



**HAL**  
open science

## Delayed maintenance modelling considering speed restriction for a railway section

Hui Shang, Christophe Bérenguer, John Andrews

► **To cite this version:**

Hui Shang, Christophe Bérenguer, John Andrews. Delayed maintenance modelling considering speed restriction for a railway section. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2017, 231 (4), pp.411-428. 10.1177/1748006X17709200 . hal-01587923

**HAL Id: hal-01587923**

**<https://hal.science/hal-01587923>**

Submitted on 25 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Delayed maintenance modelling considering speed restriction for a railway section

Hui SHANG

Université de Technologie de Troyes, France

Christophe BERENGUER

Univ. Grenoble Alpes, CNRS, GIPSA-lab, F-38000 Grenoble, France

John ANDREWS

University of Nottingham, UK

February 2017

**Warning:** This document is the last version submitted and accepted for publication ; it does NOT integrate the corrections made on the proofs during the final paper production process.

## Abstract

The deterioration of track geometry depends on several factors of which the speed of the train is one. Imposing a speed restriction can slow down the track deterioration and allows a longer survival time before a serious condition is achieved. Preventive maintenance delays can be authorized during the survival time. However, speed restrictions also reduce the system throughput. On the other hand, a longer interval between preventive maintenance activities has a lower maintenance action cost and it also enables grouping the maintenance activities to save set-up costs as well as system down time. If the repair delay is too long, it may cause unacceptable conditions on the track and lead to higher maintenance costs and accidents. Therefore, it is interesting to assess the effect of a speed restriction on the delayed maintenance strategies for a railway track section. We want to solve a maintenance optimization problem to find the optimal tuning of the maintenance delay time and imposition of a speed restriction.

To this aim, a delayed maintenance model is developed, in which track deterioration depends on the train speed and the number of passing trains. The model is used to determine an optimal speed restriction strategy and a preventive repair delay for the optimization of the system benefit and unavailability.

Coloured Petri Nets (CPN) are adopted to model the maintenance and operation of the railway track section. The CPN model describes the gradual track deterioration as a stochastic process. Different speed restriction policies and maintenance delay strategies are modelled and activated by the observed component states. Monte Carlo simulations are carried out to estimate the maintenance cost, the system benefit and the system downtime under different policies. Numerical results show the maintenance decision variable trade-off.

**Keywords** : Delayed maintenance, Coloured Petri Net, speed restriction, performance assessment, railway asset management

# 1 Introduction

The passenger comfort and safety for train travel depends on the track geometry which deteriorates over time. A defective track condition, especially track geometry faults such as poor vertical and horizontal alignment may lead to flange climb and hence derailments. These accidents due to the extreme track condition may cause fatalities and assets damage. In order to prevent these undesired issues, track visual inspection is scheduled to monitor track condition (including track gauge and cant) and appropriate maintenance actions are scheduled to control track quality to ensure safe operation and railway section availability. Railway track maintenance operations can incur significant costs; it is thus important to define the time for track maintenance and renewal to optimize maintenance decisions. Guler proposes a decision making support system for a railway section [1]. Within the decision support system, maintenance plans are scheduled based on the information collected. Several kinds of maintenance models are discussed in the literature, such as risk-based maintenance and age-based maintenance. Podofillini et al. propose a risk-based inspection model for a railway section [2]. Meier-Hirmer et al. discuss rail grinding policies for rail virtual age problem [3]. Antoni compares age-based replacement policies for signalling system [4]. Quiroga et al. consider imperfect tamping models for an age-based maintenance [5, 6]. The maintenance models mentioned above are based on the failure models assuming that the track deteriorates with time only. However, other factors affect the deterioration of railway assets and it can be interesting to consider these factors when building maintenance models. Shafiee et al. propose a usage-based maintenance model for railway track considering the degradation caused by train arrivals, and track usage becomes a decision variable in the model [7]. Zio et al. discuss a maintenance policy for a multi-component railway section, where the deterioration evolution is influenced by the trains speed [8]. Besides maintenance cost, on-time performance is also one of the evaluation criterion in the proposed model.

In other application areas, for example in the manufacturing industry, several research works have aimed at linking productivity, deterioration and maintenance: Tinga compares usage-based maintenance and load-based maintenance for a production system [9]. Yang et al. study a joint scheduling problem of maintenance and throughput considering speed configuration to reach an optimized system benefit [10].

Train speed on a line is also an important factor for both the railway service operation and maintenance. There are two ways to reduce the risk of a train derailment when a poor track geometry exists: preventive maintenance could be carried out as soon as possible, or a speed restriction imposed instead and wait for a delayed, but less expensive maintenance. Whilst a speed restriction extends the track lifetime and enables preventive maintenance to be postponed, it reduces system throughput. Thus, it is interesting to jointly consider operation and maintenance to determine the optimal tuning of maintenance delays and degraded operation. The work presented in this paper is devoted to this problem of the joint optimal tuning of delayed maintenance and the imposition of speed restrictions.

## 1.1 Problem statement

Track system consists of rails, rail joints, fastening system, sleepers, ballasts. Degradation of these main devices may cause track dangerous failures, including track geometry faults and rail failures, which may lead to the train derailment. Rail failures include rail profile problems, rail breakages and rail cracks, which can be fixed by rail grinding and rail renewal. Track gauge, cant, level and alignment are geometry parameters widely used in the literature to describe track geometry condition. There are 4 major kinds of track geometry faults such as track gauge spread, track buckle, track top and twist. Gauge spread due to poor fastening or sleeper condition can be fixed by tie-bar, spot-sleeping and track renewal. Track vertical problems (such as track twist and top) due to poor ballast condition, can be controlled by tamping and stoneblowing. In this paper, track vertical geometry problems (track top and twist) are considered as the dangerous track problems leading to railway accidents. These vertical geometry problems can be identified when track cant measurement exceeds the threshold, and in this work the cant evolution is taken as an illustrative and characteristic deterioration process of the railway track geometry.

The deterioration process for railway track geometry has been studied for a long time, with the objective to find a relationship between influencing factors (such as load, speed, materials, temperature and number of vehicles passing) and the deterioration, and several deterministic models have been proposed for the

deterioration. Kish et al. describe track buckling using a linear equation, which depends on Young's Modulus, rail temperature and train speed [11]. Similarly, a linear equation is proposed to predict the rolling contact fatigue index of the rails according to tonnages and radius of the rail [12]. Zwanenburg introduces a power function to describe the deterioration of track, for which traffic load and train speed are considered to be the main factors [13]. He also proposes another deterioration model for Switches & Crossing (S&C), which is an extension of the power function; the deterioration process of S&C component is based on the load (Million Gross Tonnage), speed, switch angles and so on [14]. All these research works show that a higher speed may lead to faster track deterioration. In addition, the loads also accelerate track deterioration.

A general observation in these works is that speed restriction slows down the deterioration, it allows a longer survival lifetime and hence preventive maintenance delays can be scheduled. Figure 1 shows a schematic view of the track deterioration process in time with and without speed restriction: it can be seen that setting a speed restriction leads to a slower deterioration process and allows for longer waiting time before the condition exceeds the failure threshold than under the normal speeds. However, a speed restriction limits the transportation capability and reduces the system throughput. Figure 1 (bottom) illustrates the relationship between the allowed speed and the system operation gain: normal speed permits higher system throughput while a speed restriction leads to lower train density and hence reduces the system benefit.

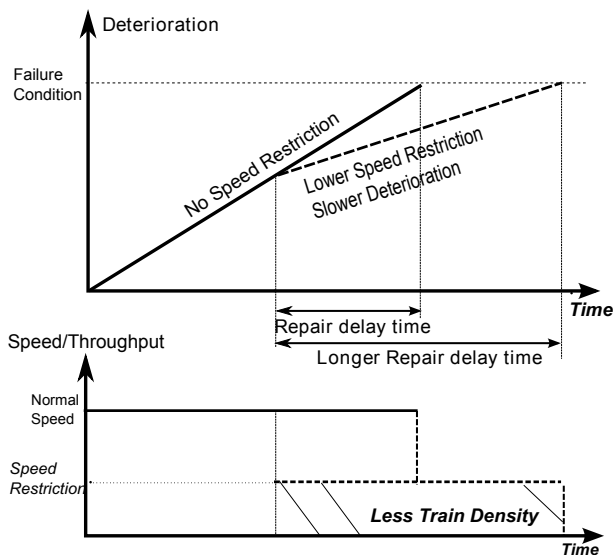


Figure 1: Effect of speed restriction on deterioration and throughput

Allowing for a longer delay before maintenance through introducing a speed reduction can be useful to better plan and organize the maintenance or to wait for the availability of maintenance resources, and thus to perform preventive maintenance at a reduced cost. It also gives more opportunities to group maintenance activities, hence saving the set-up cost and reducing the downtime due to the repairs. But reducing the speed limitation also leads to less trains per hour and thus a lower system throughput. However, speed restrictions cannot stop the geometry deterioration, so even under speed restrictions, there is still a risk that the track deterioration reaches the failure threshold and leads to corrective maintenance or an accident while waiting for the delayed preventive maintenance. Consequently, an optimal speed restriction has to be found to have a maximum system benefit and, at the same time, an appropriate preventive maintenance delay should be planned to meet the safety and operational requirements of the system.

The aim of this paper is to solve a tuning problem between speed restrictions and repair delays in order to achieve a trade-off between the imposition of a speed restriction and the duration of a preventive maintenance delay to reach an optimized system benefit and availability for a plain line railway section. It considers that the track deteriorates depending on the speed and the number of trains. A maintenance model is developed to describe the component failures and system behaviour of a multi-component railway section to solve the considered maintenance optimization problem.

## 2 Maintenance modelling and performance assessment framework

### 2.1 Global model structure

The objective of the model developed is to represent both the system behaviour (operation rules, and deterioration phenomena) and the effects of the maintenances procedures. Both sides of the model (system behaviour and maintenance) are linked through the component states. In order to assess the system performance, a two-level modelling framework is adopted to describe the system behaviour, component states and the interaction of the related maintenance and inspection processes.

The track component level, is structured around three models.

1. The *component states model* describes, for each component, the failure modes and the associated deterioration processes;
2. The *component maintenance model* includes maintenance requirements, different maintenance actions and their effects on the component state;
3. The *component operation model* describes the operation rules for the passing trains.

The railway section level (or “system level”) consists of three models.

1. *The system operation model* describes the system behaviour and the operational rules at the system level. For a railway section, the system operation rules and behaviours depend on the system structure, such as a plain line system or a system consisting of S&C. The system operation model is linked with the component operation models and captures the interactions between maintenance, inspection and component deterioration.
2. *The inspection process model* describes the inspection process for a section and its components. This inspection model needs to consider the inspection plans based on the system structure.
3. *Section maintenance model*: A multiple component section needs to consider a system maintenance decision and grouping strategies, it sends the maintenance decision of each component and waits for the system maintenance requirements.

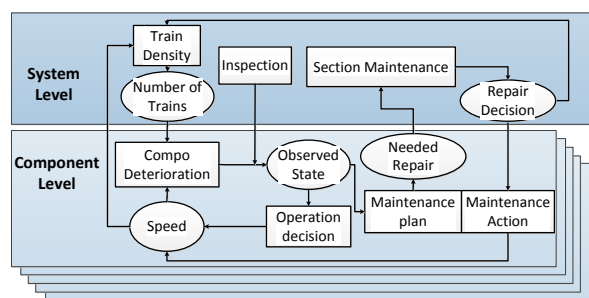


Figure 2: Framework of the multi-level model

According to the above description, the model structure for a railway section consisting of 5 track components is shown in Figure 2.

