk}}{2\sigma_{\alpha'}^2}\left(\varepsilon_{\alpha'_{svk}} - \bar{\varepsilon}_{\alpha'_{svk}}\right)^2 - \frac{1}{2\sigma_s^2}\sum_l (y_{svkl} - m_{svkl})^2\right)\right.$, $s = 1, \dots, S$, $v = 1, \dots, V_s$, $k = 1, \dots, K_s$;

ix) $\sigma^2_{\alpha'}|\boldsymbol{\theta}_{-\sigma^2_{\alpha'}} \sim \mathrm{IG}\left(c_2 + \frac{1}{2}K_s\,, d_2 + \frac{1}{2}\sum_k n_{svk}\left(\varepsilon_{\alpha'_{svk}} - \bar{\varepsilon}_{\alpha'_{svk}}\right)^2\right), s = 1, \dots, S, v = 1, \dots, V_s;$

x) $\varepsilon_{\beta'_{svk}}\big|\boldsymbol{\theta}_{-\varepsilon_{\beta'_{svk}}} \propto \exp\left(-\frac{n_{svk}}{2\sigma^2_{\beta'}}\left(\varepsilon_{\beta'_{svk}} - \bar{\varepsilon}_{\beta'_{svk}}\right)^2 - \frac{1}{2\sigma^2_S}\sum_l (y_{svkl} - m_{svkl})^2\right), s = 1, \dots, S, \quad v = 1, \dots, V_s,$

$k = 1, \dots, K_s;$

xi) $\sigma^2_{\beta'}|\boldsymbol{\theta}_{-\sigma^2_{\beta'}} \sim \mathrm{IG}\left(c_2 + \frac{1}{2}K_s\,, d_2 + \frac{1}{2}\sum_k n_{svk}\left(\varepsilon_{\beta'_{svk}} - \bar{\varepsilon}_{\beta'_{svk}}\right)^2\right), s = 1, \dots, S, v = 1, \dots, V_s;$

xii) $\alpha_{sv}|\boldsymbol{\theta}_{-\alpha_{sv}} \propto \exp\left(\sum_{k,l}(y_{svkl} - m_{svkl})^2\right) \cdot P[\alpha_{sv}]\,, s = 1, \dots, S, v = 1, \dots, V_s;$

xiii) $\beta_{sv}|\boldsymbol{\theta}_{-\beta_{sv}} \propto \exp\left(\sum_{k,l}(y_{svkl} - m_{svkl})^2\right) \cdot P[\beta_{sv}]\,, s = 1, \dots, S, v = 1, \dots, V_s;$

xiv) $\alpha'_{sv}|\boldsymbol{\theta}_{-\alpha'_{sv}} \propto \exp\left(\sum_{k,l}(y_{svkl} - m_{svkl})^2\right) \cdot P[\alpha'_{sv}]\,, s = 1, \dots, S, v = 1, \dots, V_s;$

xv) $\beta'_{sv}|\boldsymbol{\theta}_{-\beta'_{sv}} \propto \exp\left(\sum_{k,l}(y_{svkl} - m_{svkl})^2\right) \cdot P[\beta'_{sv}]\,, s = 1, \dots, S, v = 1, \dots, V_s;$

Tables

Table 1 – Comparison of the DIC for different models M1-M4 for both quality indicators: the standard deviation of longitudinal level defects (SD_{LL}) and the standard deviation of horizontal alignment defects (SD_{HA}).

Models defined from $*_{svk} = \{\boldsymbol{\alpha}_{svk}, \beta_{svk}, \boldsymbol{\alpha}'_{svk}, \beta'_{svk}\}$		**DIC**	
		SD_{LL}	**SD_{HA}**
M1:	$*_{svk} = *_{sv}$	7192	8324
M2:	$*_{svk} = *_{sv} + \varepsilon_{*_{svk}}$; $\varepsilon_{*_{svk}} \sim N(0, \sigma_*^2)$	-2794	3857
M3:	$*_{svk} = *_{sv} + \varepsilon_{*_{svk}}$; $\varepsilon_{*_{svk}} \sim RW(1)$	-3274	3406
M4:	$*_{svk} = *_{sv} + \varepsilon_{*_{svk}}$; $\varepsilon_{*_{svk}} \sim RW(2)$	-3234	3481

Note: The symbol $*$ intends to represent the parameters $\boldsymbol{\alpha}$, β, $\boldsymbol{\alpha}'$, β'.

