

Technical report 16-016

# **Probabilistic defect-based risk assessment approach for rail failures in railway infrastructure\***

A. Jamshidi, S. Faghieh Roohi, A. Núñez, R. Babuska,  
B. De Schutter, R. Dollevoet, and Z. Li

*If you want to cite this report, please use the following reference instead:*

A. Jamshidi, S. Faghieh Roohi, A. Núñez, R. Babuska, B. De Schutter, R. Dollevoet, and Z. Li, “Probabilistic defect-based risk assessment approach for rail failures in railway infrastructure,” *Proceedings of the 14th IFAC Symposium on Control in Transportation Systems (CTS 2016)*, Istanbul, Turkey, pp. 73–77, May 2016. doi:[10.1016/j.ifacol.2016.07.013](https://doi.org/10.1016/j.ifacol.2016.07.013)

Delft Center for Systems and Control  
Delft University of Technology  
Mekelweg 2, 2628 CD Delft  
The Netherlands  
phone: +31-15-278.24.73 (secretary)  
URL: <https://www.dcsc.tudelft.nl>

---

\*This report can also be downloaded via [https://pub.deschutter.info/abs/16\\_016.html](https://pub.deschutter.info/abs/16_016.html)

# Probabilistic Defect-Based Risk Assessment Approach for Rail Failures in Railway Infrastructure

Ali Jamshidi\* Shahrzad Faghih Roohi\*\* Alfredo Núñez\*  
Robert Babuska\*\* Bart De Schutter\*\* Rolf Dollevoet\*  
Zili Li\*

\* *Section of Railway Engineering, Delft University of Technology, Stevinweg 1,  
2628 CN Delft, The Netherlands*

(email: A.Jamshidi@tudelft.nl, A.A.NunezVicencio@tudelft.nl,  
R.P.B.J.Dollevoet@tudelft.nl, Z.Li@tudelft.nl)

\*\* *Delft Center for Systems and Control, Delft University of Technology, Mekelweg  
2, 2628 CD Delft, The Netherlands*

(email: S.FaghihRoohi@tudelft.nl, R.Babuska@tudelft.nl, B.DeSchutter@tudelft.nl)

---

**Abstract:** This paper develops a defect-based risk analysis methodology for estimating rail failure risk. The methodology relies on an evolution model addressing the severity level of rail surface defect, called squat. The risk of rail failure is assessed by analysing squat failure probability using a probabilistic analysis of the squat cracks. For this purpose, a Bayesian inference method is employed to capture a robust model of squat failure probability when the squat becomes severe. Moreover, an experimental correlation between squat visual length and squat crack depth is obtained in order to define four severity categories. Relying on the failure probability and the severity categories of the squats, risk of future failure is categorised in three different scenarios (optimistic, average and pessimistic). To show the practicality and efficiency of the proposed methodology, a real example is illustrated.

*Keywords:* Squat, Railway track, Bayesian inference, Failure risk, Severity analysis

---

## 1. INTRODUCTION

In the recent years, railways has been promoted in the whole world as a means of reducing road traffic congestion and emission levels. In order to keep the trains running without disruptions, an efficient maintenance policy based on risk assessment of the different components of the infrastructure is essential to anticipate problems before they occur.

Among all railway infrastructures, the track plays an important role in the entire railway system. In the Netherlands almost half of the maintenance budget is allocated to track maintenance (Zoeteman, 2014). The purpose of the budget is to keep the track at a high reliability level. Moreover, a robust track maintenance plan can facilitate infrastructure management by capturing a set of realistic cases of component degradation. Then, the infrastructure manager would be able to define which scenarios are the most relevant to consider and how to manage the track maintenance in a maintenance time horizon. As a high percentage of the railway system failures occur in the tracks, analysing the failure risk caused by surface defects is crucial for the track maintenance plan (Burstow et al., 2002; Zhao et al., 2006; Liu et al., 2001; Hassankiadeh, 2011). The idea of this paper is to analyse the effect of one common defect in railway networks called squat. To assess a defect-based risk, two major factors must be taken into account. First, the track stochastic variables such as the growth rate of defects where the rail structure deteriorates

as the traffic passes along the rails. Second, the spatial characteristics of the track since the track characteristics vary in space. The idea is to capture the evolution rate of the squat when the growth can affect the track reliability and where the track is prone to rail failure. Moreover, in extreme cases, the squat could pose a safety threat due to potential derailment (Prescott et al., 2013).