The upper part of Figure 2 represents the structure of the “system level” model. For the considered section, the “system level” model collects the repair requirements and operation rules from component models, and at the same time, it sends the maintenance decisions and operation status to the related component models. Each of the five model boxes in the lower part of Figure 2 shows the structure of the “component level” model for each track component, including component deterioration, operation decision and maintenance actions. In particular, the deterioration and its influencing factors, and the component maintenance actions are represented at this level.

## 2.2 Track component deterioration and maintenance model

The track deterioration depends on the number of passing trains and the speed of the trains. Two levels of maintenance and three speed options are scheduled for the periodically observed track condition.

In this paper, we consider track vertical geometry faults as a potentially hazardous event leading to the train derailment. As shown in Figure 3, track vertical geometry faults happen if the difference between two rails ( $cant(x)$ ) exceeds a required limits. Twist is defined to be  $cant(x) - cant(x - b)$  for a given length  $b$ .

If twist exceeds the limit value, passing trains may climb the flange and lead to derailments.

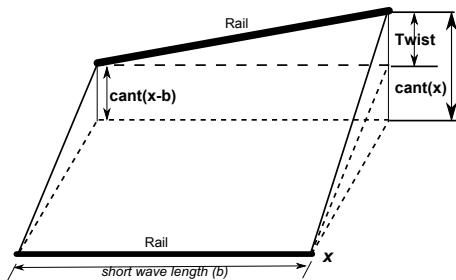


Figure 3: Illustration of track vertical faults

### 2.2.1 Track component deterioration modelling

For our maintenance evaluation objective, a deterioration process model taking explicitly into account influencing factors such as speed and MGT is required. Several deterioration models considering influencing factors have been proposed in the literature. Sadeghi et al. assume a track deterioration model considering the environment factors of the track; this model is called Track Quality Indices (TQI), which is an integer representation of track quality [15, 16]. However, the index cannot show the relationship between the deterioration and the maintenance actions. Westgeest proposes a regression model for the Key Performance Indicators (KPIs) to describe the degradation of the track geometry depending on local circumstances and the KPI maintenance models, where the KPIs are used to describe the track quality within a certain length [17]. The linear regression model, which is obtained by regression analysis according to the observation data, was used to relate the deterioration to the tonnage, tamping, subsoil and switches by Andrade [18].

These deterministic models describe the relationship between the influencing factors (environment, physical characters and structure) and the deterioration. Even though the evolution can be described in these models, they lack the ability to model the variability and the randomness of the evolution since the deterioration increments are determined and cannot model the varied deteriorations behaviour probabilistically.

Probabilistic models can be used for track lifetime modelling. For example, the Weibull distribution is used widely to describe the ageing track component lifetime. It is used to describe the ballast ageing [19, 20, 21, 22]; Antoni uses it to describe the aging behaviour of signalling devices [4]; Meier-Hirmer proposes that a residual life function of rail defects follows a Weibull distribution [3] and Patra et al. use it to describe the occurrence of the rail defect at time  $t$ , because Weibull distribution can represent many kinds of shapes by choosing different parameters, and hence many different ageing behaviours [23]. According to the paper by Audley et al., the Weibull distribution with two parameters gives the best performance to model the failure process of track with tamping in the railway [24].

However, such probabilistic lifetime models do not take into account explicitly any deterioration phenomena, and they only consider binary states of the devices: good or failed. If deterioration information can be collected, it can be worthwhile to use a stochastic process to describe the gradual degradation process. For example, the Gamma process is discussed to support maintenance decisions for a deterioration process [25]. Meier-Hirmer et al. propose an imperfect maintenance model for deteriorated track [26].

In order to represent the gradual deterioration of the track (the twist at time  $t$ ), and to take into account its influencing factors (i.e. trains number and speed), we build the following model. The trains enter the track section following a Poisson process with intensity  $\lambda$ . Denoting  $N_t$  the number of trains entering the

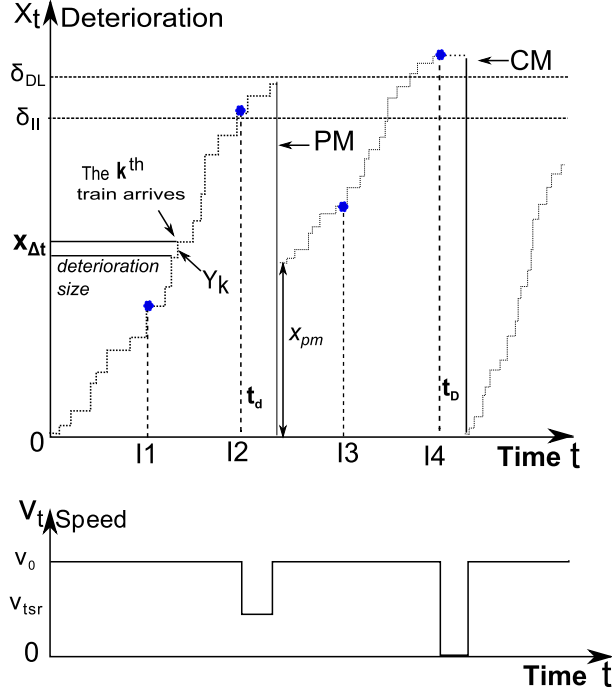


Figure 4: Relationship of system gain, deterioration and maintenance for a track component

track until time  $t$ ,  $X_t$  the state of the track (i.e. the sum of deteriorations) at time  $t$  is given by

$$X_t = \sum_{k=1}^{N_t} Y_k, \quad Y_k \sim \Gamma(\alpha(v), \beta) \quad (1)$$

where  $Y_k$  is the  $k^{th}$  random deterioration increment due to the  $k^{th}$  train passing on the section.  $Y_k$  follows a Gamma distribution  $\Gamma(\alpha(v), \beta)$ , whose shape parameter  $\alpha(v)$  depends on the speed  $v$  :

$$\alpha(v) = \alpha_0 e^{a_0 v} \quad (2)$$

The deterioration process is thus a Compound Poisson Process with Gamma distributed jump sizes. The deterioration increment for a time period  $\Delta t$  is given by:

$$X_{\Delta t} = X_{t+\Delta t} - X_t = \sum_{k=1}^{N_{\Delta t}} Y_k = \Gamma(\alpha(v)N_{\Delta t}, \beta) \quad (3)$$

$X_{\Delta t}$  represents the increment of  $X_t$  during time  $\Delta t$ , the increment depends on the number of passing trains during  $\Delta t$ . Equation 3 shows that the increment can be described using a Gamma distribution where the shape parameter is  $\alpha(v)N_{\Delta t}$ .

### 2.2.2 Track inspection modelling

Visual inspection cars can run on the line to measure the useful data for condition estimation. Figure 4 illustrates the periodic inspection on a operating calendar; “ $I_i$ ” represents the time that an inspection arrives. The inspection interval on the operation calendar is  $\theta$ . Visual inspection is assumed to be able to identify the states at time  $i\theta$ , the state identified may trigger corresponding maintenance decision.

### 2.2.3 Track maintenance model

Several techniques can be used to fix track vertical faults, such as stoneblowing, tamping and rail lifting. It has been shown that these maintenance techniques cannot fix the track to an as-good-as-new. We consider



here that tamping is the preventive maintenance action and stoneblowing is the corrective one. For example, in Figure 4, if a track defect is detected at time  $I_2$ , a preventive maintenance is arranged; once inspection identifies a track fault at time “I4”, a corrective maintenance is performed.

After a preventive maintenance (tamping), the track is to condition  $x_{pm} > 0$  that is not as-good-as-new. We assume that a corrective maintenance action can repair the track to an as-good-as-new state ( $X_{t+t'} = 0$ ) as shown in Equation 4. Figure 4 explains the triggers of preventive maintenance and corrective maintenance for defective track component. On the operation calendar, the effect of maintenance is shown in Equation 4.

$$X_{t+t'} = \begin{cases} x_{pm}, & \text{After PM} \\ 0, & \text{After CM} \end{cases} \quad (4)$$

At time  $I_2$  in Figure 4,  $X_{2\theta} \geq \delta_{II}$ , a PM is scheduled after time  $t_d$ ; after PM, at time  $t_d + I_2$ , the state of track  $X_{I_2+t_d} = x_{pm}$ . Similarly, after corrective maintenance, at time  $I_4 + t_D$ , track state  $X_{I_4+t_D} = 0$ .

### 2.3 Railway section model

The chosen railway section is a multi-component section, and as explained when introducing our global modelling framework, it is necessary to consider the behaviour of the whole system for the maintenance and performance modelling.

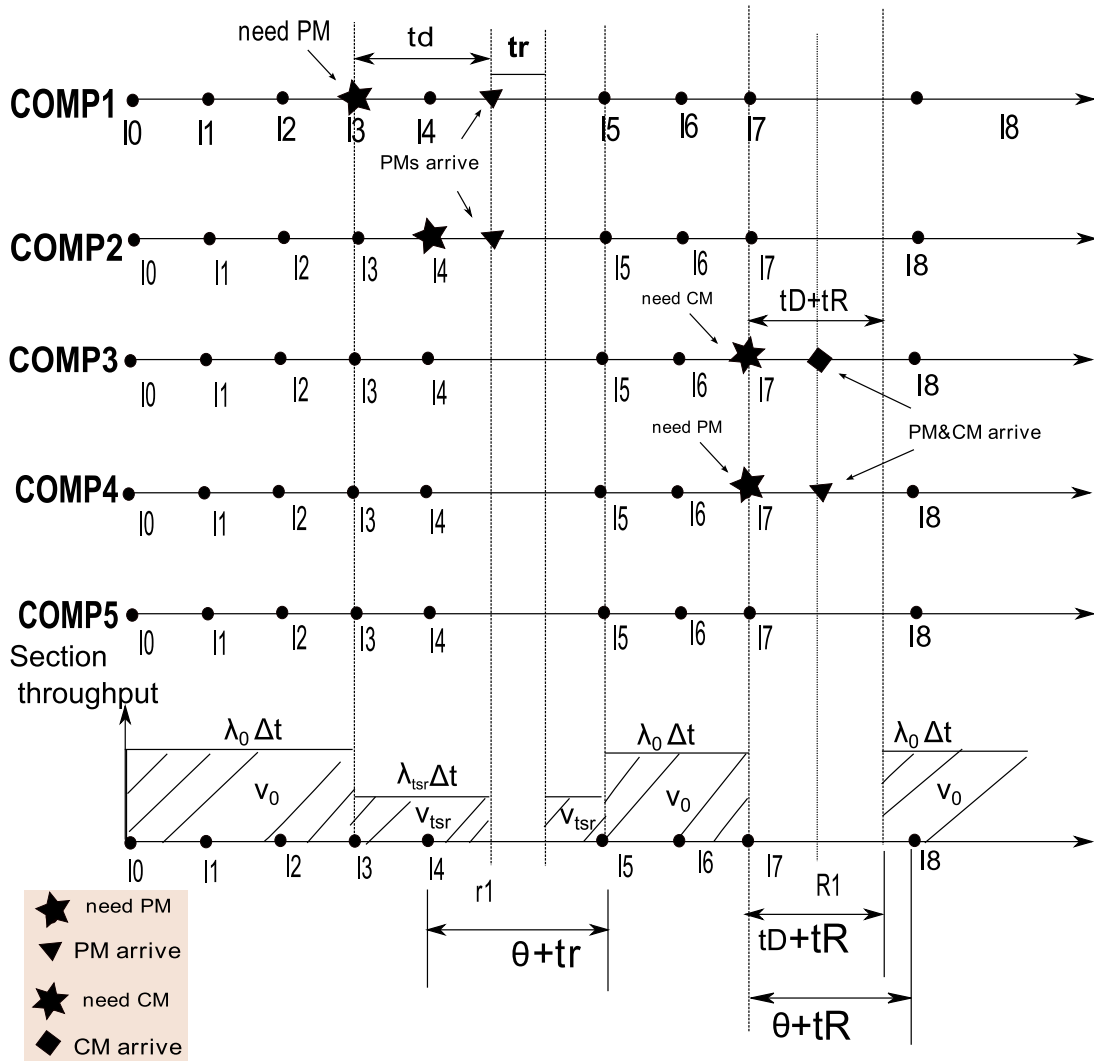


Figure 5: Scenarios of maintenance grouping for the section

### 2.3.1 Railway section operation rules

Imposing a speed restriction is one method to prevent, or at least reduce, the risk of derailment. We consider three situations as shown in Equation 5 and Figure 4. If the identified track condition exceeds the preventive maintenance threshold (i.e.  $X_t \geq \delta_{II}$ ), a speed restriction is set until the inspection results show that the condition is returned under the PM threshold. If the track condition exceeds the corrective maintenance threshold (i.e.  $X_t \geq \delta_{DL}$ ), the railway section is closed, thus the speed is set to be 0. Train speed is returned to normal after the corrective maintenance is completed.