Table 2 – Estimates of the posterior parameters for an example track segment (s=1, v=1) for both quality indicators: SD_{LL} and SD_{HA}.

Quality indicator	Models	α_{sv} (mm)		β_{sv} (mm/100MGT)		α'_{sv} (mm)		β'_{sv} (mm/100MGT)		δ_{sv}	
		Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
SD_{LL}	M1	0.3407	0.020	1.203	0.143	1.787	0.038	1.680	0.329	-0.1016	0.019
	M2	0.3065	0.024	1.469	0.133	1.379	0.067	6.406	0.488	-0.0035	0.008
	M3	0.3102	0.012	1.460	0.081	1.381	0.015	6.247	0.179	-0.0015	0.009
	M4	0.3034	0.013	1.499	0.090	1.375	0.016	6.169	0.170	0.0055	0.009
SD_{HA}	M1	0.4353	0.029	0.542	0.191	1.524	0.041	1.123	0.375	0.0172	0.024
	M2	0.4243	0.020	0.602	0.133	1.345	0.074	1.032	0.443	0.0262	0.016
	M3	0.4263	0.019	0.606	0.132	1.342	0.255	1.069	0.273	0.0279	0.015
	M4	0.4291	0.019	0.595	0.128	1.343	0.025	0.959	0.272	0.0304	0.016

Table 3 – Estimates of the spatial variance components based on model M3 with different inverse gamma prior for both quality indicators (SD_{LL} and SD_{HA}) for the particular test segment (s=1, v=1).

QI	Prior	σ_s^2 (mm^2)		$\sigma_{\alpha_{sv}}^2$ (mm^2)		$\sigma_{\alpha'_{sv}}^2$ (mm^2)		$\sigma_{\beta_{sv}}^2$ ($(mm/100\ MGT)^2$)		$\sigma_{\beta'_{sv}}^2$ ($(mm/100\ MGT)^2$)		$\sigma_{\delta_s}^2$	
		Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
SD_{LL}	A	0.0323	0.00141	0.0531	0.00816	0.3478	0.04848	1.4520	0.2379	21.950	3.609	0.0304	0.105
	B	0.0318	0.00069	0.0522	0.00818	0.3679	0.04834	1.6220	0.2489	23.100	3.758	0.1732	0.787
	C	0.0327	0.00199	0.0454	0.00729	0.3551	0.04917	3.3140	0.4299	24.610	4.025	0.2596	4.513
	D	0.0312	0.00067	0.0429	0.00723	0.3618	0.04591	16.380	1.9770	42.070	5.544	0.7139	6.716
	E	0.0326	0.00241	0.0506	0.00833	0.3522	0.04694	1.5470	0.2432	21.610	3.565	0.0121	0.013
	F	0.0318	0.00073	0.0535	0.00856	0.3535	0.04529	1.4480	0.2358	22.380	3.592	0.0947	0.725
	G	0.0316	0.00067	0.0360	0.00569	0.3048	0.03832	32.700	3.6110	57.510	6.801	0.0274	0.010
SD_{HA}	A	0.1114	0.00238	0.0052	0.00259	0.3140	0.03985	0.4275	0.0923	7.245	1.757	0.0023	0.005
	B	0.1119	0.00230	0.0040	0.00209	0.2884	0.04409	0.6264	0.1122	8.026	1.842	0.1276	1.406
	C	0.1108	0.00227	0.0038	0.00133	0.3180	0.04218	2.3540	0.3045	11.50	2.069	0.1262	1.236
	D	0.1104	0.00227	0.0097	0.00232	0.3155	0.04316	15.220	1.8170	31.50	4.277	0.5429	3.340
	E	0.1117	0.00233	0.0028	0.00156	0.3037	0.04087	0.6045	0.1056	6.914	1.609	0.0010	0.001
	F	0.1115	0.00238	0.0063	0.00274	0.3065	0.04239	0.3762	0.0845	7.481	1.762	0.0108	0.057
	G	0.1104	0.00236	0.0114	0.00206	0.2650	0.04256	31.790	3.5250	48.940	5.856	0.0269	0.009

Table 4 – Sensitivity Analysis of the model selection based on DIC for both quality indicators SD_{LL} and SD_{HA} and for different inverse gamma priors.

