



FELIX PRAUSE<sup>1</sup>, RALF BORNDÖRFER<sup>2</sup>

# **Construction of a Test Library for the Rolling Stock Rotation Problem with Predictive Maintenance**

---

<sup>1</sup>  0000-0001-9401-3707

<sup>2</sup>  0000-0001-7223-9174

Zuse Institute Berlin  
Takustr. 7  
14195 Berlin  
Germany

Telephone: +49 30 84185-0  
Telefax: +49 30 84185-125

E-mail: [bibliothek@zib.de](mailto:bibliothek@zib.de)  
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064  
ZIB-Report (Internet) ISSN 2192-7782

# Construction of a Test Library for the Rolling Stock Rotation Problem with Predictive Maintenance

Felix Prause<sup>[0000-0001-9401-3707]</sup> and Ralf Borndörfer<sup>[0000-0001-7223-9174]</sup>

Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany  
{prause, borndoerfer}@zib.de

**Abstract.** We describe the development of a test library for the rolling stock rotation problem with predictive maintenance (RSRP-PdM). Our approach involves the utilization of genuine timetables from a private German railroad company. The generated instances incorporate probability distribution functions for modeling the health states of the vehicles and the considered trips possess varying degradation functions. RSRP-PdM involves assigning trips to a fleet of vehicles and scheduling their maintenance based on their individual health states. The goal is to minimize the total costs consisting of operational costs and the expected costs associated with vehicle failures. The failure probability is dependent on the health states of the vehicles, which are assumed to be random variables distributed by a family of probability distributions. Each distribution is represented by the parameters characterizing it and during the operation of the trips, these parameters get altered. Our approach incorporates non-linear degradation functions to describe the inference of the parameters but also linear ones could be applied. The resulting instances consist of the timetables of the individual lines that use the same vehicle type. Overall, we employ these assumptions and utilize open-source data to create a library of instances with varying difficulty. Our approach is vital for evaluating and comparing algorithms designed to solve the RSRP-PdM.

**Keywords:** Rolling stock rotation planning · Predictive maintenance · Test library.

## 1 Introduction

Rolling stock rotation planning is an important topic that has been investigated by a variety of authors, see for example [11] and the citations therein. The planning of these rotations is often combined with the scheduling of maintenance but mainly preventive maintenance regimes are considered. However, predictive maintenance has some benefits compared to preventive maintenance: On the one hand, it offers the advantage that components with a high level of wear can be detected and serviced before a failure occurs. On the other hand, components that have not yet reached their end of life can continue to be used. Due to the

advances in sensor technique and machine learning necessary to process the sensor measurements, predictive maintenance can now be put into practice. There already exists a variety of approaches for determining the remaining useful life (RUL) of mechanical components but we are only aware of few articles, i.e., [3,6,7,13,14], incorporating the RUL and its predictions into rolling stock rotation planning and maintenance scheduling. Furthermore, there exists a lack of test instances concerning predictive maintenance, which was already stated by [2] claiming that the authors of predictive maintenance approaches only compare their work to other maintenance regimes instead of comparing it to the methods of other authors. In order to overcome these problems, we present the construction of a set of test instances originating from genuine timetables of a private German railroad company enriched by some realistic assumptions.

## 2 Timetables, Network & Vehicles

**Timetables, Stations and Distances** The timetables are taken from a private German railroad company and are all genuine and in operation in 2023. Information about the vehicle types operating the individual lines are taken from public sources of the railroad company. The distances between the stations of the rail network are taken from [9] using [4], while the maintenance locations origin from the official website of the railroad company. A visualization of the rail network on which the trips of the timetables are operated is given in Figure 1a.

**Costs** The costs can be distinguished into distance- and time-dependent costs, as well as failure-related costs. The first are assumed, depending on the vehicle type, to vary between 6.80 and 11.20 per km. This information is taken from [8], where the given values have been reduced by 38%, i.e., by the share of costs for profit, maintenance, and capital costs. Next, we determined the yearly costs. Therefore, we took the purchase prices of the vehicles and combined them with the associated amortization rates and expected costs for the respective maintenance facilities. These costs range from 120,000 to 1,340,000 depending on the vehicle type. The maintenance costs themselves are assumed to be 2,000 based on the assumption that the average maintenance costs for regional trains are about two-thirds of the maintenance costs for the German high-speed train ICE, see [1]. Finally, we assume that the breakdown of a vehicle causes costs of 5,000 – 50,000 depending on its number of seats.

