



HAL
open science

A Bayesian network to evaluate underground rails maintenance strategies in an automation context

Laurent Bouillaut, Olivier Francois, Stéphane Dubois

► **To cite this version:**

Laurent Bouillaut, Olivier Francois, Stéphane Dubois. A Bayesian network to evaluate underground rails maintenance strategies in an automation context. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2013, 227 (4), pp 411-424. 10.1177/1748006X13481306 . hal-00863566

HAL Id: hal-00863566

<https://hal.science/hal-00863566v1>

Submitted on 6 Oct 2023

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.

A Bayesian network to evaluate underground rails maintenance strategies in an automation context

Proc IMechE Part O:
J Risk and Reliability
227(4) 411–424
© IMechE 2013
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1748006X13481306
pio.sagepub.com


Laurent Bouillaut¹, Olivier Francois¹ and Stéphane Dubois²

Abstract

Reliability analysis has become an integral part of system design and operation. This is especially true for systems performing critical tasks, such as mass transportation systems. This explains the numerous advances in the field of reliability modeling. More recently, some studies involving the use of Bayesian networks have been proven relevant to represent complex systems and perform reliability studies. In previous works, a generic methodology was introduced for developing a decision support tool to evaluate complex systems maintenance strategies. This article deals with development of such a decision tool dedicated to the maintenance of Paris metro rails. Indeed, owing to fulfillment of high-performance levels of safety and availability (the latter being especially critical at peak hours), operators need to estimate, hour by hour their ability to prevent or to detect broken rails. To address this problem, a decision support tool was developed, the aim of this article is to evaluate, compare and optimize various operating and maintenance strategies.

Keywords

Maintenance, railway infrastructure, steel wheel metro's automation, availability, optimization, decision support, probabilistic graphical models

Date received: 29 May 2012; accepted: 13 February 2013

Introduction

Nowadays, the increasing of traffic and axle loads has led to a rail breaks growth. For safety reasons, restrictive exploitation rules are defined (a train can run on a broken rail but only after an enforcement process and with low speed) increasing delays and then, diminishing the service quality. Moreover, expensive corrective maintenance costs are needed to make up for this kind of failure. Therefore, efforts are being made for the application of reliability-based and risk-informed approaches to maintenance optimization of railway infrastructures. The underlying idea is to reduce the operation and maintenance expenditures while still assuring high safety standards.¹

Moreover, during a metro automation project, operators have to upgrade the control command of some of their lines. Indeed, some control command systems, such as track circuits, do not provide sufficiently precise localization to satisfy some exigencies, such as peak-hours intervals between trains. If this upgrade first addresses some operating problems, it has strong impacts also on rails maintenance since track circuits detect most of the broken rails, owing to their electric properties, and are therefore part of the rail diagnosis

strategy. To satisfy their high level of lines availability, operators need to evaluate the impact of all new systems in their ability to prevent and to detect broken rails.

Briefly, track circuit (TC) is a kind of electrical loop between a transmitter and a receiver using rails to transmit a “short circuit intensity” conveying information (such as speed limit) that can be “sniffed” by trains. Its normal function consists in detecting the presence of trains in its dedicated track section. Depending on if this block is free or not, the corresponding signaling indication will be permissive or restrictive (moreover the signaling permits a maximum of one train on a given TC). The TC analyses the rail impedance and so can detect some broken rails (BRs) when no trains are on the area.

¹University Paris-Est, Marne la Vallée, France

²RATP Régie Autonome des Transports Parisiens – Engineering Department, Fontenay-sous-Bois, France

Corresponding author:

Laurent Bouillaut, University Paris-Est, IFSTTAR, Grettia, Champs sur Marne F-77447, Marne la Vallée Cedex 2, France.

Email: laurent.bouillaut@ifsttar.fr

TCs finally detect almost 80% of broken rails and are actually the first contributor in BR detection.

Owing to its current automation projects, the RATP (major operator for public transportation in Paris and its surroundings) plans to upgrade the control command of some of its metro lines. The OCTYS (Open Control of Trains, Interchangeable & Integrated System) project considers switching from a track circuit train control to a communication-based train control (allowing higher speed and smaller intervals between trains that cannot be reached using TC). Then, RATP wants to evaluate the impact of control command modernization, correlated with new operating constraints. The development of a stochastic decision support tool was therefore decided to compare various scenarios (changing some operating, diagnosis and maintenance parameters).

To evaluate a given rail maintenance strategy, various indicators are needed, from annual numbers of broken rails (for the safety questions) and preventive maintenance actions, to delays before broken rails detection, including corresponding impacts on availability: the latter is expressed in the model through the related number of lost carousels (a broken rail inducing a stop of operations on the related part of the line, or for minor cases, a degraded situation with trains proceeding at reduced speed till the defect is consolidated). (The carousel (or loop) is the basic unit to express production of service for RATP. It corresponds to the round-trip journey of a given train from one extremity of the metro line.) The estimation of this last indicator requires a very high temporal precision (since it depends on the traffic that hourly changes), far beyond the degradation process dynamic. For many reasons (time computation, accuracy of parameters, learning data, etc.), the modeling of a rail degradation process with a one-hour step is impossible. The proposed decision support tool therefore had to consider evolutionary temporal granularity (changing in respect of both the rail degradation dynamic and the precision needed to estimate indicators).

To address these kinds of needs, the Diagnosis and Maintenance Group from IFSTTAR-GRETTIA has developed a generic methodology named *VirMaLab* (for Virtual Maintenance Laboratory). This approach aims to develop stochastic decision support tools for comparing, evaluating and optimizing maintenance strategies.²

Reliability analysis is an integral part of system design and operating. Moreover, it can be an input to optimize maintenance policies. Recently, dynamic Bayesian networks (DBNs) have been proved relevant to represent complex systems and perform reliability studies.^{3,4} This formalism was therefore chosen to develop our decision support tools. The major drawback of this approach comes from the constraint on the sojourn times that are necessarily exponentially distributed, as in usual Markovian approaches. To avoid this constraint, a new formalism named graphical duration models (GDMs) was introduced.⁵ This approach, based on semi-Markovian models,⁶ allows the representation

of all kind of sojourn time distributions. Then the degradation process of complex systems (multi-components, multi-states, eventually influenced by contextual variables) can be accurately modeled and thus, the related reliability indicators correctly estimated.

In this article, an extension of the commonly used *VirMaLab* formalism is introduced. To address the temporal granularity adjustment in respect of the state of the rail, a multi-models structure was developed. Usually, in *VirMaLab* applications, the whole model infers with a constant time step. Here, four models were introduced, with their own inference step fixed in accordance with the defect severity (from one month for early inner rail cracks, to one hour when broken rails occur) and their own set of diagnosis devices (all defects levels are not detected by the same means of appliances). Finally, the three first models emphasize the use of the preventive maintenance strategies on the availability of the network, whereas the last model focuses on the corrective maintenance and evaluates, hour by hour, the response of the diagnosis system in terms of broken rail detection ability.

Parameters of these models are learnt by use of REX databases and/or expert advices. Then, the global model is validated by various experiments with the standard running, diagnosis and maintenance parameters. Receiving the validation of these first results by RATP experts, new sets of scenarios can be computed, evaluating the influence of any parameters.

This article is divided in five sections. The next one introduces the context and the need of the considered study. Then, the *VirMaLab* generic approach is introduced and briefly focuses on the GDMs formalism. The proposed decision support tool, dedicated to the broken rail prevention in a context of renewal of the command-control systems of RATP steel-wheel lines, is then introduced. Finally, some conclusions and perspectives are discussed.