Prior	Models	DIC		Selected model	
		SD_{LL}	SD_{HA}	SD_{LL}	SD_{HA}
A	M1	7192	8324	M3	M3
	M2	-2794	3857		
	M3	-3274	3406		
	M4	-3234	3481		
B	M2	-2796	3857	M3	M3
	M3	-3337	3382		
	M4	-3230	3520		
C	M2	-2790	3851	M4	M4
	M3	-3106	3429		
	M4	-3107	3398		
D	M2	-2775	3856	M3	M4
	M3	-3275	3449		
	M4	-3154	3431		
E	M2	-2794	3859	M4	M3
	M3	-3083	3467		
	M4	-3217	3552		
F	M2	-2795	3865	M3	M3
	M3	-3232	3400		
	M4	-3189	3556		
G	M2	-2779	3853	M3	M3
	M3	-3288	3209		
	M4	-3214	3266		

Figures

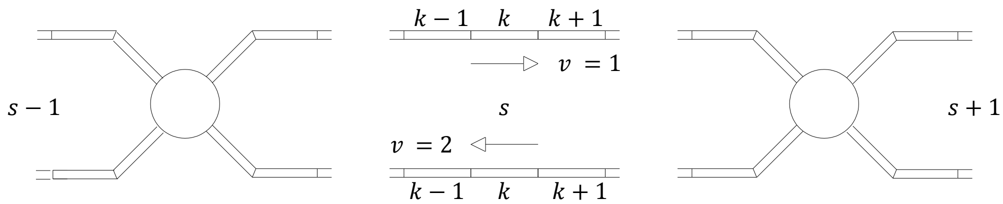

Figure 1 – A typical double track line with indices from area s, segment v and track section k.

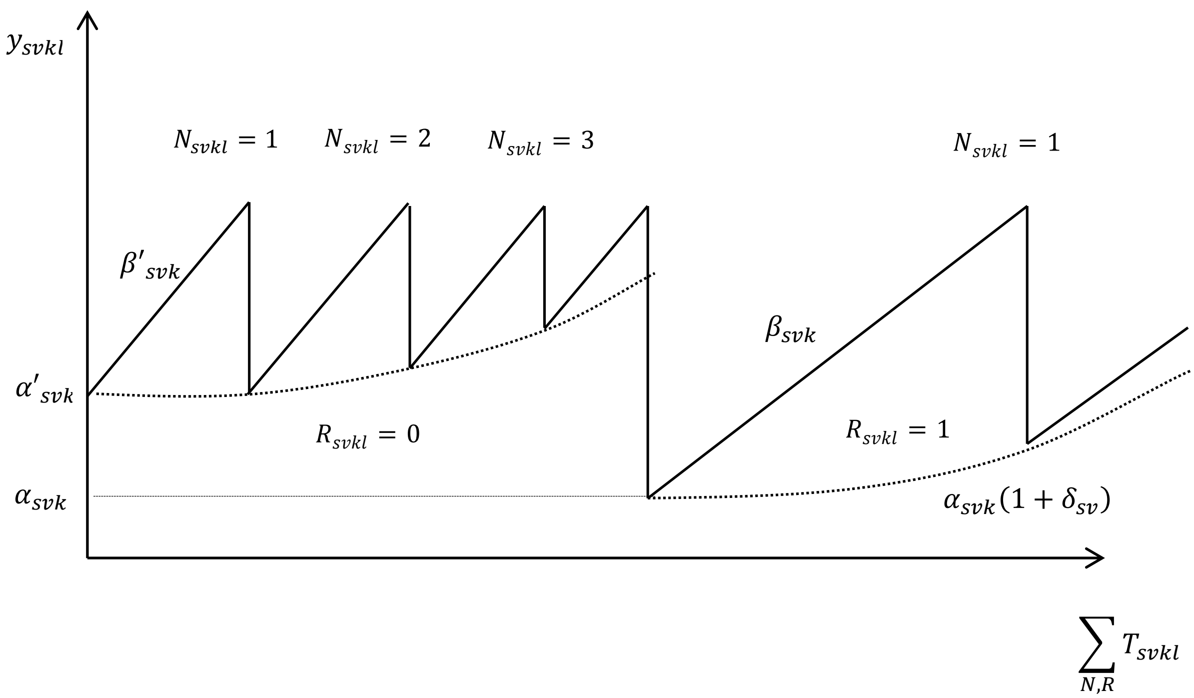

Figure 2 – Mean behaviour of the quality indicator for a given track section k in segment v in area s.