$$v_{i\theta+\Delta t} = \begin{cases} v_0, & X_{i\theta} \in (0, \delta_{II}), \Delta t < \theta \quad \text{Normal} \\ v_{tsr}, & X_{i\theta} \in [\delta_{II}, \delta_{DL}), \Delta t < \theta \quad \text{Limited} \\ 0, & X_{i\theta} \in [\delta_{DL}, \infty), \delta t < \theta \quad \text{Closed} \end{cases} \quad (5)$$

A simple train density function  $\lambda(t)$  at time  $t$  for a component is considered for the chosen section, which depends on train speed  $v_{tsr}$  in Equation 6.  $v_0$  is the normal speed,  $\lambda_0$  is the normal train density under train normal speed.

$$\lambda(t) = \frac{v_t}{v_0} \lambda_0 \quad (6)$$

The system train density function determines the loads experienced by any section and system transportation service gain. The train density for the entire chosen section  $\lambda_{sys}(t)$  is shown in Equation 7.  $\lambda_j(t)$  represents the density for the  $j$ th component.

$$\lambda_{sys}(t) = \text{Min}\{\lambda_1(t), \dots, \lambda_5(t)\} \quad (7)$$

In Figure 5, at time  $I3$ , component 1 is detected as defective and a speed restriction should be set on component 1, the minimal speed of the section is  $v_{tsr}$ , the number of trains during  $\Delta t$  is changed to  $\lambda_{tsr} \Delta t$ .

If we consider the down time for performing the repair, inspection is not periodic, as shown in Figure 5.

### 2.3.2 System maintenance strategies

The maintenance decisions for a multi-component section needs to consider the states of all components. The maintenance cost then includes the maintenance set-up cost  $c_{psu}$  (or  $c_{csu}$ ) and maintenance action cost  $c_{pm}$  (or  $c_{cm}$ ). The maintenance set-up cost arises whenever preventive maintenance or corrective maintenance is planned, it depends on the number of maintenance set-ups. The maintenance action cost depends on the number of components to be maintained in the section. If several preventive maintenance actions are carried out at the same time, only one PM set-up cost is incurred.

In addition, we assume that the preventive maintenance action cost  $c_{pm}$  is a function of the waiting time  $t_d$ : a longer maintenance delay allows to reduce the maintenance through for example: better planning and the more efficient use of the resources, hence the cost is lower as shown in Equation 8.

The overall cost of preventive maintenance includes the set-up cost  $c_{psu}$  and the cost of each preventive maintenance action  $c_{pm}(t_d)$ .  $c_{pm}(t_d)$  depends on the longest repair delay time in the group, that is  $t_d$  for component 1 at time  $I3$  in Figure 5. The longer  $t_d$  the lower maintenance cost as shown in Equation 8. If  $t_d \in (0, T^*)$ , the preventive maintenance action cost depends on the repair delay time  $t_d$ , if  $t_d \geq T^*$ , the preventive maintenance action cost equals a minimum repair value  $c_{pmin}$ .

$$c_{pm}(t_d) = \begin{cases} At_d + B & t_d \in (0, T^*) \\ c_{pmin} & t_d \in [T^*, \infty) \end{cases} \quad (8)$$

In a delayed maintenance situation, repair delays gives more opportunities to group repairs together. At the system level, we can adopt different grouping or non-grouping strategies.

**Grouping maintenance strategy** - As shown in Figure 5, the grouping maintenance strategy aims to combine possible maintenance actions to reduce the set-up cost. In Figure 5, two scenarios for the grouping maintenance strategy are shown. “COMP 1” is detected defective at time “I3” and there is a preventive maintenance to be carried out at time  $I3 + td$ , during the waiting time, “COMP 2” needs preventive maintenance, then at time  $r1$ , preventive maintenance for both “COMP 1” and “COMP 2” is carried out which takes  $t_r$  time.

There is another situation of grouping maintenance strategy: once the inspection identifies a necessary corrective maintenance, for example “COMP 3” at time  $I7$ , the section is closed immediately, if there are preventive maintenance requirements for other components, the grouping strategy may arrange these two kinds of maintenance to be performed at the same time.

**Non-grouping maintenance strategy** - Under a non-grouping maintenance strategy, we assume that the preventive maintenance actions are performed strictly with each delay time. For example, for component 2, the maintenance action will be carried out at time  $I4 + t_d$ . Similar to grouping strategy, corrective maintenance requires to close the railway section, thus the preventive maintenance for “COMP4” will be carried out together with the corrective maintenance between  $I7$  to  $I7 + t_D + t_R$ .

## 2.4 System performance evaluation

Two performance evaluation approaches can be followed to determine the optimal decision variables and find out the optimal system performance: single-objective and multiple-objective evaluation method.

### 2.4.1 Single-objective evaluation

The system benefit on an infinite time span  $\mathbb{E}B_\infty$  and the system unavailability  $\mathbb{E}Q_{avg}$  are used to evaluate the system performance. The system unavailability is given in Equation 9:

$$\mathbb{E}Q_{avg} = \lim_{t \rightarrow \infty} \frac{[N_{csu}(t) + N_{mix}(t)](t_D + t_R) + N_{psu}(t)t_r}{t} \quad (9)$$

The system downtime consists of three parts: waiting time for corrective maintenance  $t_D$ , preventive maintenance time  $t_r$ , corrective maintenance time  $t_R$ .

The system benefit  $B(t)$  is the system gain  $G(t)$  minus the system maintenance cost  $C(t)$  as shown in Equation 10.

$$\begin{aligned} \mathbb{E}B_\infty &= \lim_{t \rightarrow \infty} \frac{G(t) - C(t)}{t} \\ &= \mathbb{E}G_\infty - \mathbb{E}C_\infty \end{aligned} \quad (10)$$

The average maintenance cost rate ( $\mathbb{E}C_\infty$ ) and the system gain rate ( $\mathbb{E}G_\infty$ ) on an infinite time span are shown in Equation 11 and Equation 12. The system gain  $G(t)$  depends on the number of passing trains  $N_{train}(t)$ .  $c_{train}$  is the profit for a passing train.

$$\mathbb{E}C_\infty = \lim_{t \rightarrow \infty} \frac{C(t)}{t} \quad (11)$$

$$\mathbb{E}G_\infty = \lim_{t \rightarrow \infty} \frac{N_{train}(t)c_{train}}{t} \quad (12)$$

The maintenance cost consists of the set-up cost, maintenance action costs and the inspection cost, as shown in Equation 13.

$$\begin{aligned}
C(t) &= \underbrace{\sum_{m=1}^{N_{csu}(t)} (c_{csu} + n_m c_{cm})}_{\text{Pure corrective maintenance group}} + \underbrace{\sum_{n=1}^{N_{psu}(t)} (c_{psu} + n_n c_{pm})}_{\text{Pure preventive maintenance group}} \\
&+ \underbrace{\sum_{h=1}^{N_{mix}(t)} (c_{csu} + n_{cm,h} c_{cm} + c_{psu} + n_{pm,h} c_{pm})}_{\text{Mix maintenances group}} + N_{insp}(t) c_{insp} \\
&= [N_{csu}(t) + N_{mix}(t)] c_{csu} + \left( \sum_{n=1}^{N_{psu}(t)} n_n + \sum_{h=1}^{N_{mix}(t)} n_{pm,h} \right) c_{pm} \\
&+ [N_{psu}(t) + N_{mix}(t)] c_{psu} + \left( \sum_{m=1}^{N_{csu}(t)} n_m + \sum_{h=1}^{N_{mix}(t)} n_{cm,h} \right) c_{cm} \\
&+ N_{insp}(t) c_{insp}
\end{aligned} \tag{13}$$

In order to evaluate the average system benefit rate  $\mathbb{E}B_\infty$  and the average system unavailability  $\mathbb{E}Q_\infty$ , a simulation model is used to collect the number of trains  $N_{train}$ , the number of inspections  $N_{insp}$ , the number of maintenance set-ups ( $N_{mix}$ ,  $N_{csu}$  and  $N_{psu}$ ), the number of maintenance actions ( $N_{cm}$  and  $N_{pm}$ ).

### 2.4.2 Multiple-objective evaluation

Single objective evaluation only shows the relationship between the system benefit and the decision variables or the relationship between the unavailability and the decision variables. In some cases, the maintenance decision maker has to jointly consider the benefit and the unavailability. The multiple objective problem evaluates the system performance considering both the system benefit and unavailability at the same time, considering for example a Pareto front representation [27, 28]. We want to find the solution set  $(v_{tsr}, t_d)$  to satisfy:

if and only if there is no  $(v_{tsr}^*, t_d^*) \in (V, T_d)$

$$B(v_{tsr}, t_d) \geq B(v_{tsr}^*, t_d^*) \tag{14}$$

$$Q(v_{tsr}, t_d) \leq Q(v_{tsr}^*, t_d^*) \tag{15}$$

The solution set  $(v_{tsr}, t_d)$  is called strict Pareto front. If  $B(v_{tsr}, t_d) > B(v_{tsr}^*, t_d^*)$  and  $Q(v_{tsr}, t_d) < Q(v_{tsr}^*, t_d^*)$ , the solution set is called weak Pareto front.

## 3 Model implementation using Coloured Petri Nets

Coloured Petri Nets are a high level reliability modelling and performance assessment tool. Petri Nets have been used to describe dynamic system behaviour, they are proposed to model railway track maintenance [20, 19]. These Petri Net models have their limitation to describe the component gradual deterioration process. Coloured Petri nets rely on coloured sets to classify different token conditions, a model has used CPN to describe gradually deterioration for railway section maintenance modelling[29].