Context and needs of the study

The needs of public transportation (in terms of capacity) increases constantly in the Paris metro (the current traffic is around 1.5 billion trips per year). The RATP decided, therefore, to modernize the train command control for some of its most congested steel-wheel lines. The final aim of these automation projects is to be able to reach train intervals close to 90 s (roughly 15% traffic increasing). For that purpose the company has to deploy a continuous train speed control instead of a sector-based train positioning.

To be able to take into account the impact of various factors, the RATP engages a research project that consists in the implementation of decision support software that models the entire diagnosis and maintenance actors together with the rail degradation depending on some operating rules.

Introduction of the system

To precisely identify the impact of rail flaws on safety and availability of the railway system, the assessment of the current broken rail monitoring process regarding the new CBTC (communication based train control) system has to be made. Depending on the nature of defects, the disturbance will be more or less penalizing for the passenger service. A statistical model of the rail defects evolution should help the identification of the most critical flaws. Figure 1 introduces the consequence of inner crack spreading, inducing a broken rail.

To simplify the analysis, the rail states along the main deterioration process are clustered into four classes:

- OK (the rail has no defect);
- X1 (internal cracks, larger than 2 mm);
- X2 (internal cracks and some surface cracks, less than 30 mm);
- BR (broken rail).

Currently, the diagnosis of rail defects results from the combination of diagnoses from the four actors (detecting devices or specific staffs) involved in the broken rail monitoring process: these actors are characterized by different inspection periodicities and different detection efficiencies according to the type of defect and its location in the rail.

- A specific ultrasound vehicle (USV) dedicated to preventive maintenance actions is equipped with ultrasonic sensors. It diagnoses the rail on average twice a year. It is the only device able to detect the four considered classes of defects.
- Some walking survey teams (WT) are passing along the lines on average once or twice a month. They only can detect BRs and, in our study, have therefore a corrective function.
- During the passenger service, metro drivers (Drv) sometimes feel some shocks from the rails, and so, can contribute through their reporting to the detection of some BR.
- The TC normal function is the detection of trains on a given rail section. Depending on if this block is free or not, the corresponding signaling indication will be permissive or restrictive (moreover the signaling permits a maximum of one train on a given TC). As it analyses the rail impedance, track circuits can detect some BRs when no trains are on the area. TCs finally detect almost 80% of broken rails and are actually the first contributor in BR detection.

As briefly explained in the introduction, in its metro automation programs, RATP has to renew the command control system. With new implemented systems, the role of track circuits has changed in terms of signaling: whereas in past systems the TC ensure the basic detection of all trains, with new systems TC are useful



Figure 1. A broken rail owing to the spreading of an inner crack.

only for the detection of non-equipped or failed trains. But with lines modernized with the new CBTC system, the occurrence of a broken rail is a more critical event than in the past because of the reduced interval between trains and the fact that several trains could be present on the same TC section. The prevention of BR and a minimum number of false alarms for BR are key points to fulfill availability requirements. Our study aims to furnish a decision support tool allowing (across various indicators) to evaluate and compare some simulation runs and maintenance strategies.

A great amount of information, distributed among databases (from signaling and track departments) and expert advices, are available and will be used to learn probabilistic parameters of the proposed model (introduced in ‘Application of the *VirMaLab* approach to rail maintenance Optimization’). But, this information is sometimes uncertain, imprecise or even missing. For all of these reasons, the formalism of the Bayesian network theory, introduced in ‘The considered formalism: Bayesian networks’, offers an adequate framework to represent our system and its maintenance.

Hypothesis and needs of the study

A metro line is constituted of hundreds of elementary rail sections (between 5 and 18 meters long), with various ages, various states, etc. For complexity reasons, the development of a degradation model of a complete line is, therefore, unrealistic. To tackle this issue, a metro line (or portion of line) is assumed to be the sequence of a set of independent elementary rail sections (as illustrated in Figure 2), 18 m long. The proposed model focuses on one of these elementary sections and then, results are extrapolated to larger line sections to obtain reliability indicators on the considered portion of line.

To facilitate the use of this model for both maintenance operators and managers of the automation lines

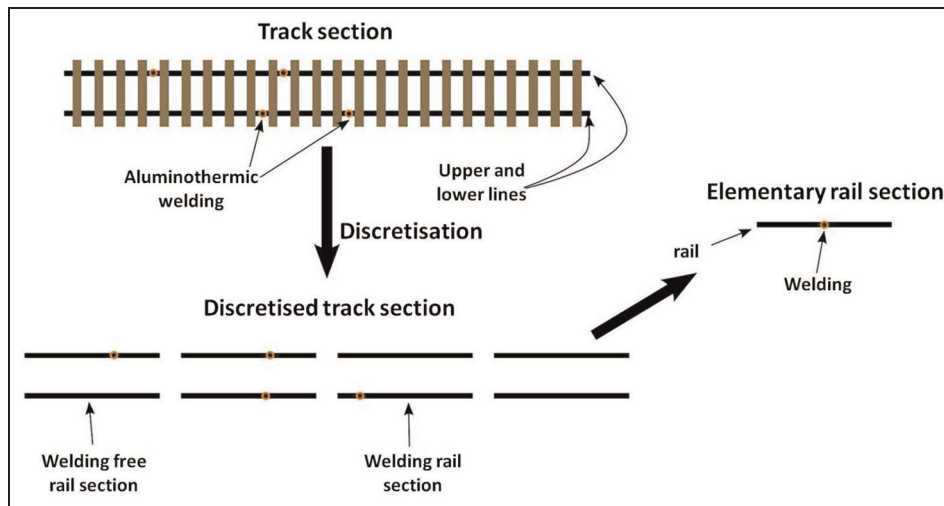


Figure 2. Discretization hypothesis for modeling a track section using an elementary rail section.

project, RATP asked for a user-friendly interface. It allows determining the following parameters.

- The considered line (among the 11 RATP steel-wheel underground lines).
- The rail context: the whole line or only the curved rails (possibly only the upper rail).
- The critical curve radius. It determines the set of curves on which a BR could have particularly critical consequences in terms of passengers' safety.
- The rail quality. For different reasons, an operator can decide to change the steel stiffness. Consequently, the rail degradation process must reflect the stiffness of the corresponding steel.
- Operations and rolling stock specifications: service periods (including the different peak and off-peak hours' ones: during one day, six periods are defined to reflect the transportation offer that varies according to time) and their associated headways, commercial speed (used as mean speed for the trains running in scenarios), length and axle load. These parameters influence the rail degradation speed and are also necessary to evaluate some final indicators.
- Diagnosis parameters: good detection and false alarms rates, USV and WT auscultation periods, parameters of the TC technologies encountered on the considered underground line.

When all parameters are defined, the inference can begin. Owing to the modeling complexity, the computation of an experiment can be quite long (around 2 h). Moreover, to ensure good results, a very high number of temporal steps are considered. For each simulation, the behavior of the elementary rail portion is analyzed over 50,000 years (indicators for the complete line are therefore obtained by meanings over 100 y). A solution based on an exact inference algorithm⁷ was studied, but the currently used structure of our network does

not allow the implementation of such algorithms. Nevertheless, the accuracy of the obtained results was proved during the study by computing the same experiment many times. The same results were systematically obtained.

After discussions and assessments with RATP experts, the final indicators must provide information about the following.

- Broken rails: number of annual BR on the line, mean time (in hours) before their detection and the contribution of each diagnosis device in these detections. Number of annual 'dangerous' BR (occurring in a critical 'in curve' context).
- Maintenance actions: number of annual preventive and corrective maintenance actions, ratio of actions triggered by false alarms and their distribution among all diagnosis devices.
- Loss of production: when a BR occurs, the operations' rules induce speed reductions or stopping of service in the corresponding area. Then, the operations' performance defined by the train headway cannot be maintained and the number of running trains strongly decreases. RATP names this loss of trains running 'monthly lost carousels'.