In this paper, risk of rail failure is assessed relying on a probabilistic approach using a Bayesian inference method. The Bayesian approach provides robust inferences together with a more realistic treatment of growth rate uncertainties. A few studies have been carried out on the application of Bayesian methods in safety of railway infrastructures. Andrade et al. (2015) employ Hierarchical Bayesian models to predict the evolution of the main quality indicators related to railway track geometry degradation including the standard deviation of longitudinal level defects and the standard deviation of horizontal alignment defects. The goal is to use the modelled indicators in planning of track maintenance operations. An investigation on railway ballast failures is done by Lam et al. (2014) using Bayesian inference to analyse uncertainty induced by measurement errors of vibrations in the ballast failure zones. Two integrated frameworks for track degradation and rail maintenance decisions are proposed relying on Bayesian networks in (Bouillaut et al. 2008; Mahboob, 2014). A nonparametric Bayesian approach with a Dirichlet Process Mixture Model is used to facilitate reliability analysis in a railway system by Mokhtarian et al. (2013). Train acci-



Fig. 1. Example of severe squats on a rail

dent consequences can be modelled by Bayesian networks where human errors and track degradation are addressed (Bearfield et al., 2005, Marsh, 2004; Castillo et al., 2015). This paper is organised as follows. In Section 2, a short background on the squats is presented. Section 3 addresses the Bayesian model of rail failure. Section 4 presents the risk assessment model together with a real-life example. Finally, in Section 5, conclusions are presented.

## 2. SQUAT IN RAILWAY INFRASTRUCTURES

Surface defects can affect track availability. Those rolling contact fatigue (RCF) defects can be classified as rail corrugation, squats, head checks, shatter cracking, vertical splits, head horizontal splits, and wheel burns (Magel, 2001). Appearance of those defects results in the increase of maintenance operations needed, more frequent track monitoring required, and rail failure when not detected in time in the worst case.

In this paper, we investigate squats, which are surface-initiated defects. The squats are observed in tracks, either ballast tracks or slab tracks, and in all possible traffic volumes (Kaewunruen et al., 2014). Fig. 1 shows a reference photo of severe squats with cracks already propagated beneath the rail surface.

Typically, the squats evolve from indentations into defects with surface cracks growing along the depth beneath the rail surface (Li et al., 2010). Once the squat gets severe in terms of crack depth and visual length, the train ride quality and safety become considerably low (Remennikov and Kaewunruen, 2008). In practice, squats can be detected and analysed using different methods, such as inspection using human inspectors, on-board measurements via photo/video records, axle box acceleration (ABA) measurements, and other non-destructive testing (NDT), such as ultrasonic and eddy current testing. While axle box acceleration (ABA) measurements are efficient in detecting both early stage and severe squats (Molodova et al. 2014; Li et al. 2015), in this paper the focus is the analysis of severe defects in terms of crack lengths. Thus, we rely on ultrasonic and surface photos of the defects.

Ultrasonic (US) testing is currently one of the most extensively employed automatic inspection technique for squats. This method can only be used to reliably detect cracks with depths higher than 4 mm, depending on the instru-

ments. When a rail includes squats with cracks larger than 4 mm, the evolution of the defects generates a potential risk of the rail failure. This paper employs US measurements to model crack growth of squat. In the next two sections, the rail failure probability model is presented.

## 3. BAYESIAN MODEL FOR RAIL FAILURE

Bayesian methods are widely used as a statistic technique to evaluate robustness in stochastic data behaviours in particular, for analysis of hazard rates with a small number of data samples. Potential benefits of the Bayesian approach in comparison with the usual Maximum Likelihood Estimate (MLE) method are computationally explained by Ahn et al. (2007). The MLE is an effective tool to estimate hazard rate as long as a sufficient amount of data is available. Using the MLE, a single point value for the failure rate, which maximises the likelihood function, can be estimated. However, our prior beliefs about the likely values for the failure rates are not injected into the estimation model with the MLE. In contrast to the MLE, Bayesian inference treats failure rates as random variables. Thus, the difference is that in the Bayesian model, the estimation output is a probability density function rather than a single point as in the MLE.