### 3.1 Modelling tool: Coloured Petri Nets

A Coloured Petri Net (CPN) consists of 9 elements as shown in Equation 16 [30].

$$CPN_t = (P, T, A, \Sigma, V, C, G, E, I) \tag{16}$$

$CPN_t$  is a timed CPN,  $P$  is set of places,  $T$  is set of transitions,  $A$  is set of arcs,  $\Sigma$  is set of coloured set,  $V$  represents the set of variables.  $G$  is the set of guard functions,  $I$  is the initial marking of the net,  $C$

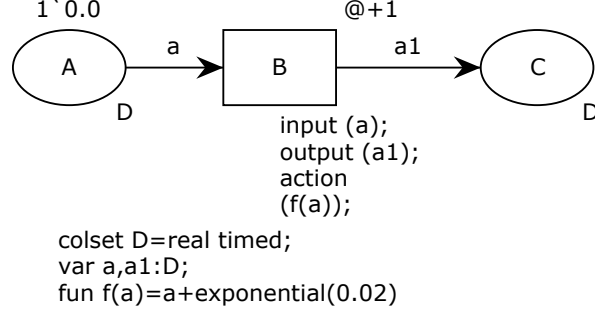


Figure 6: Elements in Coloured Petri Nets

is the set of coloured functions and  $E$  is the set of the arc expression functions,  $E(p, t)$  represents the arc from a place to a transition and  $E(t, p)$  represents the arc from a transition to a place.

The above elements can be represented in graphs: circles represent places, such as “A” and “C” in Figure 6. Rectangles represent transitions (“B” in Figure 6). Places and transitions are connected by arcs, represented by arrows with single or double direction. Besides places, transitions and arcs, Figure 6 shows other elements in CPN: “D” is a timed color set, the type is real; the tokens in them belongs to color set “D”; variables “a” and “a1” belongs to color set “D”, they bring the value of tokens from A to B or from B to C. In addition, “B” carried out function  $f(a)$  when it fires. Initial marking of “A” is 0.0 and initial marking of “C” is empty  $\Phi$ .

### 3.2 Marking process and data collecting process in CPN

The marking process and the data collecting process in a CPN are useful for system performance evaluation. The marking of a timed CPN is defined as shown in Equation 17.

$$M(p) \in C(p)_{\text{TMS}}, p \in P \quad (17)$$

The timed marking of place  $p$  at time  $t^*$  is a pair  $(M(p), t^*)$ , where  $C(p)_{\text{TMS}}$  in equation 17 is a timed coloured set with timestamps, and  $t^*$  is the value of global clock.

Once the binding elements of a transition satisfy the firing condition, the transition fires and the marking changes. For example, in Figure 6, the binding element for transition B firing is  $M(A) \neq \Phi$ . Once there is a token in place “A”, transition B fires and changes the marking of the system as follows:

$$(M_0(A), 0) \xrightarrow{B} (M_1(A), 0) \Rightarrow (0.0, 0) \xrightarrow{B} (\Phi, 0).$$

There are two ways to find the optimal solution with CPNs, one way is to find the solution via a Reachability Graph (RG), but the RG for a timed CPNs is too large to find the solution. Another way produces an approximate solution by running simulations; the data collection process for the CPNs is used to monitor system behaviour and collect data for performance evaluation.

There are 3 useful functions for data collection in CPN tools: function  $pred()$  is used to determine when to collect data; function  $obs()$  collects the data we need and function  $breakpoint()$  stops simulations as shown in Figure 7.  $M_i, M_j$  and  $M_k$  are markings, and  $(t, b)$  represents the combination of transition and binding elements between two markings.

The function “ $pred()$ ” is shown in Table 1, “ $size(New\_Page'B\_1\_mark) \geq 1$ ” means the marking of place B is larger than 1 and it turns on the monitor; if not (i.e. “ $|obsBindElem\_ = false$ ”), the monitor is off.

if the marking  $M_j$  appears during the simulation, functions  $obs()$  collects the data, we can collect the number of marking  $M_j$ , the value of tokens in the marking  $M_j$  and so on with the function  $obs()$

Function  $breakpoint(M_k)$  helps us to stop the simulation once there is marking  $M_k$ .

### 3.3 CPN models for 5-components railway section

Using coloured Petri nets, we implement a maintenance model for a section consisting of 5 components using the section global assumptions and framework shown in Figure 2.

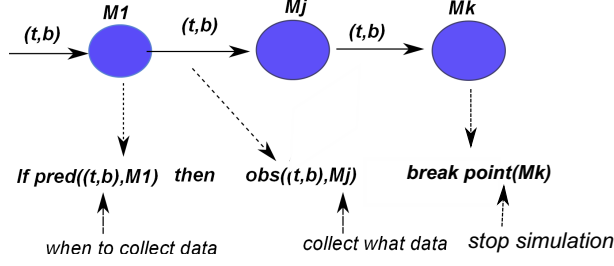


Figure 7: Data collection process in CPNs

Table 1: Function  $\text{pred}()$  definition

The Predicate function:	
fun	$\text{pred}(\text{bindelem}, M1 : \text{type}) =$
let	
fun	$\text{predBindElem}(Mj(t,b))$
	$= \text{size}(M1) \geq 1$
	$ \text{predBindElem}_- = \text{false}$
in	
	$\text{predBindElem bindelem}$
end	

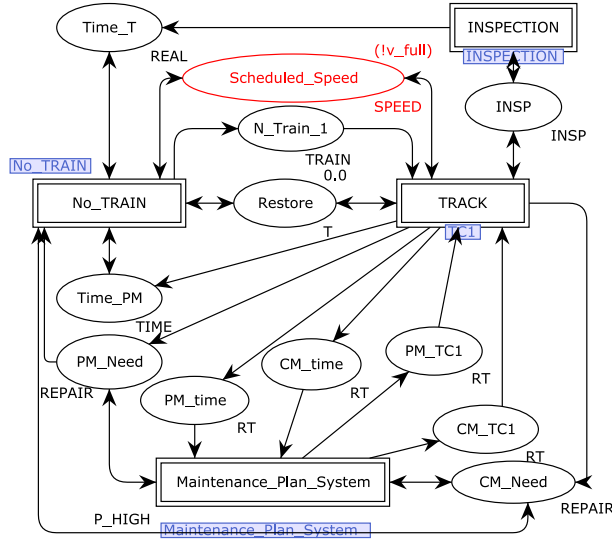


Figure 8: System level CPN model with one TRACK component module

Figure 8 shows a part of the CPN model. Each substitution transition represents the corresponding modules in Figure 2: substitution transition “No\_TRAIN” models “Train density” at system level in Figure 2; substitution transition “TRACK” models component level and substitution transition “Maintenance\_Plan\_System” is for the module “Section Maintenance” in Figure 2. There are still four substitution transitions “TRACK\_1” for the whole model which are not shown since they are the same as “TRACK”, and there are places for each “TRACK\_1”, for example, place “N\_Train\_1”, “Scheduled\_Speed  $i$ ”, “PM\_TC1” and “CM\_TC1”.

### 3.3.1 Track deterioration CPN model

Figure 9 shows the CPN model for track deterioration in substitution transition “TRACK”, initial marking for the place “Track” is  $(\text{state}, t) = (0.0, 0.0)$ , which means at time  $t = 0.0$ , the track cant state = 0.0.

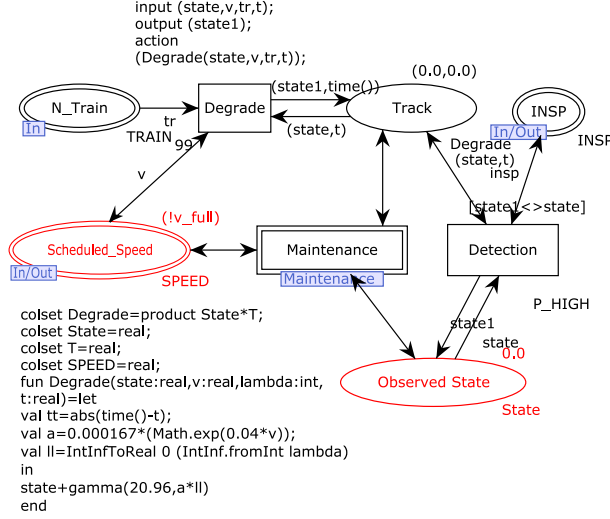


Figure 9: Track deterioration CPN model

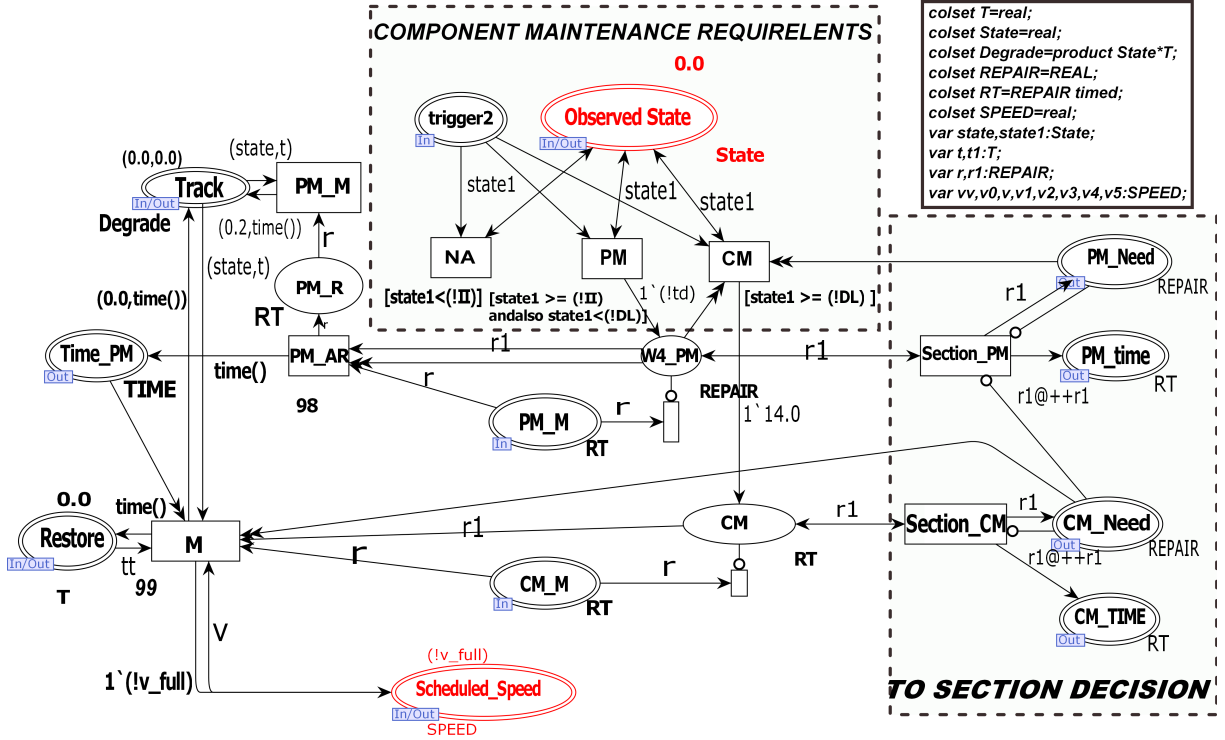


Figure 10: Component Maintenance CPN module

Then track condition for any time  $t$  is modelled, the deterioration increment for every  $\Delta t$  is generated by transition “Degrade” in Figure 9. In this paper, we assume that the deterioration of 5 components in the chosen section depends on speed and the number of passing trains. In order to model varied deterioration evolutions with the same characteristics factors for the 5 components, function  $Degrade(state : real, v : real, lambda : int, t : real)$  is defined according to Equation 3. Token in place “N\_Train” represents the number of trains during a period and token in place “Scheduled\_speed” represents the passing train speed. “IntInfToReal 0  $a$ ” changes the data type of  $a$  to be real type; similarly, IntInf.fromInt  $a$  is used to change integer variable  $a$  to be infinite integer. In function  $Degrade(state, v, lambda, t)$ , function “gamma( $\beta, \alpha$ )” generates a random value following a Gamma distribution with a shape parameter  $\beta$ , and scale parameter  $\alpha$ . The random value represents the deterioration increments, which is varied for  $\Delta t$  and for every substitution

transition “TRACK”, and thus track components have varied deterioration behaviours.