The VirMaLab approach

Introduction of the generic approach

During a previous study, a model named *VirMaLab* (for virtual maintenance laboratory) was proposed, this model being able to settle the optimal diagnosis parameters and the most adapted maintenance policy for a predetermined context and according to running constraints.

Figure 3 introduces this generic approach used for building such a decision support tool in order to determine optimal maintenance strategies.

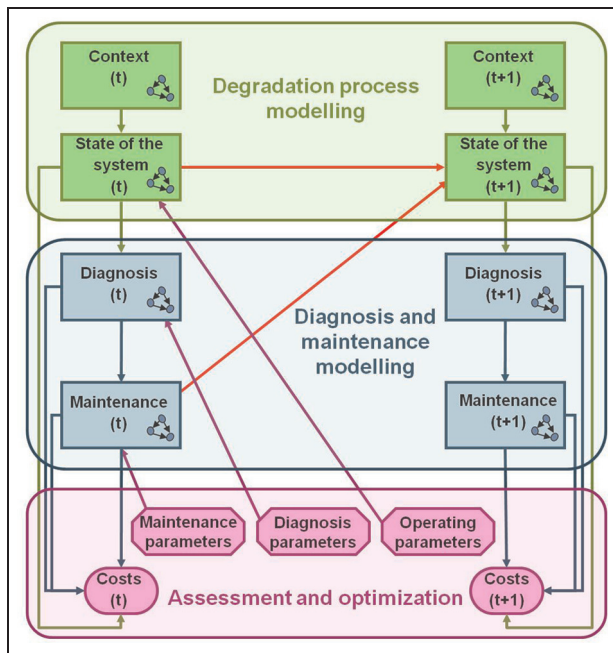


Figure 3. Generic approach for the *VirMaLab* maintenance decision support tool.

This approach is divided into three steps, described in the following sub-sections:

- modeling the degradation process;
- modeling the diagnosis and maintenance strategy;
- optimizing maintenance parameters.

Then, with such a decision support tool, one can evaluate various maintenance strategies and determine, for given cost functions, the best set of maintenance and diagnostic parameters. It can be applied to simple systems but also to multi-states and multi-component systems (with eventually interacting components). The learning of such modeling can be done with both expert advices and REX (Return of experience) databases.

Modeling the degradation process. The first step of *VirMaLab* consists in the mathematical modeling of the physical state of the system and the associated time evolution of its various components. This means being able to determine whether the system is still fault free after a given running time, knowing its initial state and the running parameters. On the contrary, if a defect appears, it is therefore necessary to determine, among a predefined list of possible defects, in which damaged state the system is.

Many studies already dealt with this topic. Cox models,⁸ or more generally proportional hazard regression models,⁹ are interesting tools for analyzing influence of contextual variables on degradation processes. Stochastic processes (such as Gamma or Poisson processes) are also frequently used in the reliability

field¹⁰ and find interesting application for railway maintenance modeling.¹¹ Finally, one of the most commonly used approaches is based on the Markov chains formalism, frequently modeled by Bayesian networks.¹² This approach is perfectly adapted for modeling transitions of a multi-states (and multi-components) system. Nevertheless, this approach needs a Markovian hypothesis. This means that sojourn times in each state must necessarily be exponential. If some systems confirm this assumption, most industrial applications underline non-Markovian behaviors (in the railway field, Weibull distribution are commonly used). In this case, a Markovian degradation process modeling can introduce non-negligible biases. Then, the estimated maintenance parameters can be far from optimal results.¹³

According to our bibliographical study, it seems that none of these approaches allows taking into account both the influence of contextual variables and multi-components systems having each its own degradation process (with all kinds of sojourn time's distributions).

To overcome this drawback, some semi-Markovian approaches can be considered.⁶ In *VirMaLab*, most of the previously introduced approaches can be implemented to model the degradation process in respect of the system's properties and needs of the considered study. Moreover, an original semi-Markovian modeling of degradation processes is also introduced, able to fit all kinds of sojourn time distributions: the GDMs,⁵ introduced in 'Graphical duration models'.

Modeling the diagnosis and maintenance strategy. The second step consists in the modeling of diagnostic devices (detection rate, false alarms rate, etc.) and their setting parameters (periodical auscultations, etc.). According to the results of the diagnosis' measurement campaign, each reference frame recommends the use of a maintenance action adapted to the current estimated state of the system. When a maintenance action is realized, the state of the system and its degradation process has to be updated to take into account the corrective action impact.

Optimizing maintenance parameters. The final aim of a *VirMaLab* decision support tool is providing evaluations and comparisons of various maintenance policies in respect of decision makers' needs (availability, safety, costs, etc.). Then, the last step of the modeling consists of integrating all necessary costs (running, unexpected stops, maintenance, diagnosis, etc.) characterizing the system.

For a given cost function, multi-objectives optimization algorithms can generally determine the optimal diagnosis and maintenance parameters. Some works on this subject are in progress, using genetic algorithms¹³ with specific crossover operators (BLX α or multi-points) and mutation operators (Gaussian or bit-flip), chosen in respect of the potential solutions coding (real or binary).

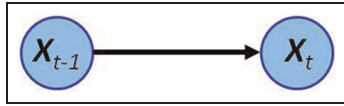


Figure 4. Dynamic Bayesian network modeling a Markov chain.

The considered formalism: Bayesian networks

Probabilistic graphical models, or Bayesian networks (BN), are mathematical tools relying on both the probability theory and the graph theory. They allow to qualitatively and quantitatively represent uncertainty. Basically, BNs are used to compactly describe the joint distribution of a collection of random variables $X = (X_1, \dots, X_N)$ taking their values in $\mathcal{X} = \{\mathcal{X}_1, \dots, \mathcal{X}_M\}$.

Formally, a BN denoted by \mathcal{M} is defined as a pair $(\mathcal{G}, \{p_n\}_{1 \leq n \leq N})$, where $\mathcal{G} = (\xi, \varepsilon)$ is a directed acyclic graph (DAG) and $\{p_n\}_{1 \leq n \leq N}$ a set of conditional probability distributions (CPD) associated with the random variables. These distributions aim to quantify the local stochastic behavior of each variable. The graph nodes and the associated random variables being both represented by $\xi = \{X_1, \dots, X_N\}$. ε is the set of edges encoding the conditional independence relationships among these variables. Finally, \mathcal{G} is said to be the qualitative description of the BN.

Besides, both the qualitative (i.e. \mathcal{G}) and quantitative (i.e. $\{p_n\}$) parts can be automatically learnt, if some complete or incomplete data or experts opinions are available.¹⁵

Using a BN is also particularly interesting because of the easiness for knowledge propagation through the network. Indeed, various inference algorithms allow the computation of the marginal distribution of any sub-set of variables. The most classical one relies on the use of a junction tree.⁷

Finally, note that such modeling is able to represent dynamic systems (e.g. which contain variables with time-dependent distributions) via the DBN solution.¹⁶ A DBN is a convenient extension of the BN formalism to represent discrete sequential systems. Indeed, DBNs are dedicated to model data that are sequentially generated by some complex mechanisms (time-series data, bio-sequences, number of mechanical solicitations before failure, etc.). It is, therefore, frequently used to model Markov chains. Figure 4 illustrates this property, introducing a DBN modeling the Markov Chain of the sequence $X = (X_1, \dots, X_N)$ taking its values in the set \mathcal{X} . This DBN is described by the probabilities that quantify the transitions from one state of \mathcal{X} to another. More precisely, a DBN defines the probability distribution of a collection of random variables $(X_t)_{t \geq 1} = (X_{1,t}, \dots, X_{t,t}^D)_{t \geq 1}$ where t is the discrete time index.