In Bayesian inference, prior knowledge and beliefs about unknown parameters are represented by the probability density distribution  $\pi_0(\lambda)$ , and statistical observations  $y$  have the likelihood  $f(y|\lambda)$  where  $\lambda$  is the failure rate. Then, according to Bayes' theorem, the posterior distribution of rail failure probability is expressed as:

$$\pi(\lambda|y) = \frac{f(y|\lambda)\pi_0(\lambda)}{f(y)} \propto f(y|\lambda)\pi_0(\lambda) \quad (1)$$

Let us assume that the failure probability is constructed by considering a nonlinear regression model over the crack depth. The data include observations of the crack depth, the number of cracks with the same depth, and the number of cracks with the growth above 4 mm (see Fig. 2). The nonlinear regression model shows the likelihood distribution of parameters  $a$  (intercept) and  $b$  (slope) in the Bayesian inference model:

$$f(y|(a, b)) = \exp(-1/(a + b \cdot y)) \quad (2)$$

where  $y$  is the crack depth. When no prior information is available about the values of parameters  $a$  and  $b$ , we assume uniform prior distributions (Faghih-Roohi et al., 2014):

$$\pi_0(a) = \text{Uniform}(A_1, A_2) \quad (3)$$

$$\pi_0(b) = \text{Uniform}(B_1, B_2) \quad (4)$$

By Bayes' theorem, the joint posterior distribution of the model parameters is proportional to the product of the likelihood and the priors. Monte Carlo methods are often used in Bayesian data analysis to describe the posterior distribution. The objective is to generate random samples from the posterior distribution and use them when it is not possible to compute analytically the posterior distribution. For this purpose, a slice sampling algorithm is chosen to obtain  $N$  samples of the distribution with an arbitrary density function (Neal, 2003). The slice sampling algorithm is a type of a Markov Chain Monte Carlo (MCMC) algorithm. Among all the MCMC methods

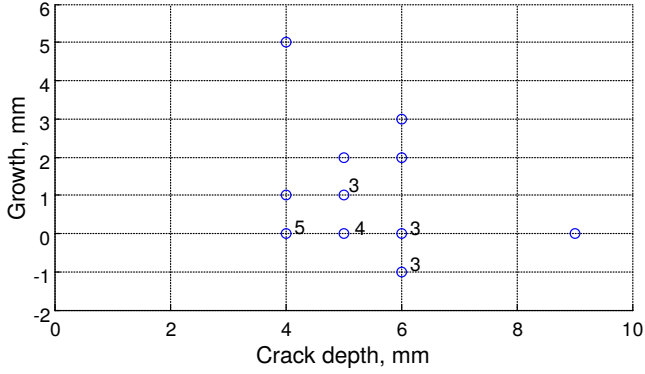


Fig. 2. The cracks length over 4 mm versus crack growth. Numbers indicate the occurrence of the data point.

such as Gibbs sampling, and the Metropolis–Hastings algorithm, slice sampling is easier to implement as only the posterior needs to be specified (Gilks, 1996).

#### 4. FAILURE RISK ASSESSMENT

##### 4.1 Rail Failure Probability

In this section, the risk model is presented. First, failure probability is calculated by considering squats with over 4 mm in crack depth measured by US. The probability of failure indicate how likely is a squat to develop into a rail break in the future. To evaluate the failure probability, we consider squats with crack depths ranging from 1 mm to 9 mm. By measuring the depths every one year, we see how many cracks have reached depth of 4 mm or even more, and how the cracks growth over time. Then, we enumerate the squats with the same growth and crack depth, to capture the typical behaviour of squats in the particular track. Fig. 2 shows the occurrence of cracks of more than 4 mm over a track segment of around 2.35 kilometres during a period of 4 years. The Mega Gross Tone (MGT) is equal to 3.719 per year in this track. The data collected in the Fig. 2 is used to estimate the Bayesian parameters,  $a$  and  $b$ , in order to capture the failure rate. The idea is to use crack depths for several different squats over time to calculate growth with regards to number of the squats with same growth in depth.

The posterior distribution of the regression parameters ( $a, b$ ) is calculated based on the MCMC simulation generated in one thousand samples. Fig. 3 and Fig. 4 show how the parameters ( $a, b$ ) vary over the samples. The posterior distributions show updated state of the mean value and the level of the uncertainty of the model parameters. As seen in the figures, the purpose is to check for convergence using sample means. This produces a smoother plot than the raw sample traces, and can make it easier to identify and understand any non-stationarity. The first fifty values of Fig. 3 are not comparable to the rest of the figure. However, the rest of each plot shows that the parameter posterior means have converged to stationarity.