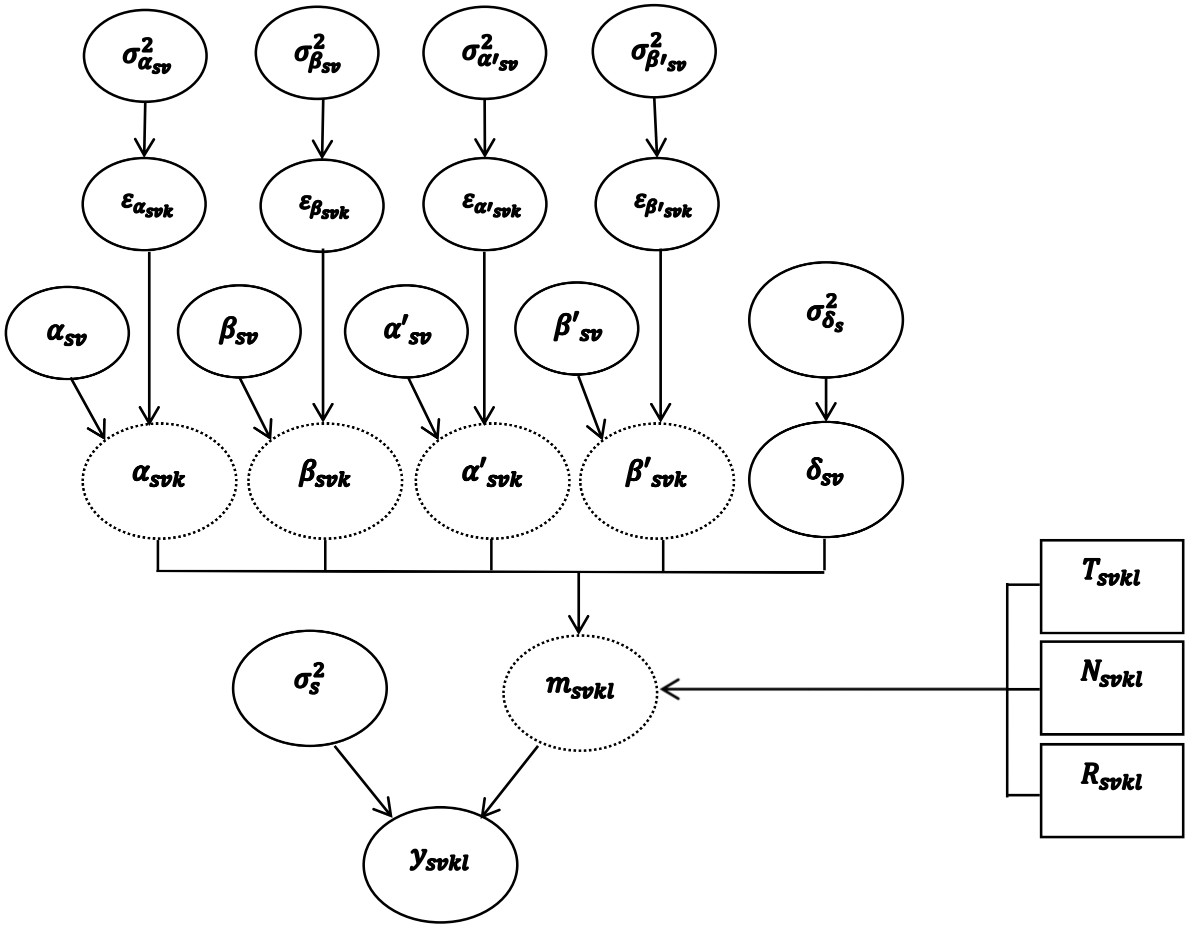

Figure 3 – Directed Acyclic Graph (DAG) for the proposed hierarchical Bayesian models M2, M3 and M4.