Graphical duration models

The graphical duration model (GDM) is a specific DBN, using semi-Markov models. The main idea is the

introduction of remaining time variables into the graph that allows the modeling of multi-state systems featuring complex sojourn times. Figure 5 shows a GDM in its DBN form. The left side of the network gives the GDM initial model, while the second one depicts its transition model.

The solid lines define the basic structure; dashed lines indicate optional items and thick edges characterize dependencies between time slices. The model handles two kinds of variable.

- $(X_t)_{1 \leq t \leq T}$ represents the system state over a sequence of length T , whose cardinality is the number of states of the system.
- $(X_t^D)_{1 \leq t \leq T}$ represents the remaining time before a system state modification (remaining sojourn time). These variables are called duration variables. Their cardinality is determined by the user in respect of a chosen memory depth.

Optionally, it is possible to introduce a context description of the studied system by means of a prior graphical model \mathcal{M}_{Z_t} . It aims to define the distribution of a possible collection of context variables (covariates) $Z_t = (Z_{p,t})_{1 \leq p \leq P}$ (one variable at least) that works on variable state X_t and/or duration variable X_t^D .

The DAG of a GDM shows that the current system state X_t depends on the previous system state X_{t-1} , the previous remaining duration X_{t-1}^D and, optionally, on contextual variables $Z_{n,t}$. On the other hand, the current duration variable X_t^D is dependent on the previous duration variable X_{t-1}^D , the current state X_t and, optionally, on the previous state X_{t-1} and some contextual variables $Z_{n,t}$.

Consequently, the process (X_t) (respectively, (X_t^D)) is not Markovian since $X_{t-1} \not\perp\!\!\!\perp X_{t+1} | X_t$ (respectively, $X_{t-1}^D \not\perp\!\!\!\perp X_{t+1}^D | X_t^D$). Where the notation $A \perp\!\!\!\perp B$ means that variables A and B are statistically independent.

On the other hand, the GDMs graphical structure leads to

$$(X_{t-1}, X_{t-1}^D) \perp\!\!\!\perp (X_{t+1}, X_{t+1}^D) | (X_t, X_t^D) \quad (1)$$

So, the set (X_t, X_t^D) engendered by a GDM is Markovian, despite (X_t) being not. The GDM generalizes the recent studies on discrete semi-Markovian processes.¹⁷

On the practical point of view, this approach allows specifying arbitrary state sojourn time distributions in contrast with a classic Markovian frame work in which all durations have to be exponentially distributed. This modeling is, therefore, particularly interesting as soon as the question is to capture the behavior of a given system subjected to a particular context and a complex degradation distribution. More details on this GDM (quantitative description, optional context description, etc.) can be found in Donat et al.⁵

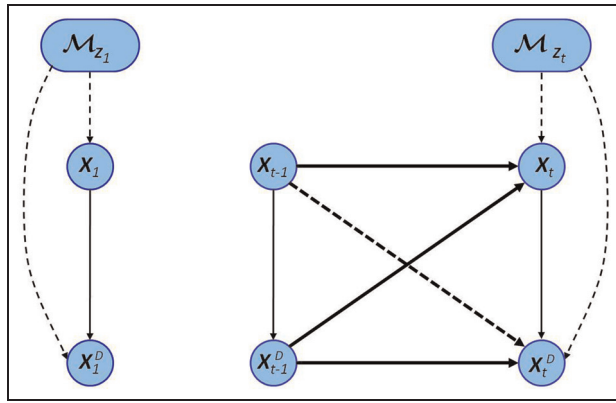


Figure 5. Graphical duration model in the form of a dynamic Bayesian network.

Generic structure of the Bayesian network for VirMaLab decision support tools

Based on details introduced in the previous sections, the generic structure of *VirMaLab* decision support tools can be proposed, presented by Figure 6. The three steps described in Figure 1 can be clearly found.

First, the degradation process of each component, eventually influenced by exogenous variables (sub-networks M_Z). It can be modeled either by a “classical” approach (stochastic processes, Markov chains, etc.) if some assumptions on the degradation’s nature afford it or using GDM (Graphical Duration Model) (if we do not have any idea on the degradation’s exact nature or if its process is quite complex). Nevertheless, for the second approach, some REX data are necessary. Note that

the possible interaction of some components on other degradation processes is not represented on this picture but can be integrated to the *VirMaLab* modeling.

Then, the maintenance strategy is described with both nodes D^i representing all diagnosis devices or staffs (specific to one component or shared by some of them) and nodes A describing maintenance actions. Note that diagnosis devices can be characterized by single variables (as in Figure 6) or eventually by more complex sub-networks precisely describing the whole diagnosis procedure. All kind of maintenance actions can be integrated to the modeling: correctives and condition-based actions (triggered in respect of results provided by diagnosis devices) and systematic actions (scheduled using the node ρ as a switch whatever the results brought by the diagnosis).

Finally, various costs can be integrated: in respect of the current state of the components (C^{X_i}), costs involved by diagnosis procedures (C^{Diag_i}) or by maintenance actions (C^{Maint_i}). These C nodes are similar to utility nodes, introduced in the Bayesian networks theory to provide decision support tools named influence diagrams.¹⁵

Application of the VirMaLab approach to rail maintenance optimization

StatAvaries, a multi-models extension of the VirMaLab approach

The adopted modeling is based on the *VirMaLab* generic approach, introduced in ‘Introduction of the generic approach’. Nevertheless, the aims of this study do

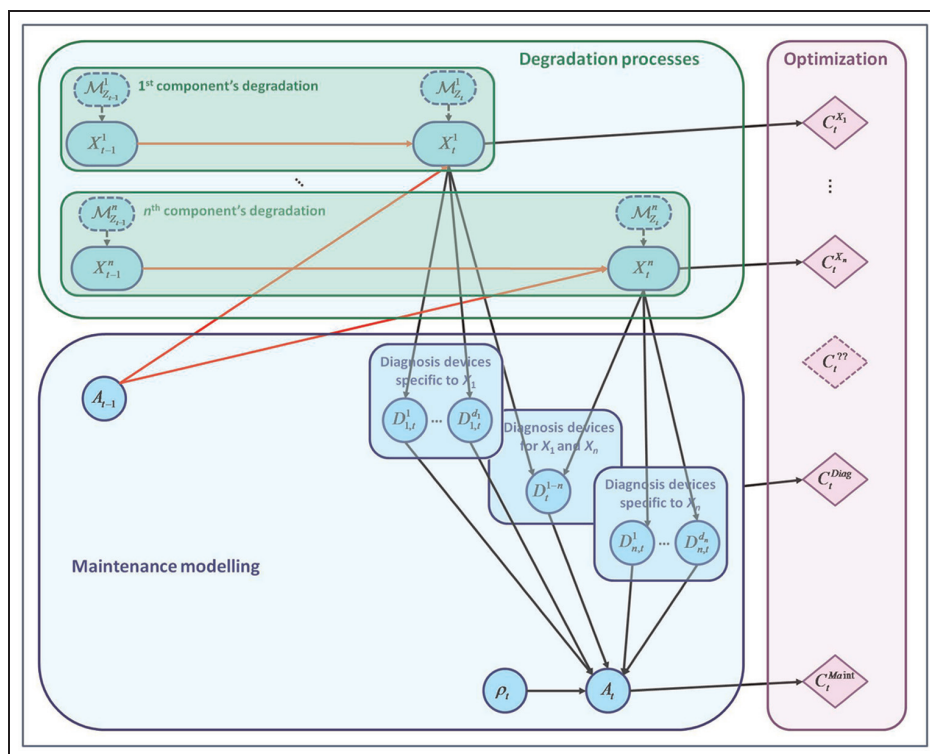


Figure 6. Generic structure of DBN developed as *VirMaLab* tools for maintenance optimization.

not include economic parameters (only the indication about the impact on production in terms of carousel is highlighted). So, only the first two blocks (*Degradation process* and *Maintenance modeling*) will be considered.