The probability failure regression models resulted from  $N$  samples of MCMC simulation are depicted in Fig. 5, where  $N$  is equal to 1000. The idea of Fig. 5 is to show how squat will be prone for rail break in the future. In this figure, the non-linear regression models

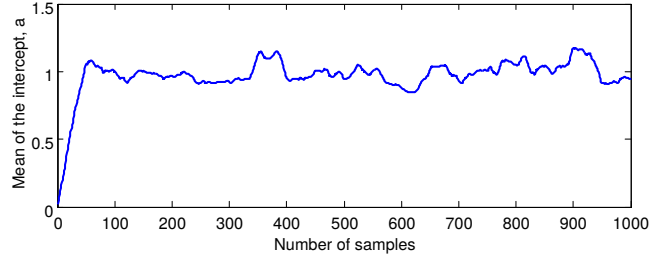


Fig. 3. Posterior distributions of regression parameter  $a$

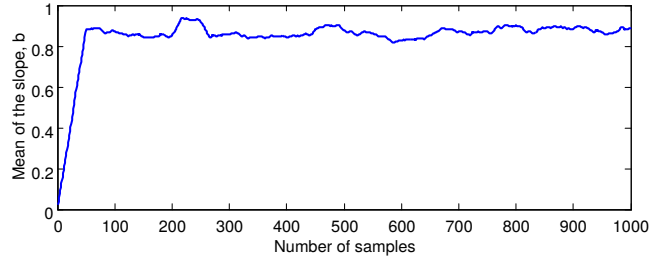


Fig. 4. Posterior distributions of regression parameter  $b$

Table 1. Bayesian point and interval estimates

Scenario	Parameter	Mean	95 % confidence interval
$s_1$	$a_1$	0.8001	[0.9860, 1.0089]
	$b_1$	0.8000	
$s_2$	$a_2$	1.0009	[0.8589, 0.8667]
	$b_2$	0.8624	
$s_3$	$a_3$	1.9592	
	$b_3$	1.1346	

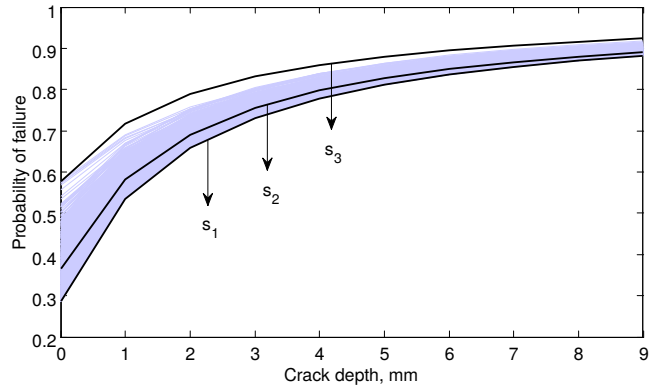


Fig. 5. Bayesian estimates for rail failure probability after a time horizon sufficiently long to guarantee a timely maintenance.

of  $s_1, s_2, s_3$  are used to reflect the optimistic failure scenario, the average scenario and the pessimistic scenario, respectively. Thus, relying on the figure, for each available crack depth, the probability of the squat to develop into a rail break is estimated within a time horizon, sufficient to guarantee a timely maintenance. For example, the squats with crack depth 7 mm induces rail to be broken if we do not maintenance operations in a long time horizon, with probabilities increasing according the scenario: 0.8554 for  $s_1$ , 0.8668 for  $s_2$  and 0.9068 for  $s_3$ . Point estimates and Bayesian confidence intervals, representing uncertainty about parameters after data analysis are presented in Table 1.

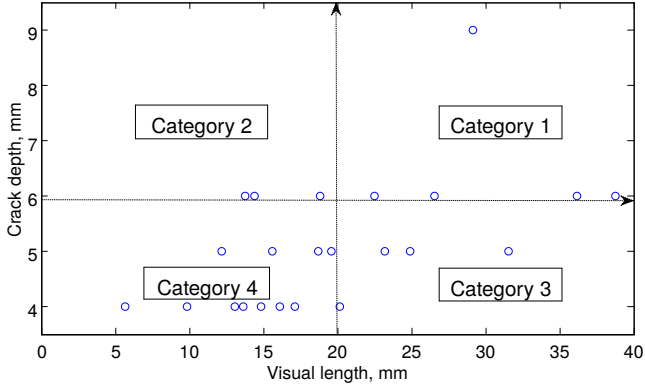


Fig. 6. Experimental categories of squat based on crack depths and the visual lengths

#### 4.2 Squat Severity Analysis

As the visual lengths of squats follow a specific growth model classifying the squats according to its severity, in this section, a relation on how the visual lengths and the crack lengths are linked to each other is investigated. The idea is to classify the severity of the squat when it is getting worse in terms of both, the visual length and the crack length. For this purpose, the visual length and the crack depth of 36 squats were registered every six months over 2 years (see Fig. 5).