Binding elements for transition “Degrade” is  $M_{t_{train}}(N\_Train) \neq \Phi$ ,  $M_{t_{train}}(Track) \neq \Phi$  and  $M_{t_{train}}(Scheduled\_S)$ . According to the marking process in CPN, the marking of place “Track” can be:

$$\begin{aligned}
 & \left( \underbrace{(0.0, 0.0)}_{\text{marking } M(\text{track})=(\text{state}, t)}, \underbrace{0.0}_{\text{global time } t^*} \right) \xrightarrow{\text{Degrade}} \\
 & \left( \underbrace{(0.0 + \text{gamma}(20.96, a * ll))}_{\text{state1}}, \underbrace{\text{time}(), t_{train}^*}_{t_1^*} \right)
 \end{aligned} \tag{18}$$

In Figure 9, component inspection is also modelled by transition “detection”: once there is a token in place “INSP” and state  $\neq$  state1, transition “Detection” fires and updates marking in place “Observed State”. Since both the transition “Degrade” and the transition “Detection” need tokens in place “Track”. The transition firing priority is used to order the firings. It is assumed that transition “Degrade” has a higher firing priority than transition “Detection”, which means that if the binding elements for both transitions are satisfied, transition “Degrade” fires before transition “Detection”. This represents that track deteriorates before the inspection arrives.

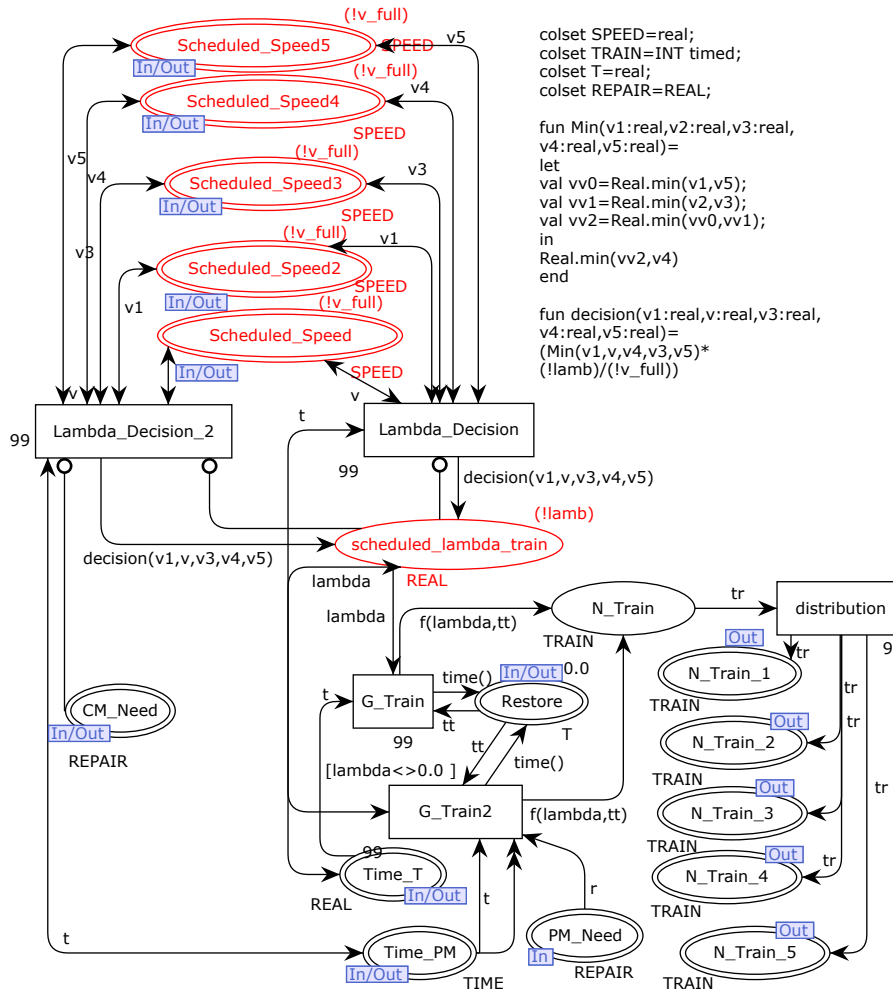


Figure 11: System operation CPN model describing the number of passing train according to train speed on the section



### 3.4 Component maintenance CPN model

The component maintenance requirements are modeled at component level in the system CPN model. Figure 10 is the CPN model for grouping maintenance at component level. Place “Track” here is the same as in Figure 9, a token in this place represents the track component condition.

Figure 10 illustrates the details of the component maintenance module with the grouping strategy (substitution transition “Maintenance” in Figure 9), including the maintenance plan and the maintenance action. The component maintenance requirement depends on the token in place “Observed State” which is the inspection result for a component. Since place “Observed State” belongs to a untimed coloured set and there are two arcs with two directions linking transition “NA” ( or “PM” or “CM”) and place “Observed State”, the untimed coloured set and the arcs may trap the CPN model into a dead-lock, place “trigger2” has a token after an inspection which helps to control the component decision transitions firing after the inspections.

Three transitions (“NA”, “PM” and “CM”) are used for decision making, the transition guards decide the firing condition: “!II” represents the preventive maintenance threshold  $\delta_{II}$  and “!DL” is the corrective maintenance threshold  $\delta_{DL}$ .

If the observed state  $state1$  satisfies the guard function  $state1 \in [\delta_{II}, \delta_{DL}]$ , transition “PM” fires and the delay time “!td” is sent to place “W4\_PM”. If this is the first preventive maintenance requirement since the last corrective maintenance, transition “Section\_PM” fires and it sends a token to place “PM\_Need” which indicates the section needs preventive maintenance, and the preventive maintenance time is set in place “PM\_time”. If this is not the first preventive maintenance requirement on the section since previous corrective maintenance then there was at least one defect identified previously and the section is now awaiting the maintenance. Thus the transition “Section PM” does not fire. Place “PM\_M” is the preventive maintenance decision from the system level. If there is a valid token in place “PM\_M” and the tokens in place “W4\_PM”, transition “PM\_AR” fires and consumes all the tokens in place “W4\_PM”. Transition “PM\_M” fires and recovers the track state to be 0.2.

This model also considers the situation in which the repair delay is longer than the inspection interval, i.e.  $t_d \geq \theta_{insp}$ . The inspection results do not change the preventive maintenance time and the reset arc helps to empty the place “W4\_PM”.

If the inspection identifies the failure  $state1 \geq (!DL)$ , or the defective state deteriorates to be a failure, transition “CM” fires and makes place “W4\_PM” and “PM\_Need” to be empty, which means that there is no PM requirement. At the same time, the section needs corrective maintenance (token in place “CM\_Need” and the maintenance time is controlled by tokens in place “CM\_TIME”). Place “CM\_M” represents the section decision. Once there are tokens in both place “CM” and “CM\_M”, transition “M” fires and recovers the track to be as good as new (i.e.  $state = 0.0$ ).

A monitor of “M” is set to collect the number of corrective maintenance activities, i.e.  $(\sum_{m=1}^{N_{csu}(t)} n_m + \sum_{h=1}^{N_{mix}(t)} n_{cm.h})$ . Similarly, a monitor of “PM\_M” is used to collect the number of preventive maintenance activities, i.e.  $(\sum_{n=1}^{N_{psu}(t)} n_n + \sum_{h=1}^{N_{mix}(t)} n_{pm.h})$ .

#### 3.4.1 System operation CPN model

In order to model the effect of MGT on track deterioration, system behaviours such as the number of passing trains and train speed are described in system operation CPN model.

Figure 11 shows the details of the system behaviour CPN model. It models the number of trains passing the chosen section for a period, the number of trains for the whole section is represented by the token value in place “N\_Train”. The token in place “scheduled\_lambda\_train” represents the train density; the initial marking for it is “!lamb”. Transition “G\_Train” calculates the number of trains passing with normal speed the section in a duration  $t$  which is the input place “Time\_T” from upper CPN module. The duration is the inspection interval  $\theta$ . When there is preventive maintenance, the duration is not the inspection interval, so transition “G\_Train2” generates the number of trains for the degraded operation modes. Function  $f(lamb : real, t : real)$  models the random arrivals of trains with the arrival rate  $\lambda$ . When there is a corrective maintenance needed or repairs are carried out, the railway line is closed so there is no trains and the number of trains is not needed.

Since the section is a plain line section, the speed of the whole section depends on the minimum speed. Transition “lambda\_Decision\_2” and “lambda\_Decision” determines the speed for the section with a function “decision(v1,v2,v3,v4,v5)”. In addition, the number of trains are the same for each component in this series section, transition “distribution” sends the number of trains from place “N\_Train” to “N\_Train\_i”.

We set a monitor with both “G\_Train” and “G\_Train2” to collect the total number of trains  $N_{train}$ .

### 3.4.2 System inspection CPN model

Figure 12 shows the inspection CPN model, the initial marking of place “INSP\_n” is a token with the timestamp “!theta”. Once the model global time reaches “!theta”, transition “INSP\_Arr” fires and sends the inspection arrival time to place “Time\_T”, which is an output of the inspection module. At the same time, each component module can have an inspection arrival token in place “INSP i” to enable the detection process. Transition “INSP\_lv” represents the inspection car leaving the chosen section and enables another inspection after “!theta”.

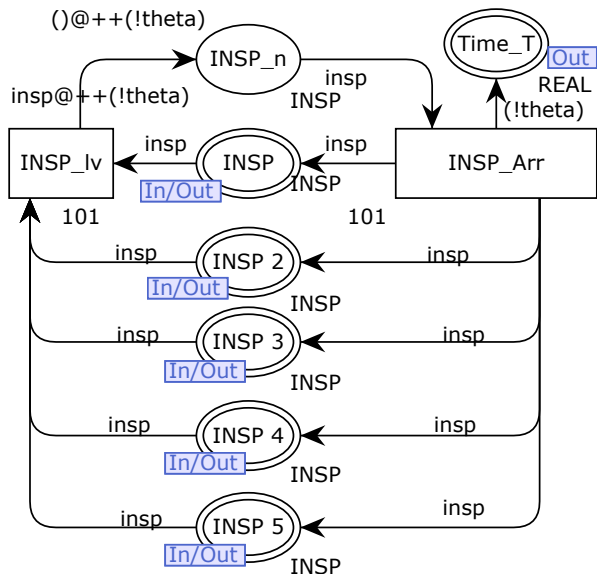


Figure 12: Inspection CPN model for the chosen section

A simulation monitor is set at place “INSP\_Arr” to collect the number of inspection  $N_{insp}$  for each simulation.