Contrary to the preliminary study,¹⁸ the second specificity of this application consists in the necessity of being able to evaluate, hour by hour, the consequences of a BR (especially during peak hours) when the rail degradation speed unit should be between the week and the month. As previously explained, the development of a rail degradation process model with a one-hour step is therefore absolutely unsuitable (in terms of complexity, accuracy, etc.). To overcome this problem, an original multi-models approach (introduced in Figure 7) is proposed. Each state of the rail S_i , taking its values in $\{ok, X_1, X_2, BR\}$, is characterized by a *VirMaLab* Bayesian network with a sojourn time distribution S_i^D , learnt from REX databases. Owing to the complexity of the proposed model's graphical structure, computation could not use an exact inference algorithm. The approximate version of the junction tree algorithm¹⁵ was therefore considered.

The first two layers of the structure are characterized by one-month iterations. The third one, dedicated to X_2 defects, has a one-week step. Finally, the last network, evaluating the ability of the system to detect BR, has a one-hour step.

The three first sub-models (networks $OK-X_1$ to X_2-BR), detailed in '*VirMaLab* preventive maintenance network', deals with the rail's preventive maintenance strategy. They aim to evaluate and quantify the ability of the system to prevent broken rails. The last

sub-model, introduced in figure 9 focuses on BRs detection and corrective actions.

VirMaLab preventive maintenance network. This first model (introduced in Figure 8) deals with the rail's preventive maintenance strategy. As a *VirMaLab* modeling, it is constituted of two blocks.

The first one describes the degradation process of the rail, using the GDM formalism (introduced in 'Graphical duration models'). The rail degradation can be influenced by several contextual variables, such as the type of rolling stock (changing from one line to another), the curve radius (and we give the possibility of considering only the inner or upper rail) and the steel's stiffness. These contextual variables are summarized in the sub-network "*Context*".

The second block of this model describes the diagnosis actors (devices or staff) and the maintenance strategy. Three actors trigger periodic auscultations of the rails: the USV, WT and Drv, whom presence depends on the state of the traffic, with peak hours, operation closure at night, etc.). Nodes "*USV presence*", "*WT presence*" and "*Drv presence*" model this presence on a given portion of rail a time t .

The modeling of the last device, meaning the track circuits (TC), is a little more complex. Indeed, several track circuit technologies are implemented on each line of the RATP network; each technology having different failure rates (if a TC is down, it is unable to detect any BR), different mean length, etc. The availability of TC is modeled by node "*TC*", depending on the node

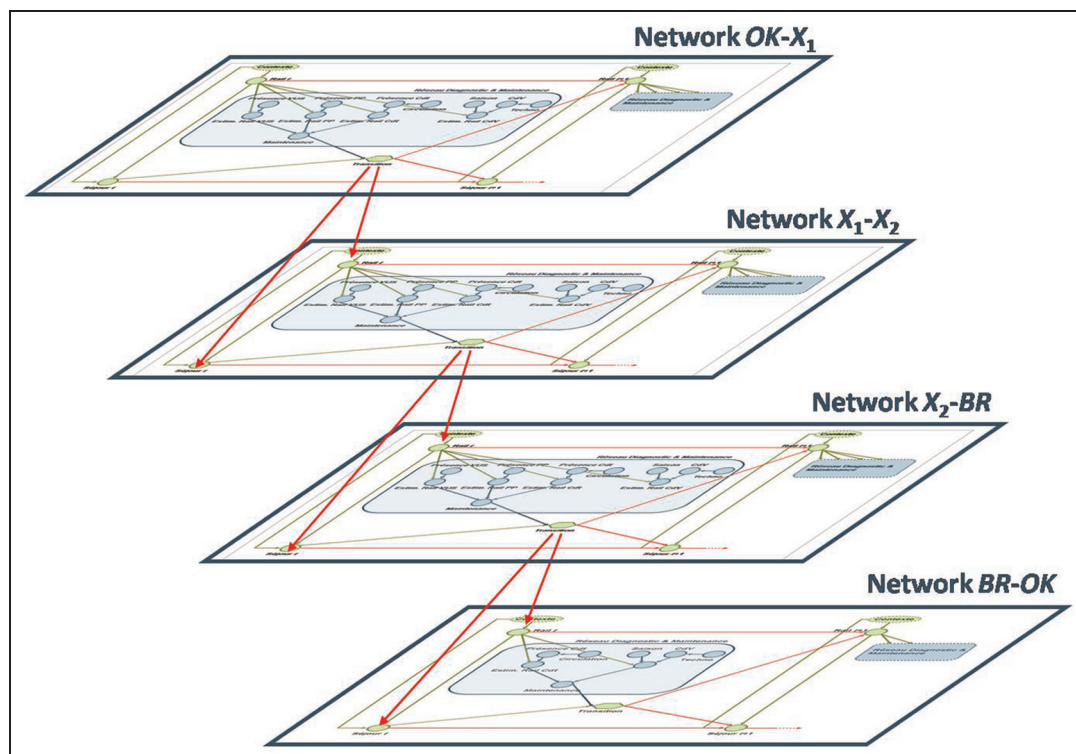


Figure 7. Multi-models structure of the *VirMaLab* decision support tool StatAvaries.

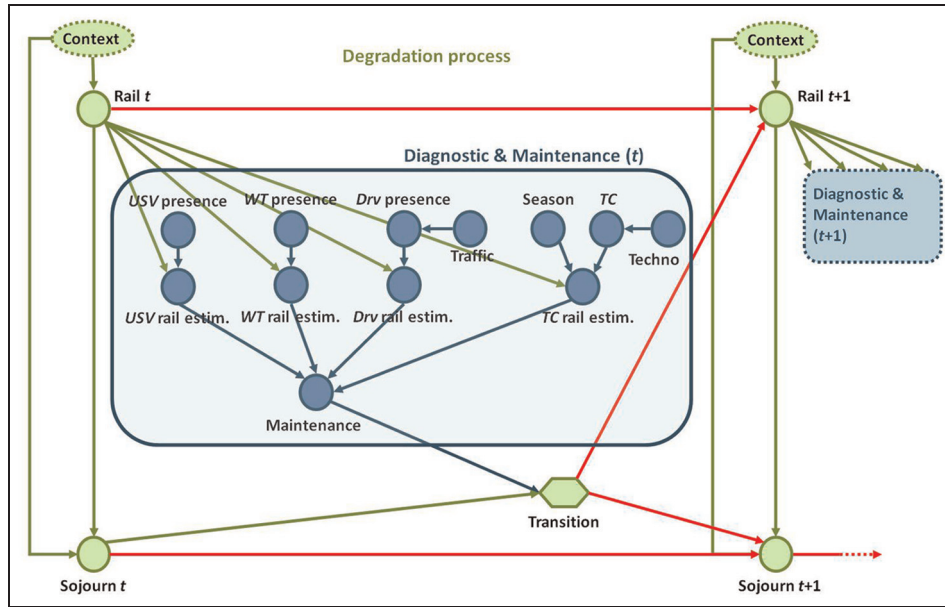


Figure 8. Structure of the *VirMaLab* model for the first three layers of the *StatAvaries* multi-models.

Table 1. Description of nodes from the *VirMaLab* model introduced in Figures 8 and 9.