As depicted in Fig. 6, the squat growth space is divided into four categories representing the squat severity. The reason behind specifying the category boundaries is that the squats with visual length above 20 mm will potentially reappear after a grinding operation. Thus, Category 1 shows the most severe growth of squats where the crack length and visual length both are sufficiently high to require maintenance as soon as possible. Contrary to the Category 1, Category 4 is a safe category reflecting all the squats which are located in the early stage of growth. There are a few squats observed in categories 2 and 3. Even though the squats situated in Category 2 are in early stage of growth in terms of visual length, the crack depth are considerably high. In Category 3 the visual length is high whereas the crack depth is below 4 mm.

#### 4.3 Rail Failure Risk

Relying on the failure probability scenarios and the severity categories, the risk can be defined as:

$$R_s^z = w_z \int_{l_{z_i}}^{l_{z_{i+1}}} P_s(l) dl \quad (5)$$

where  $R_s^z$  is the rail failure probability for scenario  $s$  and Category  $z$ ,  $w_z$  is the severity weight of Category  $z$  and  $[l_{z_i}, l_{z_{i+1}}]$  is the interval of crack depths that defines Category  $z$ . The idea is to use the failure probability of the crack depth  $l$ ,  $P_s(l)$ , for each category of severity,  $z$ , by considering the severity weight  $w_z$ . Table 2 shows the resulting risk values of each scenario at different categories. To illustrate, as expected, the risk of failure in scenario  $s_3$  for Category 1 is the highest where the crack are most severe both in the length and the depth while the risk for scenario  $s_1$  in the Category 4 contains the lowest value.

Table 2. Failure risk results

Risk Scenario	Category 1	Category 2	Category 3	Category 4
$s_1$	0.7758	0.5172	0.1621	0.0811
$s_2$	0.7862	0.5241	0.1654	0.0827
$s_3$	0.8173	0.5449	0.1753	0.0876

The failure risk values can be used as risk Key Performance Indicators (KPIs) to address health condition of the rail, so to keep informed infrastructure manager of the status of track. In combination with other KPIs as defined in Jamshidi et al. (2015), the risk values can be employed to support a condition-based maintenance plan.

## 5. CONCLUSIONS

In this paper, a probabilistic approach is used to model rail failure considering the squat growth. A Bayesian method was employed to make robust failure estimation, including optimistic, average and pessimistic scenarios. Furthermore, uncertainties of the method are also obtained and used to calculate Bayesian confidence intervals per failure scenario. Then, according to where the squat is in the severity categories, the rail failure risk is obtained per failure scenario. In future studies, we will develop the methodology to analytically predict the rail failure over a time horizon using risk key performance indicators relying on different measurement sources. Parameters like mechanical strength values, material properties and geometrical values of the rail such as area of cross section, can give further details about the way the crack will evolve over time. The evaluation on how those parameters influence the risk assessment is part of the further research.

## ACKNOWLEDGEMENTS

This research is supported by ProRail and MAGW, and by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs.

## REFERENCES

- Ahn, S., Park, C., and Kim, H. (2007). Hazard rate estimation of a mixture model with censored lifetimes. *Stochastic Environmental Research and Risk Assessment*, 21(6), 711–716. Doi: 10.1007/s00477-006-0082.
- Andrade, A. R., and Teixeira, P. F. (2015). Statistical modelling of railway track geometry degradation using hierarchical Bayesian models. *Journal of Reliability Engineering & System Safety*, 142, 169–183. Doi: 10.1016/j.res.2015.05.009.
- Bearfield, G., and Marsh, W. (2005). Generalising event trees using Bayesian networks with a case study of train derailment. *The 35th International Conference on Computer Safety, Reliability, and Security*, (3688), 52–66. Trondheim, Norway, 20–23 September. Doi: 10.1007/11563228.
- Bouillaut, L., François, O., Leray, P., Aknin, P., and Dubois, S. (2008). Dynamic Bayesian networks modelling maintenance strategies: Prevention of broken rails. *8th World Congress on Railway Research*, Seoul, South Korea. 18–22 May 2008.
- Burstow, M. C., Watson, A. S., and Beagles, M. (2002). Application of the whole life rail model to control rolling contact fatigue. *Proceedings of the International Conference Railway Engineering*, London, UK, July 2002.