### 3.4.3 System maintenance decision CPN model

Each component model has its own maintenance requirements, system maintenance decisions need to satisfy all the component maintenance requirements. There are three situations for the system maintenance planning: pure corrective maintenance, pure preventive maintenance and mixed maintenance. They are arranged by the combination of tokens in places “CM\_Need”, “PM\_Need” and “PM\_time” in Figure 13.

Figure 13 illustrates the system maintenance planning CPN model. It shows the details of the substitution transition ‘Maintenance\_Plan\_System’ which makes the maintenance decisions for the whole railway section. Transition “set\_Plan” fires when a pure preventive maintenance group is needed, the firing time of which is controlled by the place “PM\_time”. It sends tokens to the place “PM\_TC*i*” which is the preventive maintenance demand connected to the component maintenance modules. Transition “Set\_plan\_CM” fires when it is time for the pure corrective maintenance group. Similarly, it sends the corrective maintenance demands (the tokens) to the place “CM\_TC*i*”. Transition “Set\_Plan\_CM&PM” fires when a mixed group of maintenance actions are implemented. Since this is a mixed maintenance group, the transition sends the maintenance requirements to all the places “PM\_TC*i*”s and “CM\_TC*i*”s.

We set the monitors for three transitions (“set\_Plan”, “Set\_plan\_CM” and “Set\_Plan\_CM&PM” ) separately to collect the number of setups (i.e.  $N_{psu}(t)$ ,  $N_{csu}(t)$  and  $N_{mix}(t)$  ) in Equation 13.

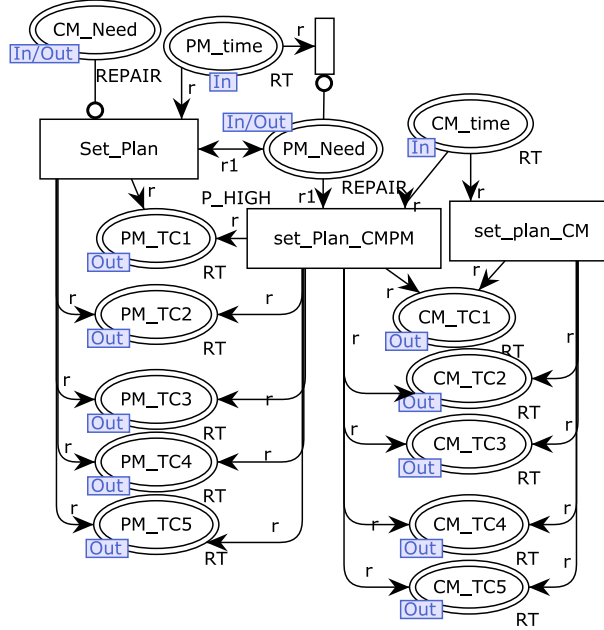


Figure 13: CPN module for Maintenance Plan System

## 4 Numerical results and analysis

In order to have a tuning of speed restriction  $v_{tsr}$  and preventive maintenance delay  $t_d$ , the developed model is populated using the parameters in Table 2 (the values in this table are illustrative for the simulation to demonstrate our CPN model) and Monte Carlo simulations are carried out to evaluate the model.

Table 2: Parameters used in simulations: unit of time: days, unit of cost: euros, unit of speed: km/h

	Value		Value
$\delta_{II}$	0.9 mm	$\delta_{DL}$	1.2 mm
$\theta_{insp}$	15 days	$t_D$	14 days
$v_0$	100 km/h	$\lambda_0$	12
$c_{pmin}$	1,000 euros	$\alpha_0$	0.000167
$a_0$	0.04	$\beta$	20.96
$t_r$	3 days	$t_R$	6 days
$A$	-183.3	$B$	12000 euros
$x_{pm}$	0.2mm		

Firstly, convergence is determined from the simulations, Figure 14 and 15 show the convergence for the long term system benefit  $\mathbb{E}B_\infty$  and system unavailability  $\mathbb{E}Q_\infty$ , the configurations are  $\delta_{II} = 0.9$ ,  $t_d = 14$  and  $v_{tsr} = 10$ .

The length of simulation should be determined according to simulation convergence to find the optimal system performance on an infinite time span.

System benefit  $B$  converges at the length of 500 years as shown in Figure 14; unavailability  $Q$  converges after 700 years as shown in Figure 15 with the speed  $v_{tsr} = 10$  and repair delays  $t_d = 14$  and  $t_D = 14$ .

The following simulations are set to run with the length of 10,000 years to estimate the maintenance cost and the unavailability.

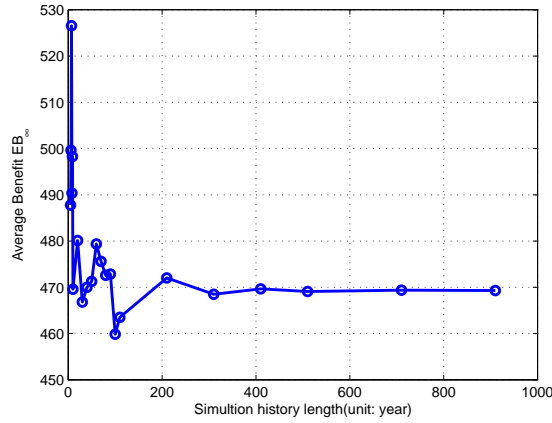


Figure 14: Convergence of the estimated system benefit with the simulation length

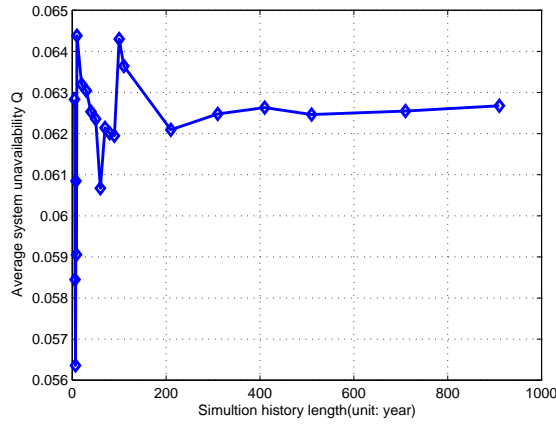


Figure 15: Convergence of the estimated system unavailability with the simulation length

#### 4.1 Results for the grouping strategy

Considering the simulation convergence, Monte Carlo simulations are carried out with the length of simulation to estimate the performance evaluation.

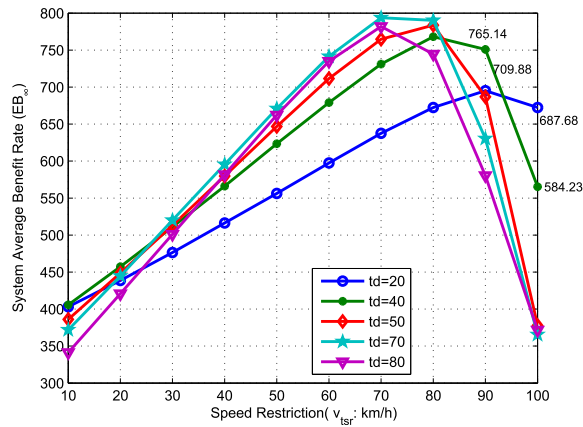


Figure 16: System Benefit  $EB_{\infty}$  against  $v_{tsr}$  for grouping maintenance strategy

System average benefit is one the metrics for maintenance performance evaluation. Figure 16 shows the

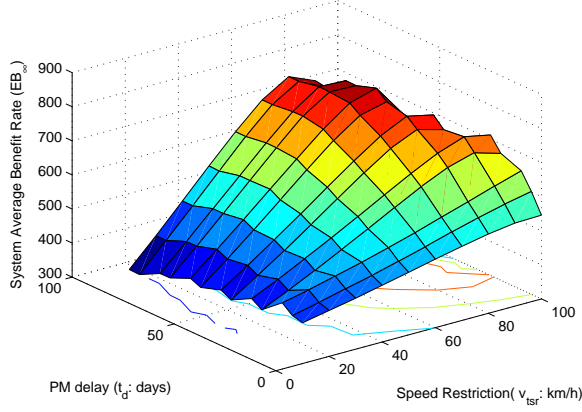


Figure 17: System benefit  $\mathbb{E}B_\infty$  as a function of  $v_{tsr}$  and  $t_d$  for grouping maintenance strategy

system benefit  $\mathbb{E}B_\infty$  as a function of speed restriction  $v_{tsr}$ , it illustrates 5 curves for the given repair delays and their maximum benefits explicitly. For example, when  $t_d = 20$ , the maximum system benefit is 709.88 if the speed restriction is  $v_{tsr} = 90$ . The relationship of system benefit and repair delays can also be seen in Figure 16, for the speed restriction  $v_{tsr} = 90$ , the system benefit is 765.14 if the repair delay is  $t_d = 40$ , but for the speed restriction  $v_{tsr} = 100$ , the system benefit is 687 when  $t_d = 20$ , it is higher than the benefit 584 when  $t_d = 40$ ; the results are inverse. They show that for a given speed restriction, a longer preventive maintenance delays  $t_d$  will not lead to a higher system benefit, since the preventive maintenance delays may lead to more corrective maintenance which is much more expensive than the preventive maintenance.

In order to show the relationship between system benefit, repair delay and speed restriction, Figure 17 shows the system benefit as a function of speed restriction  $v_{tsr}$  and repair delays  $t_d$ , where speed restriction is  $v_{str} \in [10, 100]$  and the range of repair delays is  $t_d \in [0, 85]$ . The convex surface shows that the system benefits like a ladder against the PM delays.

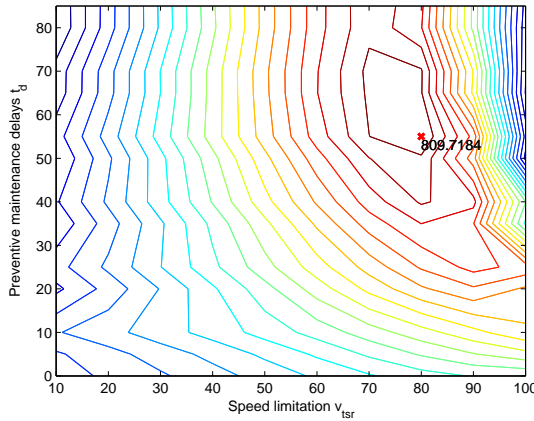


Figure 18: Contour plot of system benefit  $\mathbb{E}B_\infty$  as a function of  $v_{tsr}$  and  $t_d$  for grouping maintenance strategy

Figure 18 plots the contour of system benefit  $\mathbb{E}B_\infty$  as a function of  $t_d$  and  $v_{tsr}$ . It shows the optimized system benefit ( $\mathbb{E}B = 808.7$ ) explicitly when the speed restriction  $v_{tsr} = 80$  and the repair delay  $t_d = 55$ .

System maintenance cost as a function of speeds and repair delays is shown in Figure 19. A optimal maintenance cost can obtained if the speed restriction is 10 and the repair delay is 85. For the given preventive maintenance delay, the maintenance cost increases along with the speed, since the higher speed may cause more failures and defects.