Node	States	Description
USV presence	Yes No	Turns periodically to “Yes”, in respect of the USV auscultation period
USV rail estim	OK X ₁ X ₂ BR NM	Estimation of the state of the rail by the USV
WT presence	Yes No	Turns periodically to “Yes”, in respect of the WT auscultation period
WT rail estim	OK X ₂ BR NM	Estimation of the state of the rail by WT
Traffic	Timetable slot 1 ... Timetable slot 6	A day is divided into six timetable slots, describing peak hour, night...
Drv presence	Yes No	Turns periodically to “Yes”, in respect of the USV auscultation period
Drv rail estim	OK BR NM	Estimation of the state of the rail by drivers
Season	Warm Cold	
Techno	CVCM UM71 ITE Ruta NCO 50 SigTay ITE with joints	Six TC technologies used in the metro network
TC	ON HS	Describe is the TC is working or not
TC rail estim	OK BR	Estimation of the state of the rail by TC
Transition	Yes No	Activated when the sojourn time of the current states (node <i>Sojourn</i>) turns to 0

USV: ultrasound vehicle; WT: walking survey teams; BR: broken rail; NP: Not present; Drv: drivers; TC: track circuit.

“*Techno*”, listing all technologies encountered on the considered metro line. Moreover, the analysis of RATP databases underlines that, during warm seasons, the rail dilatation helps keeping the electric contact of many BR. In this case, the TC ability to detect BR registers a 50% decrease. This property was introduced by node “*Season*”.

Table 1 introduces a brief description of each node of both models introduced in Figures 8 and 9.

All four diagnosis actors provide an estimation of the current state of the rail, modeled by nodes “--- rail estim.”. Their CPT (Conditional Probabilities Table) contains the confusion matrix quantifying their good detection rates, false alarm rates and non-detection

rates for each state of the rail. These conditional probabilities were obtained by RATP expert advices.

As an illustration of this process, Table 2 introduces the conditional probability table of node “*USV rail estim.*”. State “*NP*” (for “not present”) describes the answer of node “*USV rail estim.*” when the ultrasonic vehicle is not present on the considered portion of rail (node “*USV presence*” = “*No*”). If an auscultation is running (“*USV presence*” = “*Yes*”), the USV confusion matrix is introduced then. Parameters τ_{VUS}^1 , τ_{VUS}^2 and τ_{VUS}^{BR} are the USV false alarm rates (respectively for defects X₁, X₂ and broken rails) while Θ_{VUS}^1 and Θ_{VUS}^2 are the USV non-detection rates (respectively for rail defects X₁ and X₂). All these rates were

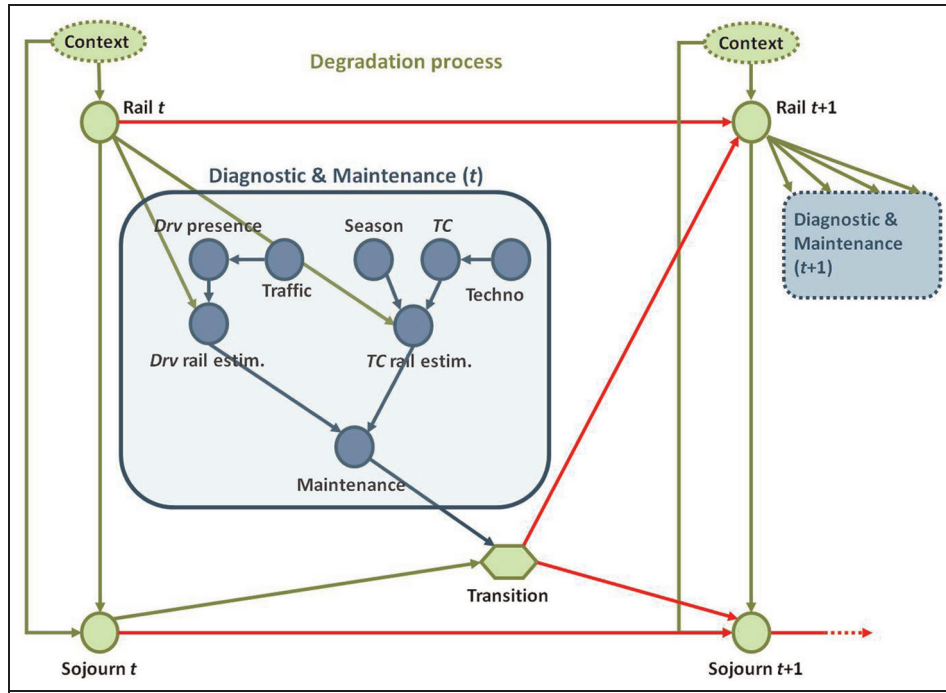


Figure 9. Structure of the model for the ‘one-hour step’ slice of the StatAvaries multi-models.

Table 2. CPT of diagnosis node “USV rail estim.”.

USV presence t	Rail t	USV rail estim. t				
		OK	X ₁	X ₂	BR	NP
Yes	OK	$1 - \tau^1_{VUS} - \tau^2_{VUS} - \tau^{BR}_{VUS}$	τ^1_{VUS}	τ^2_{VUS}	τ^{BR}_{VUS}	0
	X ₁	Θ^1_{VUS}	$1 - \Theta^1_{VUS}$	0	0	0
	X ₂	Θ^2_{VUS}	0	$1 - \Theta^2_{VUS}$	0	0
	BR	0	0	0	1	0
No	OK	0	0	0	0	1
	X ₁	0	0	0	0	1
	X ₂	0	0	0	0	1
	BR	0	0	0	0	1

BR: broken rail; USV: ultrasound vehicle; NP: not present.

estimated by RATP expert advices, sometimes reinforced by RATP REX databases.

Finally, diagnosis results influence the maintenance decision (node “Maintenance”). As only one maintenance action is considered (the replacement of the damaged portion of rail), when a maintenance action is performed, it is assumed that the system turns to the “ok” state in a single iteration.

VirMaLab corrective maintenance network. This second model, introduced in Figure 9, focuses on BRs detection and corrective actions.

Indeed, it evaluates both the ability of the system to detect a BR and the impact of such an event on the global indicators (time before detection, availability by the corresponding number of impacted revenue service trains through the assessment of the loss of production, etc.).

This network is activated when a BR occurs. To evaluate the time before the detection, only the actors with a short reaction time are considered. USV and WT diagnosis, whose checking periods are on a relatively long-term basis, are therefore not taken into account and drivers are compared with a real-time device. The actors with a short reaction time to BR are Drv and TC. Indeed, in service periods, Drv systematically diagnoses the rail (a train runs at least each 105 s–115 s and the model iteration is one hour long).

Experiments running. Each experiment begins in the ok state. Then, a sojourn time T_{ok} is generated in respect of ok^D . During this period, some indicators are computed (mainly false alarms). After T_{ok} iterations (i.e. T_{ok} months in terms of rail degradation) an inner crack appears and the second network is activated. A new duration T_{X1} is generated on respect of X_1^D . To allow

quick degradations modeling, this duration can be null. In such cases, the third network is immediately activated. On the other hand, during these T_{X1} months, the preventive maintenance strategy is evaluated analyzing false alarms, good detections, etc. If a maintenance action is performed, the system is reinitialized during the next iteration.

If no preventive maintenance action is done, when the X_1 sojourn time ended, the third network is activated with a new duration T_{X2} . The preventive maintenance policy is evaluated one more time. If the X_2 defect is not detected and/or if no preventive maintenance action is performed, the rail will break after T_{X2} weeks. Then, the last network is activated. No sojourn time is generated since the BR is a blocking state. Indeed, the only way to end this state is the detection of the defect and the replacement of the rail.

As an illustration, the next section will introduce results obtained in some of the dozens of scenarios investigated for RATP. The aim of this article is not to list all the results obtained during the study, but to introduce the *VirMaLab* multi-models extension, illustrated by some experimental examples. For more information on some of the obtained results, readers can of course contact the authors.