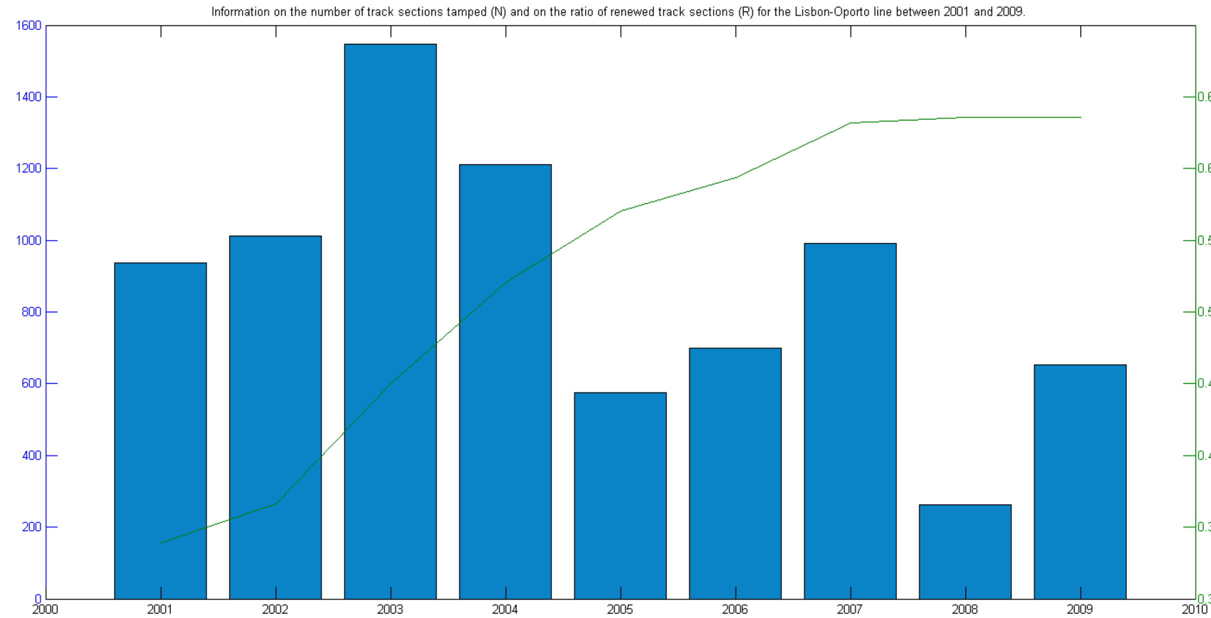

Figure 4 – Information on the number of track sections tamped (N) and on the ratio of renewed track sections (R) for the Lisbon-Oporto line between 2001 and 2009.

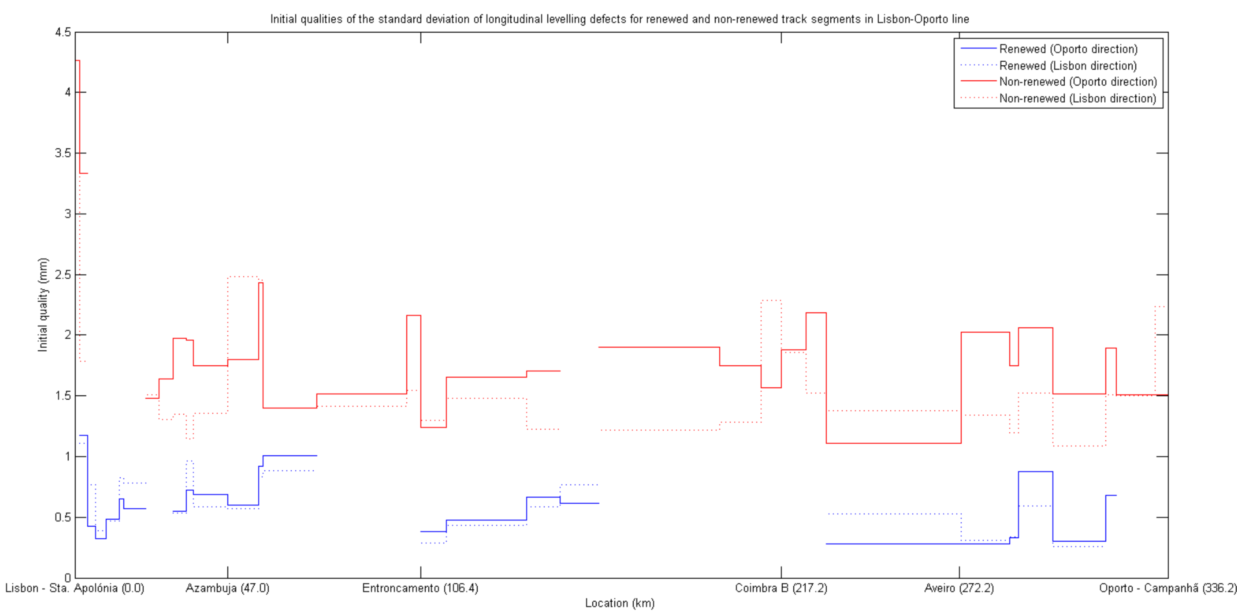

Figure 5 – Initial qualities of the standard deviation of longitudinal level defects for renewed (α) and non-renewed (α') track segments in Lisbon-Oporto line.