For a given speed restriction (for example  $v_{tsr} = 100$ ), system maintenance cost against the floor of ratio

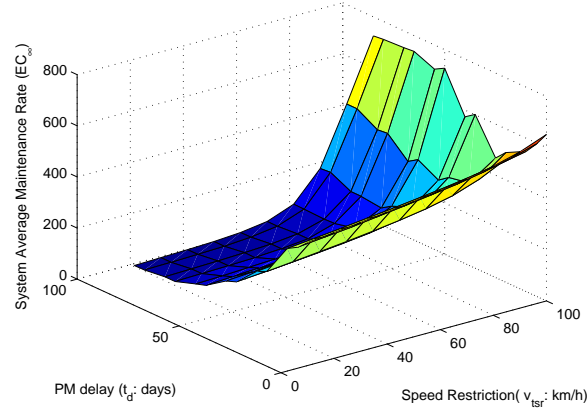


Figure 19: Maintenance cost  $\mathbb{E}C_\infty$  against  $v_{tsr}$  and  $t_d$  for grouping maintenance strategy

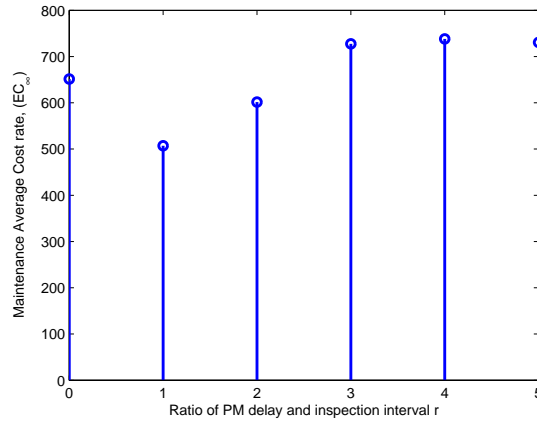


Figure 20: Plot of  $\mathbb{E}C_\infty$  against  $r = \lfloor \frac{t_d}{\theta_{insp}} \rfloor$  given that  $v_{tsr} = 100$  for grouping maintenance strategy

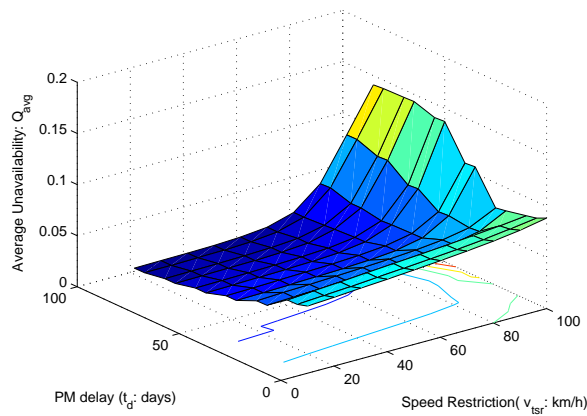


Figure 21: System unavailability  $\mathbb{E}Q_{avg}$  as a function of  $t_d$  and  $v_{tsr}$  for grouping maintenance strategy

of repair delay and inspection interval (i.e.  $r = \lfloor \frac{t_d}{\theta_{insp}} \rfloor$ <sup>1</sup>) can be seen in Figure 20. An optimal maintenance cost for a given speed restriction  $v_{tsr} = 100$  is obtained when the ratio is 1. It means an optimal repair delay can be scheduled for a given speed restriction. Thus, the optimal repair delays for  $v_{tsr} = 10$ , but the

<sup>1</sup>Notation  $\lfloor a \rfloor$  represents the largest integer not greater than  $a$

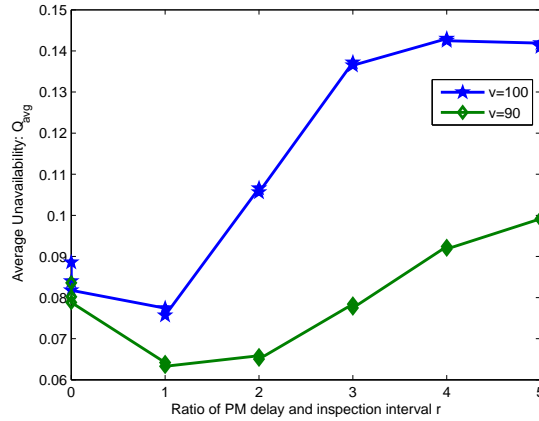


Figure 22:  $\mathbb{E}Q_{avg}$  given that  $v_{tsr} = 100$  and  $v_{tsr} = 90$  for grouping maintenance strategy

track is assumed to be repaired within a period not longer than 3 months when the defect is identified, so the results are shown in Figure 19 when  $t_d \in [0, 85]$  days.

Figure 21 shows the average system unavailability  $\mathbb{E}Q_{avg}$  on an infinite time span, as a function of the speed restriction  $v_{str}$  and repair delay  $t_d$ . The range of the speed restriction is  $v_{tsr} \in [10, 100]$  and the range of repair delay is  $t_d \in [0, 85]$ .

Figure 22 shows system unavailability as a function of the ratio  $r$  of PM delay and inspection interval, for  $v_{tsr} = 90$  and  $v_{tsr} = 100$ . Some points are overlapped at each  $r$ . The minimum unavailability for each curve can be obtained when  $r = 1$ , which is similar to the results obtained for the system maintenance cost.

Figure 23 is the plot of the optimal speed restriction for PM delays to optimize system benefit. When the PM delay is 10 days, the optimal speed restriction is 100 km/h; if the PM delay is 85 days, the optimal speed restriction is 70 km/h.

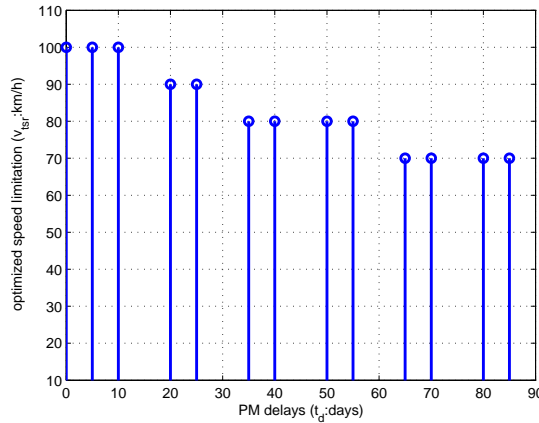


Figure 23: Optimal  $v_{tsr}$  for PM delay  $t_d$  for grouping maintenance strategy

Figure 24 illustrates the Pareto front for the grouping maintenance strategy, which is the solution for multiple objectives evaluation. It shows if  $v_{tsr} = 80$  and  $t_d = 55$ , we can have the highest  $\mathbb{E}B$  with a reasonable  $\mathbb{E}Q$ .

## 4.2 Results for the non-grouping strategy

Simulations for the maintenance strategies without grouping are carried out. Similar to the grouping maintenance strategy, the maximum system benefits are explicitly shown. Figure 25 plots the system benefit as a function of preventive maintenance delay  $t_d$  and speed restriction  $v_{tsr}$ . The maximum system benefit is

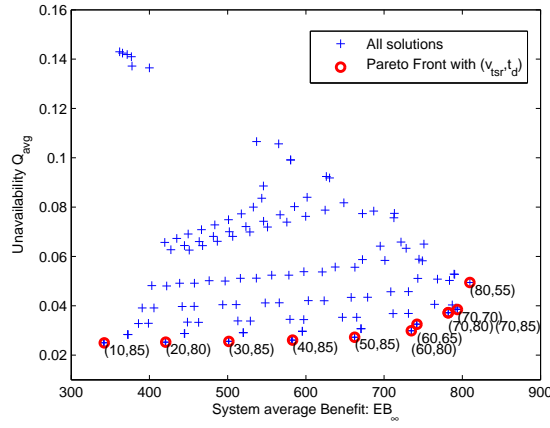


Figure 24: Pareto front for the grouping maintenance strategy solution with the parameter  $(v_{tsr}, t_d)$

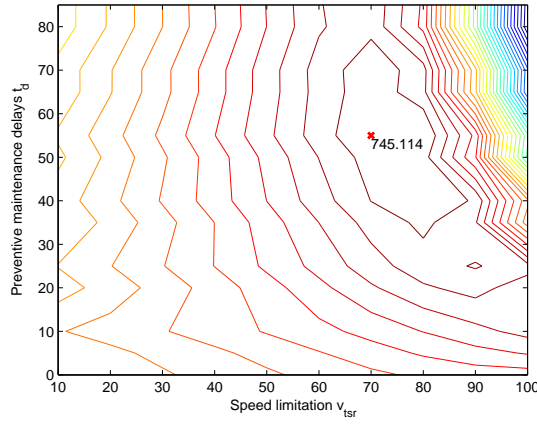


Figure 25: Plot of  $\mathbb{E}B_\infty$  as a function of speed restriction  $v_{tsr}$  and preventive maintenance delay  $t_d$  for non-grouping maintenance strategy

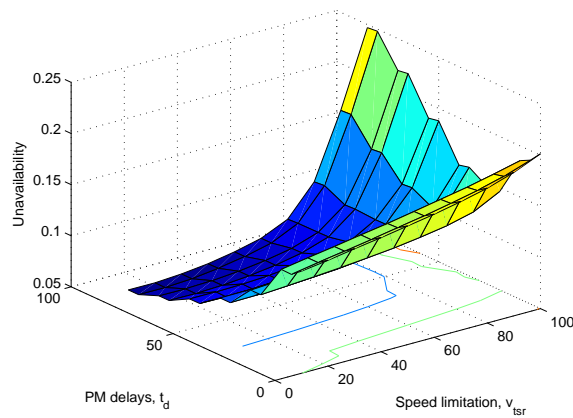


Figure 26: Plot of system unavailability as a function of speed restriction  $v_{tsr}$  and preventive maintenance delay  $t_d$  for non-grouping maintenance strategy

shown explicitly: when  $v_{tsr} = 70$  and  $t_d = 55$ , the maximum system benefit is  $B = 745.114$ .

Figure 26 illustrates the plot of system unavailability as a function of speed restriction  $v_{tsr}$  and preventive maintenance delay  $t_d$ . It shows that the system unavailability increases against speed restriction, and for a



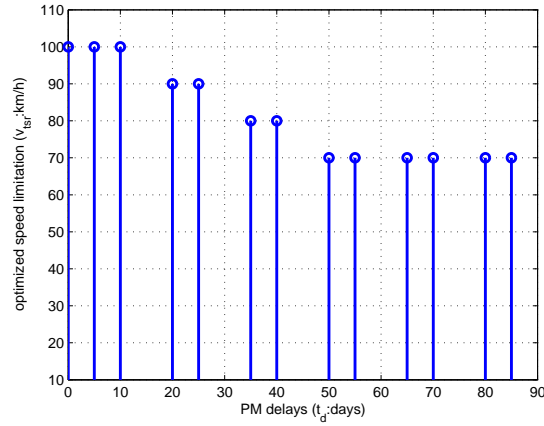


Figure 27: Optimal speed restriction  $v_{tsr}$  for PM delays under non-grouping maintenance strategy

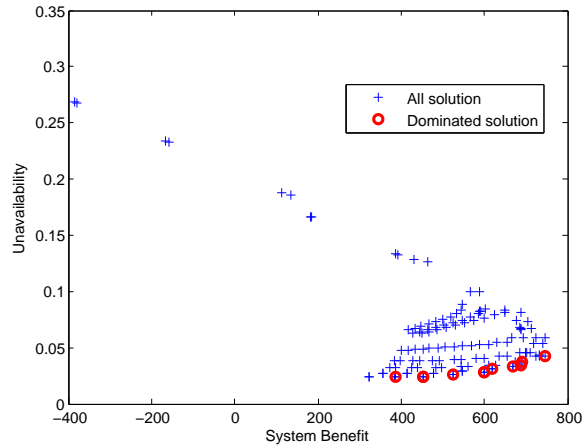


Figure 28: Multi-objective solutions and Pareto front for non-grouping maintenance strategy

given speed restriction  $v_{tsr}$ , an optimal system unavailability can be obtained.