Illustration by some results

This section introduces some of the obtained results, detailed for Line 7 of the Paris metro network. The aim is not to detail all results of the study, but to give some illustrations of the proposed methodology. For confidentiality reasons, exact values of indicators are deleted. Nevertheless, the interest of these results lays in the evolution of indicators according to variations of main components of the model, and in the comparison between the reference scenario and the prospective ones.

It is worth first of all turning one's attention to the fact that, before running these experiments, a long validation process was achieved. This process aimed to validate the model parameters (all conditional

probabilities learnt from both expert advices and REX databases) by comparing the obtained results (number of broken rails, lost loops, maintenance actions, etc.) for a reference scenario with their known values. This reference scenario depicts the situation (running, diagnosis and maintenance) the RATP was currently facing.

Impact of preventive maintenance in various running contexts. In this study, one of the considered scenarios deals with the influence of the USV auscultation period on maintenance actions and the network's availability, in respect of the traffic frequency. Figure 10 focuses on the impact of the systematic preventive maintenance actions (triggered by USV inspections) on the annual number of broken rails.

For this experiment, the ultrasonic auscultation period was changed, with value in $[2T_o, T_o, T_o/2, T_o/3, T_o/6]$ (where T_o is the reference period, currently used on Line 7), and balanced by two various running settlements:

- the first one analyses the influence of the USV auscultation period with the current traffic frequency (peak hours around 114 s)
- the second one prospects what could be the influence of the USV auscultation period if the traffic was increased by 25% (the peak hours headway being at 90 s).

We can note that, as expected and for both traffic conditions, the more frequently ultrasonic equipment sound the infrastructure, the more preventive actions will be planned (blue curves). Early defects are therefore more easily diagnosed, and then corrected before they turn to the critical state of broken rail (red curves).

Moreover, the increase of trains' traffic accelerates the rail degradation process. Indeed, with the current auscultation period T_o , the annual number of broken rails increases by 50% with a higher traffic, inducing an unacceptable availability level of the infrastructure

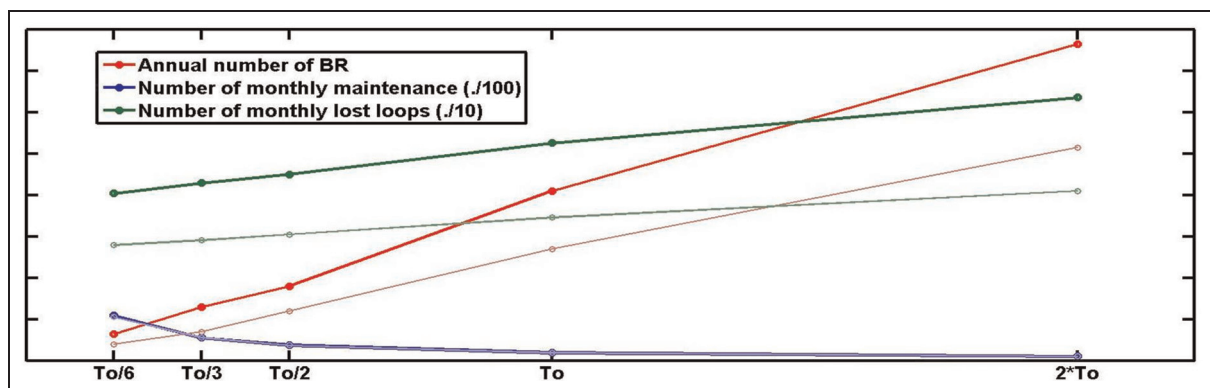


Figure 10. Influence of the USV period (in months) on rail's degradation – application to Line 7 in both actual running parameters (light curves) and increased traffic (bold curves).

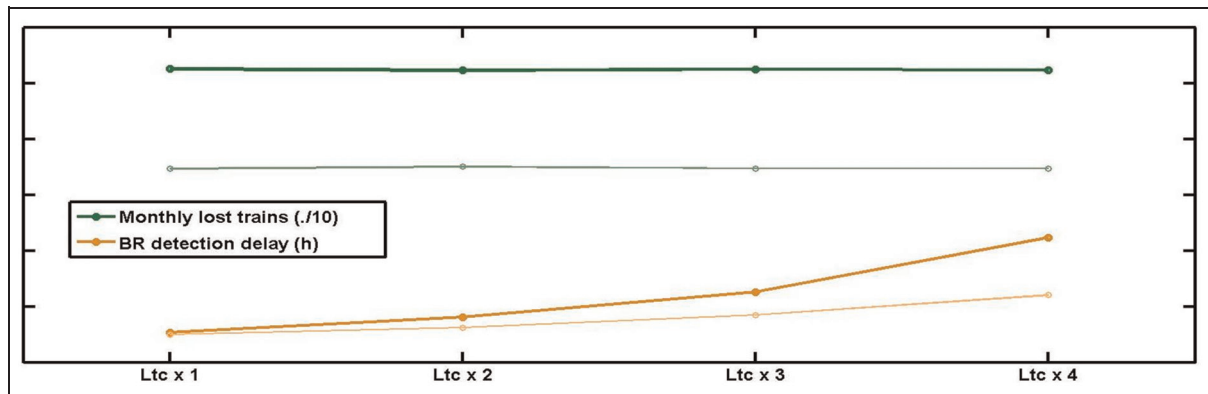


Figure 11. Influence of the TC mean length on broken rail's detection – application to Line 7 in both actual running parameters (light curves) and increased traffic (bold curves).

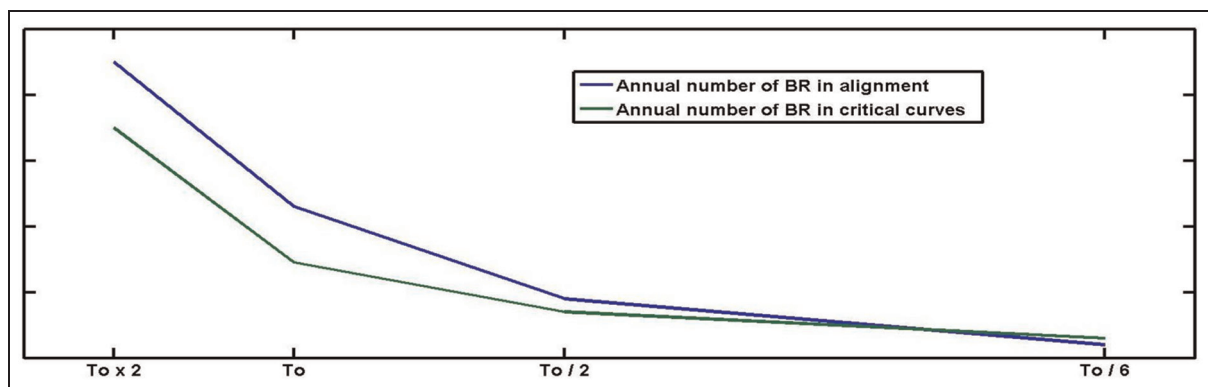


Figure 12. Influence of the USV period on rail's degradation for various curves contexts – application to Line 7.

(comparison of bold and light red curves). Blue and red curves (both bold and light lines) seem to testify that traffic only impacts the end for the rail degradation process (transitions from X_2 defects to BR) and not early defects. This assumption was confirmed by mean life times of each defect classes (X_1 , X_2 and BR), learned from RATP REX databases.

Then, if RATP wants the traffic to increase substantially on its renewed lines, as expected by its on-going programs for modernizing the command-control systems, the developed model shows that it is recommended to adapt accordingly the preventive maintenance strategy; for example by increasing the USV auscultation frequency.

Influence of track circuits mean length in various running contexts. Figure 11 shows the evolution of the estimated number of mean lost exploitation loops and the broken rails detection time in respect of the TC mean length. For this experiment, the mean length was changed, with value in $[L_{TC}, 2 L_{TC}, 3 L_{TC}, 4 L_{TC}]$ (where L_{TC} is the reference TC mean length on Line 7), and balanced by two various running settlements (used and described previously).