- Castillo, E., Calviño, A., Grande, Z., Sánchez-Cambronero, S., Gallego, I., Rivas, A., and Menéndez, J. M. (2015). A Markovian-Bayesian Network for Risk Analysis of High Speed and Conventional Railway Lines Integrating Human Errors. *Journal of Computer-Aided Civil and Infrastructure Engineering*, 1467-8667. Doi: 10.1111/mice.12153.
- Faghih-Roohi, S., Xie, M., and Ng, K. M. (2014). Accident risk assessment in marine transportation via Markov modelling and Markov Chain Monte Carlo simulation. *Ocean Engineering*, 91, 363-370. Doi: 10.1016/j.oceaneng.2014.09.029.
- Gilks, W., Richardson S., and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London.
- Hassankiadeh, S. J. (2011). Failure Analysis of Railway Switches and Crossings for the purpose of Preventive Maintenance, *M.Sc thesis*. KTH Stockholm Sweden.
- Jamshidi, A. Núñez, A. Li, Z., and Dollevoet, R. P. B. (2015). Maintenance decision indicators for treating squats in railway infrastructures. *Proceedings of the 94th Annual Meeting of the Transportation Research Board*, Washington, USA, January 11-15 2015.
- Kaewunruen, S., and Ishida, M. (2014). Field monitoring of rail squats using 3D ultrasonic mapping technique, *Journal of the Canadian Institute for Non-Destructive Evaluation*. 35(6), 5-11.
- Lam, H. F., Hu, Q., and Wong, M. T. (2014). The Bayesian methodology for the detection of railway ballast damage under a concrete sleeper. *Journal of Engineering Structures*, 81:289-301. Doi: 10.1016/j.engstruct.2014.08.035.
- Li, Z., Molodova, M., Núñez, A., and Dollevoet, R. (2015), Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure. *IEEE Transactions on Industrial Electronics* 62(7): 4385-4397. Doi:10.1109/TIE.2015.2389761
- Li, Z., Zhao, X., Molodova, M., Dollevoet, R. P. B. (2010). Squat treatment by way of minimum action based on early detection to reduce life cycle costs. *Proceedings of the Joint Rail Conference JRC*, Urbana, Illinois, 27-29 April.
- Liu, X., Barkan, C., and Saat, M. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. *Transportation Research Record: Journal of the Transportation Research Board* (2261),178-185.
- Magel, E. E. (2001). Rolling Contact Fatigue: A Comprehensive Review, *Report, DOT/FRA/ORD-11/24*, Federal Railroad Administration.
- Mahboob, Q. (2014). A Bayesian Network methodology for railway risk, safety and decision support, *Doctoral dissertation*, TU Dresden.
- Marsh, W., and Bearfield, G. (2004). Using Bayesian networks to model accident causation in the UK railway industry. *Probabilistic Safety Assessment and Management*, 3597-3602. Doi: 10.1007/978-0-85729-410-4\_575.
- Mokhtarian, P., Namzi-Rad, M. R., Ho, T. K., and Suesse, T. (2013). Bayesian nonparametric reliability analysis for a railway system at component level. *IEEE International Conference on Intelligent Rail Transportation (ICIRT)*, 197 – 202, Beijing, 30 August-1 September. doi: 10.1109/ICIRT.2013.6696293.
- Molodova, M., Li, Z., Núñez, A., and Dollevoet, R. P. B. (2014) Automatic detection of squats in the railway infrastructures. *IEEE Transactions on Intelligent Transportation Systems* 15(5):1980-1990. Doi: 10.1109/TITS.2014.2307955
- Neal, R. M. (2003). Slice sampling. *Annals of Statistics*, 31 (3), 705-741.
- Prescott, D., and Andrews, J. (2013). Modelling maintenance in railway infrastructure management. *In Reliability and Maintainability Symposium (RAMS)*, Orlando, FL.
- Zhao, J., Chan, A. H., and Stirling, A. B. (2006). Risk analysis of derailment induced by rail failures-a probabilistic approach. *In Reliability and Maintainability Symposium, 2006. RAMS'06 Annual*, 486-491. 23-26 January, Newport Beach, CA. Doi: 10.1109/RAMS.2006.1677421
- Zoeteman, A., Dollevoet, R. P. B., and Li, Z. (2014). Dutch research results on wheel/rail interface management: 2001-2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 228(6), 642-651. Doi: 10.1177/0954409714524379.