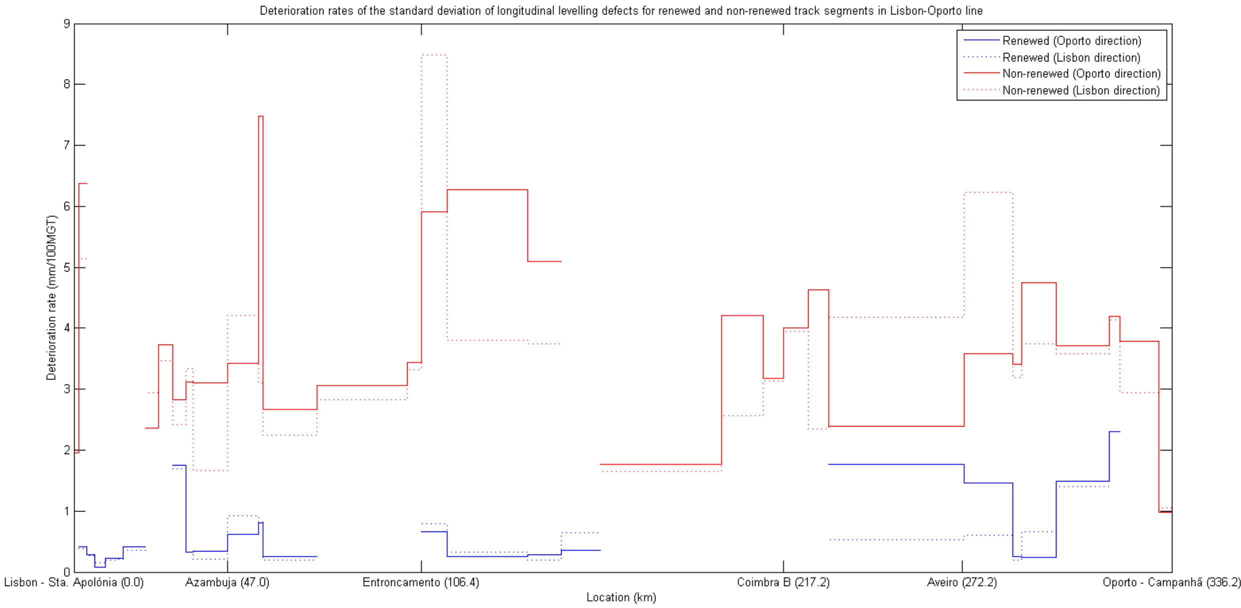

Figure 6 – Deterioration rates of the standard deviation of longitudinal level defects for renewed (β) and non-renewed (β') track segments in Lisbon-Oporto line.