Figure 27 shows the optimal speed for different PM delays if the non-grouping strategies is adopted. If the repair delays are short, such as  $t_d = 10$ , the section does not need a speed restriction imposed; if the repair delays is longer than 50 days, it is better to carry out speed restriction  $v_{tsr} = 70$  to optimize system benefit.

All solutions for the multi-objective evaluation and Pareto fronts for the non-grouping maintenance strategy are shown in Figure 28.

Table 3: The results comparison.  $G$ :system gain;  $C$ : maintenance cost;  $B$  system benefit;  $Q$ : system unavailability

		$v_{tsr}=80, t_d=55$	$v_{tsr}=70, t_d=55$
' $t_d$ ' grouping	<b>G</b>	1028.26	955.00
	<b>C</b>	218.54	167.89
	<b>B</b>	<b>809.72</b>	<b>787.11</b>
	<b>Q</b>	0.049	0.037
no grouping	<b>G</b>	996.85	917.41
	<b>C</b>	251.43	172.3
	<b>B</b>	<b>744.87</b>	<b>745.11</b>
	<b>Q</b>	0.058	0.042

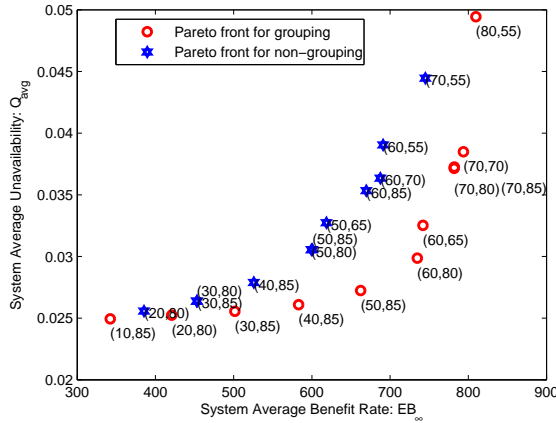


Figure 29: Comparison of the Pareto fronts of both strategies with the parameter  $(v_{tsr}, t_d)$

### 4.3 Comparison of both maintenance strategies

Table 3 shows the comparison of single objective performance for the grouping and non-grouping strategy. Besides system benefit and system unavailability, we also list system gains and maintenance cost in this table. In Table 3, shows the comparison of single objective performance for the grouping and non-grouping strategy. Besides system benefit and system unavailability, we also list system gains and maintenance cost in this table.

Figure 29 shows the comparison of the Pareto fronts of the grouping solution and the no grouping solution. The grouping strategy leads to both higher benefit and lower unavailability for the multiple objective optimization.

## 5 Conclusion and perspective

Speed is one of factors which affects the track deterioration and also impacts on railway operation. Therefore, imposing a speed restriction reduces the rate of track deterioration so that it extends the component survival time, but it reduces system throughput and hence system service performance (gain). An optimal speed configuration is needed to have a balance between the maintenance cost and the system gain. In addition, the longer survival time allows planning for a longer repair delay. During the delay time we can make better repair plans with a lower repair price. In addition, the delays enable maintenance activities to be grouped to save maintenance set-up costs.

A maintenance optimization problem is discussed in this paper which will determine an optimal tuning of the speed restriction and the repair delay to maximize the system benefit and minimize the system unavailability. A 5 component section is modelled in this paper, the deterioration process of these components depends on speed of the passing trains. In order to focus on the effect of speed on track deterioration, the component deterioration factors are assumed to be speed and the number of trains, and a Compound Poisson process with Gamma distributed jump size is used to described the random evolution increments. CPN Tools is adopted to model this 5 component section. The CPN model describes the deterioration as a stochastic process using coloured sets and coloured set functions. Coloured set function “ $Degrade(state, v, \lambda, t)$ ” describes the varied deterioration behaviours probability depending on the parameters  $v$  and  $\lambda$ , thus different track components have varied deterioration evolution with the same factors. A two level CPN model describes both the system behaviour and the maintenance process. In this model, speed restriction  $v_{tsr}$  and preventive maintenance delay  $t_d$  are the maintenance decision variables; system benefit  $EB_\infty$  and system unavailability  $EQ_{avg}$  are used to evaluate the system performance as a single objective evaluation and a multiple objective evaluation.

Monte Carlo simulations are carried out to find out the optimal  $v_{tsr}$  and  $t_d$ . For the single objective evaluation, the results explicitly show the optimized system benefit and the corresponding optimal speed

restriction and optimal repair delays. The plots of optimal speed against repair delays show that the slower speed restriction is needed for the longer repair delays to maximize the system benefit. The simulation results also show that the grouping strategy can lead to a higher system benefit, lower maintenance cost and lower unavailability than a strategy without grouping policies. The grouping strategies are shown to effectively reduce the set-up costs. Considering both the benefits and the system unavailability at the same time, can be accomplished by deriving the Pareto optimum from the solutions. The parts of the Pareto Fronts for both maintenance strategies indicate that grouping can lead to more efficient maintenance solutions.

In the future, the work will be extended to account for the dependencies in the operational speed in the system operation rules. In addition, more complex system structures will be considered, for example a greater number of sections and the inclusion of switches to change the direction of the train onto other lines.

## References

- [1] Guler H. Decision Support System for Railway Track Maintenance and Renewal Management. *J Comput Civ Eng* 2012; 27(3): 292–306.
- [2] Podofillini L, Zio E and Vatn Jr. Risk-informed optimisation of railway tracks inspection and maintenance procedures. *Reliab Eng Syst Saf* 2006; 91(1): 20–35.
- [3] Meier-Hirmer C and Poulighy P. Impact of preventive grinding on maintenance costs and determination of an optimal grinding cycle. In *Safety, Reliab. Risk Anal. Theory, Methods Appl. - Proc. Jt. ESREL SRA-Europe Conf., 22-25 September, 2008*, volume 4. Valencia; Spain, pp. 3183–3189.
- [4] Antoni M. The ageing of signalling equipment and the impact on maintenance strategy. In *Comput. Ind. Eng. 2009. CIE 2009. Int. Conf.*, volume 2. IEEE, pp. 1179–1184.
- [5] Quiroga LM, Wegele S and Schnieder E. Benefit of railway infrastructure diagnosis systems on its availability. In *2nd IFAC Work. dependable Control Discret. Syst. 2009*. pp. 146–150.
- [6] Quiroga LM and Schnieder E. A heuristic approach to railway track maintenance scheduling. *WIT Trans Built Environ* 2010; 114: 687–699.
- [7] Shafiee M, Patriksson M and Chukova S. An optimal age-usage maintenance strategy containing a failure penalty for application to railway tracks. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 2014; 0(0): 1–11. DOI:10.1177/0954409714543337.
- [8] Zio E, Marella M and Podofillini L. Importance measures-based prioritization for improving the performance of multi-state systems: application to the railway industry. *Reliab Eng Syst Saf* 2007; 92(10): 1303–1314.
- [9] Tinga T. Application of physical failure models to enable usage and load based maintenance. *Reliab Eng Syst Saf* 2010; 95(10): 1061–1075.
- [10] Yang Z, Djurdjanovic D and Ni J. Maintenance scheduling for a manufacturing system of machines with adjustable throughput. *IIE Trans* 2007; 39(12): 1111–1125.
- [11] Kish A and WClark D. Track Buckling Derailment Prevention Through risk based train speed reduction. In *AREMA 2009 Annu. Conf. Expo*. pp. 1–21.
- [12] Dirks B and Enblom R. Prediction model for wheel profile wear and rolling contact fatigue. *Wear* 2011; 271(1-2): 210–217.
- [13] Zwanenburg WJ. A model for the life expectancy of railway switches and crossings for maintenance and renewal planning in asset management systems. *WIT Transactions on The Built Environment* 2008; 103: 765–773.

- [14] Zwanenburg WJ. *Modelling Degradation Processes of Switches & Crossings for Maintenance & Renewal Planning on the Swiss Railway Network*. PhD Thesis, Delft University of Technology, 2009.
- [15] Sadeghi J and Askarinejad H. Development of improved railway track degradation models. *Struct Infrastruct Eng* 2010; 6(6): 675–688.
- [16] Sadeghi J and Askarinejad H. Development of track condition assessment model based on visual inspection. *Struct Infrastruct Eng Maint , Manag , Life-Cycle Des Perform* 2011; 7(12): 895–905.
- [17] Westgeest FP, Dekker R and Fischer RH. Predicting rail geometry deterioration by regression models. In *ESREL 2012-Advances Safety, Reliab. Risk Manag. June 25-29, 2012*. Taylor & Francis Group, pp. 926–933.
- [18] Andrade AR and Teixeira PF. A Bayesian model to assess rail track geometry degradation through its life-cycle. *Res Transp Econ* 2012; 36(1): 1–8.
- [19] Andrews J. A modelling approach to railway track asset management. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 2013; 227(1): 56–73.
- [20] Andrews J, Prescott D and De Rozières F. A stochastic model for railway track asset management. *Reliab Eng Syst Saf* 2014; 130: 76–84.
- [21] Prescott D and Andrews J. A railway track ballast maintenance and inspection model for multiple track sections. In *11th Int. Probabilistic Saf. Assess. Manag. Conf. Annu. Eur. Saf. Reliab. Conf. 2012, PSAM11 ESREL 2012; 25 June*, volume 1. pp. 41–50.
- [22] Caetano LF and Teixeira PF. Optimisation model to schedule railway track renewal operations: a life-cycle cost approach. *Struct Infrastruct Eng* 2014; : 1–13.
- [23] Patra AP, Kumar U and Larsson-Kråk PO. Assessment and improvement of railway track safety. In *Proc. - 9th Int. Heavy Haul Conf. "Heavy Haul Innov. Dev. IHHA 2009; Shanghai; China; 22-25 June 2009*. pp. 91–98.
- [24] Audley M and Andrews J. The effects of tamping on railway track geometry degradation. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 2013; 227(4): 376–391.
- [25] Grall A, Bérenguer C and Dieulle L. A condition-based maintenance policy for stochastically deteriorating systems. *Reliab Eng Syst Saf* 2002; 76(2): 167–180.
- [26] Meier-Hirmer C, Riboulet G, Sourget F et al. Maintenance optimization for a system with a gamma deterioration process and intervention delay: application to track maintenance. *Proc Inst Mech Eng Part O J Risk Reliab* 2009; 223(3): 189–198. DOI:10.1243/1748006XJRR234.
- [27] Caramia M and Dell’Olmo P. *Multi-objective management in freight logistics increasing capacity, Service level and Safety with optimization Algorithm*. Springer London Ltd, 2008.
- [28] Ngatchou P, Zarei A and El-Sharkawi A. Pareto Multi Objective Optimization. In *Proc. 13th Int. Conf. on, Intell. Syst. Appl. to Power Syst. 6-10 Nov. 2005*. IEEE, pp. 84–91.
- [29] Shang H, Bérenguer C and Andrews J. Delayed maintenance modeling with speed restriction for railway section. In *Saf. Reliab. Methodol. Appl. - Proc. Eur. Saf. Reliab. Conf. ESREL 2015*. London: Taylor & Francis, pp. 1029–1038.
- [30] Jensen K and Kristensen L. *Colored Petri Net–Modelling and validation of concurrent systems*. Springer-Verlag Berlin, 2009.