First, we can note that the numbers of broken rails and lost loops logically only depend on the traffic rate

and not on the TC length (comparison between bold and light curves).

The interesting information concerns the BR detection time, strongly impacted by the track circuit length as TC could not detect a broken rail while a train is over it (the longer they are, the smaller their detection ability will be).

Such a result provides some significant information to help decision makers taking care of the size of their track circuits in respect of the adopted trains' interval.

Impact of preventive maintenance on critical broken rails. Another interesting indicator deals with critical occurrences of broken rails, taking place in curves with small curvature radius (in that the probability for a derailment is higher). As introduced previously, mechanical constrains on the rail are higher in curves. So, this context triggers a quicker degradation and a then higher number of internal cracks (with finally a higher number of BR). As all the rolling stock's weight leans on the curves upper rails, most of in curves BR occurs on these upper rails and can therefore involve dangerous situations.

Figure 12 introduces the influence of preventive maintenance on the localization of BR and on their number. The improvement of preventive maintenance

shows a decrease of 'in risky curves' BR numbers with finally a nearly equivalent behavior of both contexts (alignment and risky curves).

Impact of the inhibition of some diagnosis systems on BR mean time detections. The modernization of command-control systems in RATP raises some questions related to the currently involved actors in the process permitting the detection of BR: drivers and TC. At the moment, no corresponding project is foreseen, but the implementation of a driverless command-control system on a steel-wheel line could occur in the future; another possible evolution for signaling in RATP could be the replacement of TC by other detection means, such as axle counters (less demanding in terms of maintenance than TC), as TC have a reduced role in newly deployed command-control systems. To help explore these prospective scenarios, the developed model permits some experiments that quantify the impact of removing drivers or track circuits from the process of detection of BR. In both cases, the BR detection mean times increase by 33.8% to 60.7%, according to the inhibited system and the considered traffic frequency...

Such results confirm that, if TC and drivers are no longer necessary to drive and control trains for some automation contexts, their ability to detect quickly broken rails makes them indispensable in terms of infrastructure's safety and availability: it means that their possible removal from the process permitting the detection of BR would need to be compensated by other dedicated means.

Conclusions

In this article an original maintenance strategy modeling was introduced, dedicated for the prevention and detection of broken rails, in a context of renewal of the command-control systems for Paris steel-wheel metro lines.

This modeling is based on a generic approach named *VirMaLab* using the DBN theory, with a modular approach. Thus, the proposed modeling can be divided into sub-networks, eventually interconnected, describing the rail degradation process, the different diagnosis actors (devices and staff) and finally, the maintenance actions decision.

The originality and innovation of this work is that, if the application introduced in this article deals with the railway infrastructure, the considered approach is generic and can easily be extended to all kind of maintenance processes modeling for determining maintenance and/or diagnosis optimal parameters.

Moreover, if *VirMaLab* considers various degradation models (stochastic processes, Markov chains, Cox models, etc.) a specific Bayesian network is proposed. This semi-Markovian approach, named GDM, ensures an accurate degradation process modeling, whatever sojourns times distributions in all system's states. If the

algorithm complexity is higher with GDM, they are a good solution when the exact nature of the degradation process is unknown or, if it is too complex, to be correctly modeled by other approaches.

Finally, the multi-models extension introduces a multiple temporal sampling, satisfying both the degradation dynamic and the accuracy required to quantify correctly the impact of broken rails and their related false alarms, in order to support RATP in its decisions.

As an illustration of this approach, some results are introduced, focusing on the influence of some diagnosis parameters (ultrasonic diagnosis periodicity, inhibition of some broken rails detectors, traffic frequency, etc.) on safety and availability indicators (annual numbers of broken rails and preventive maintenance actions, broken rails detection mean time, etc.). These experiments were realized for the Line 7 of the Paris metro network.

These results illustrate the ability of the approach to simulate all kinds of scenarios, modifying maintenance decisions, diagnosis parameters or running variables and to provide decision support to maintenance and operating decision makers.

One last advantage of the introduced method leads in the fact that all new information (from databases or expert advices) or modification of the diagnosis process can easily be taken into account to amend the modeling.

Finally, the integration of meta-heuristics in the inference algorithm was recently made, furnishing useful tools to determine, in respect of some predetermined criteria, the optimal diagnosis and/or maintenance parameters.

More details on another *VirMaLab* railway application, named *SpecifRail*, can be found in Bouillaut and Quost.¹⁹ It aims to optimize the compromise between replacements of defected rails and the complete line renewal (in respect of both economical constraints and network availability). A study, dedicated to the maintenance optimization of Bombardier rolling stock's users' access systems, is still in progress. It aims to develop a methodology defining dynamic maintenance strategies changing in respect of the system's behavior.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Podofilini L, Zio E and Vatn J. Risk-informed optimization of railway tracks inspection and maintenance procedures. *Rel Engng Sys Saf* 2005; 93(12): 20–35.
2. Bouillaut L, François O, Aknin P, et al. *VirMaLab* (virtual maintenance laboratory): a decision support tool for optimization of maintenance strategies. *Recherche, Transport, Sécurité* 2011; 27: 241–257.
3. Langseth H and Portinale L. Bayesian networks in reliability. *Rel Engng Sys Saf* 2007; 92(1): 92–108.

4. Celeux G, Corset F, Lannoy A, et al. Designing a Bayesian network for preventive maintenance from expert opinions in a rapid and reliable way. *Rel Engng Sys Saf* 2006; 91(7): 849–856.
5. Donat R, Leray P, Bouillaut L, et al. A dynamic Bayesian network to represent discrete duration models. *Neurocomputing* 2009; 73: 570–577.
6. Limnios L and Oprisan G. *Semi-Markov processes and reliability. Statistics for industry and technology*. Springer, 2001.
7. Lauritzen SL and Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. *J Royal Stat Soc* 1988; 50(2): 157–224.
8. Cox DR. Regression models and life-tables. *J Royal Stat Soc* 1972; 34(2): 187–220.
9. Kay R. Proportional hazard regression models and the analysis of censored survival data. *App Stats* 1977; 26(3): 227–237.
10. Van Noortwijk JM. A survey of the application of gamma processes in maintenance. *Rel Engng Sys Saf* 2009; 94(1): 2–21.
11. 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 IMechE, Part O: J Risk and Reliability* 2009; 223(3): 189–198.
12. Aven T and Jensen J. *Stochastic models in reliability. Stochastic modelling and applied probability*. Springer, 1999.
13. Bouillaut L, Donat R, Neji A, et al. Estimation of multi-components system's reliability: comparison of two graphical model approaches. In: *13th IFAC symposium on information control problems in manufacturing*, Moscow, Russia, 3–5 June 2009.
14. Jensen FV. *An introduction to Bayesian networks*. UCL Press, 1996.
15. Jordan M. *Learning in graphical models*. Cambridge, MA: MIT Press, 1999.
16. Murphy KP. *Dynamic Bayesian networks: representation, inference and learning*. PhD thesis, University of California, Berkeley, 2002.
17. Barbu V, Boussemart M and Limnios N. Discrete time semi-Markov processes for reliability and survival analysis. *Comm Stats – Theory and Methods* 2004; 33(11): 2833–2868.
18. Bouillaut L, François O, Leray P, et al. Dynamic Bayesian networks modeling maintenance strategies: prevention of broken rails. In: *8th world congress on railway research*, Seoul, Korea, 18–22 May 2008.
19. Bouillaut L and Quost X. 6th PCRd European project *UrbanTrack. Predictive and preventive maintenance of metro and commuter tracks: Maintenance optimisation*. Final report. 2009.