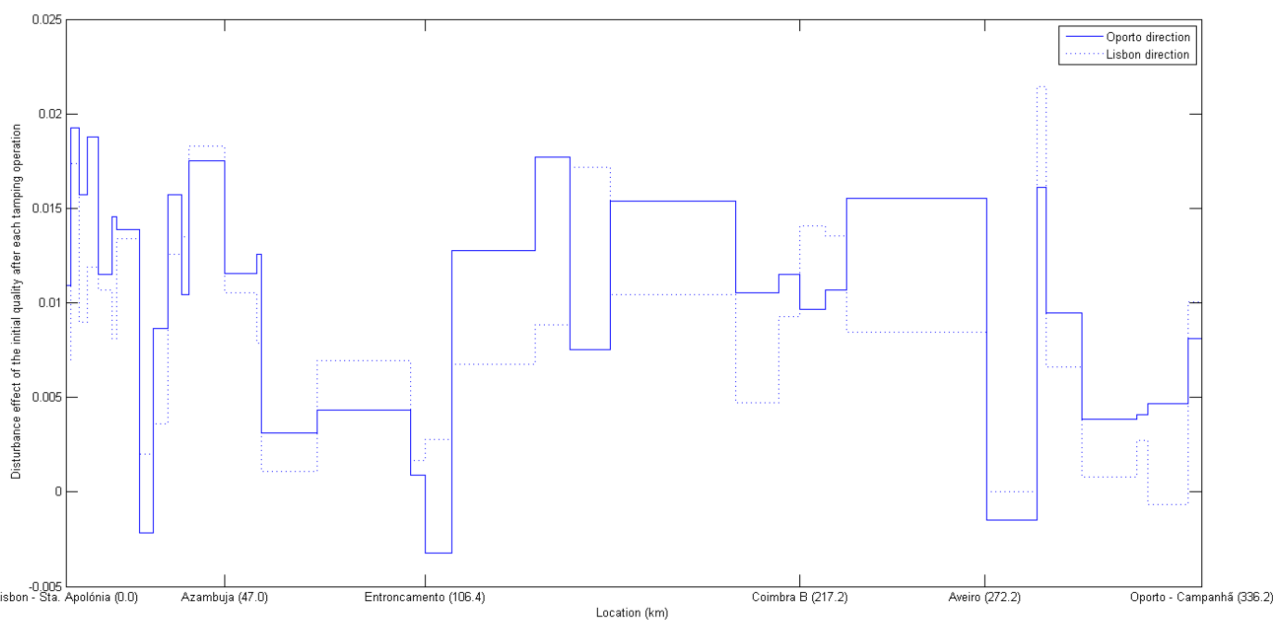

Figure 7 – Disturbance effect (δ) of the initial quality after each tamping operation for the standard deviation of longitudinal level defects.

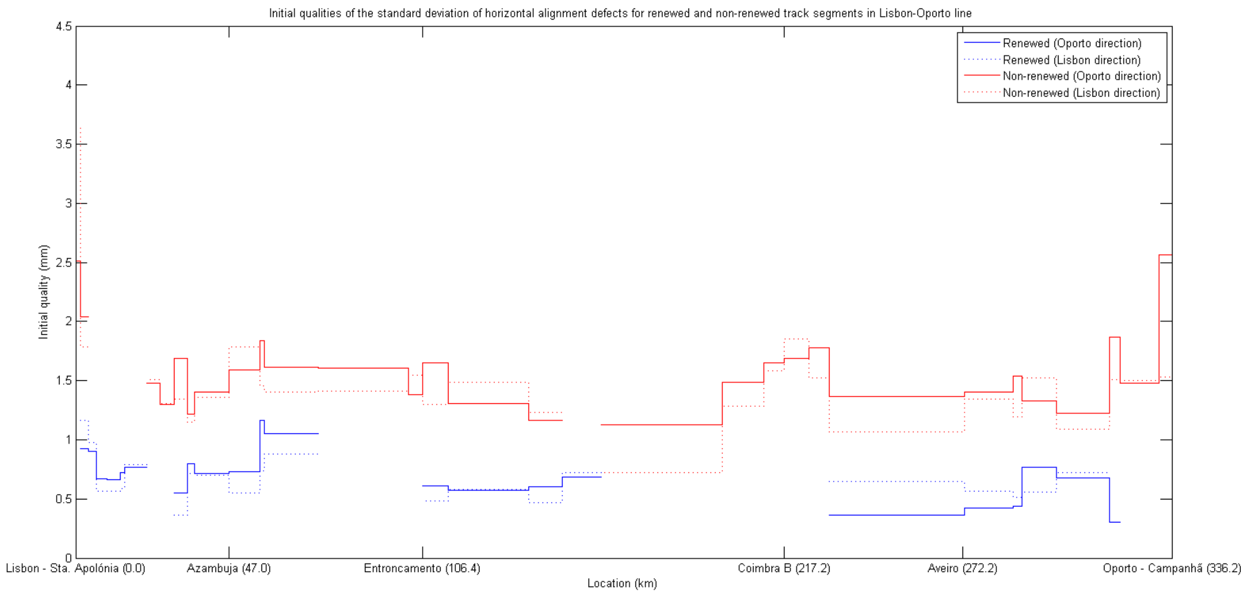

Figure 8 – Initial qualities of the standard deviation of horizontal alignment defects for renewed (α) and non-renewed (α') track segments in Lisbon-Oporto line.

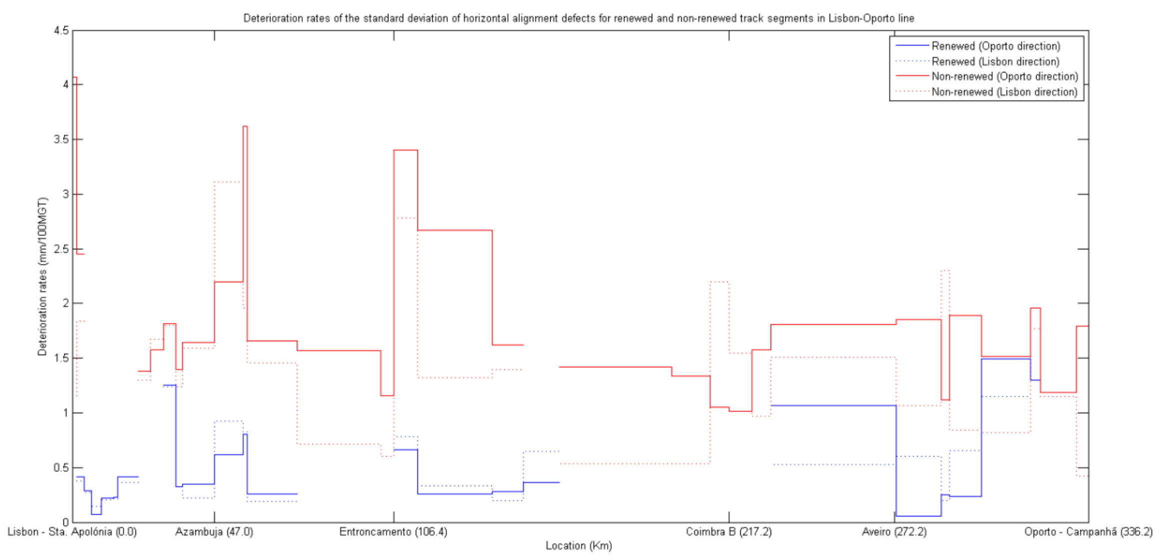

Figure 9 – Deterioration rates of the standard deviation of horizontal alignment defects for renewed (β) and non-renewed (β') track segments in Lisbon-Oporto line.

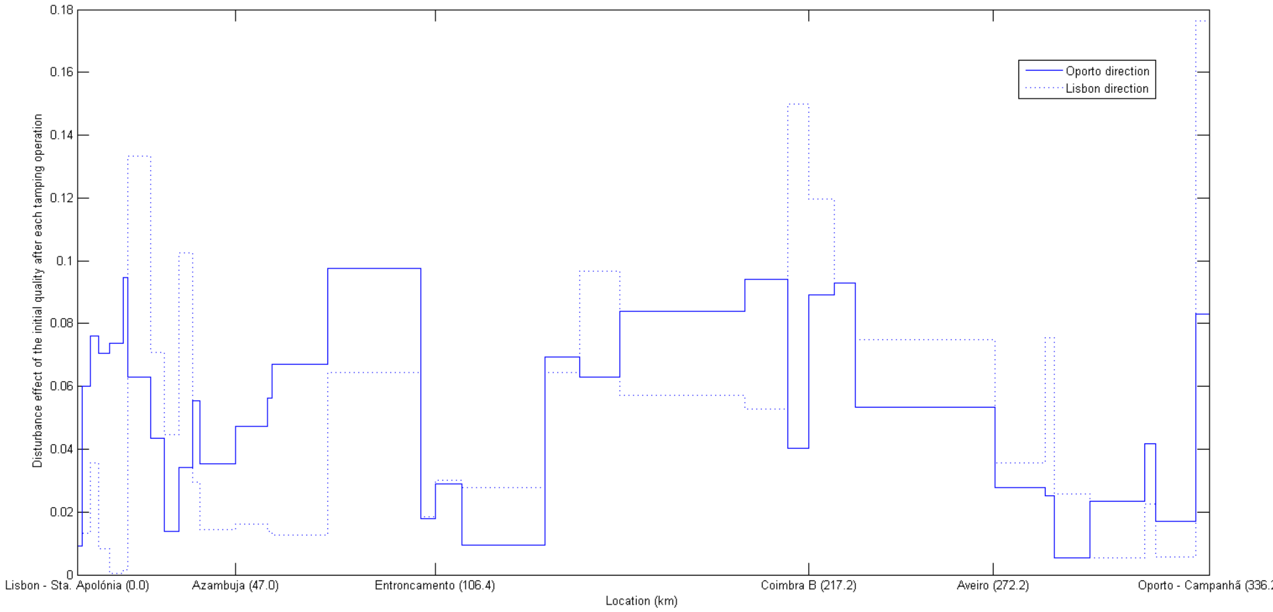

Figure 10 – Disturbance effect (δ) of the initial quality after each tamping operation for the standard deviation of horizontal alignment defects.