**Maintenance & Vehicle Specifications** We assume that maintenance services on the train doors take 02:15 hours. The considered vehicle types are taken from the railroad company’s website and we have used the manufacturer’s information on the drive type of the trains and the number of their seats. Furthermore, we assume that the driving speed of the vehicles is 80 km/h when deadheading.

### 3 Modeling the Degradation

**Degradation Caused by Trips** In this article, the components of interest are the train doors. We measure their service life and the load they endure by the number of opening-closing cycles they undergo. The number of possible cycles is taken from [12], who observed that 4,630 opening-closing cycles led to three failures, i.e., a failure occurred approximately every 1,500 cycles. We assume that the number of cycles that occur at each stop of a trip varies between zero and four, depending on the number of passengers waiting at the platform of the station. To determine this number, we take data on the hourly passenger volume, multiply it by the total number of stops occurring at each station, and scale it s.t. the maximal value is equal to three. This procedure yields an individual parameter  $\lambda \in (0, 3]$  for each of the stations that we associate with a Poisson distribution. Afterwards, to obtain the degradation caused by each trip, we sample from the probability distribution function (PDF) of each station, where we round values greater than four down to four, and sum up all samples. This procedure is repeated multiple times and gives rise to a histogram, to which we fit a normal distribution describing the number of cycles that occur during each trip. Here, we utilize the Poisson distribution since it is widely used to model the arrival of passengers, see e.g. [10], and the normal distribution since Poisson distributions are stable and approximately normal for large values of  $\lambda$ .

**Health States of the Vehicles** Since the health states of the vehicles cannot be measured directly, i.e., we have to rely on indirect measurements like counting the number of operated cycles or evaluating sensor measurements, they have to be assumed as uncertain and considered as random variables. The same applies to the degradation caused by the trips since the number of opening-closing cycles that occur during their operation can only be estimated. Thus, in the case of a linear degradation behavior, we would need to determine the convolution of both PDFs. Measurement errors are often modeled by normal distributions and we chose to model the number of opening-closing cycles occurring during the operation of the trips as normally distributed as well. Since normal distributions are uniquely determined by their mean  $\mu \in \mathbb{R}$  and variance  $\sigma^2 \in \mathbb{R}_{>0}$ , we identify each random variable by its corresponding parameters. Furthermore, we assume that the health state of a vehicle is a random variable  $H_v \sim \mathcal{N}(\mu_v, \sigma_v^2)$ , where  $H_v = 1$  indicates that  $v$  is as good as new and  $H_v \leq 0$  corresponds to a failure of the vehicle. Hence, the failure probability of vehicle  $v$  is given by

$$\mathbb{P}_f(\mu_v, \sigma_v^2) := \mathbb{P}[H_v \leq 0] = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{-\mu_v}{\sqrt{2\sigma_v^2}} \right) \right).$$

Next, we describe the degradation function  $\Delta$  mapping the parameters of a vehicle and those of the trip it performs to the parameters of the health state after the trip is operated. This degradation function needs to take into account that already worn out vehicles deteriorate faster than new or recently maintained

ones. Additionally, it should exhibit a non-linear behavior since mechanical components usually show non-linear degradation patterns, as for example described by [5]. Therefore, we define  $\Delta$  as follows:

$$\Delta : \mathbb{R}^2 \times \mathbb{R}_{>0}^2 \rightarrow \mathbb{R} \times \mathbb{R}_{>0},$$

$$\begin{pmatrix} \mu_v \\ \mu_t \\ \sigma_v^2 \\ \sigma_t^2 \end{pmatrix} \mapsto \begin{pmatrix} \mu_v - \frac{1}{7,500(1+\mu_v)} - \frac{3\mu_t}{5,000} \\ \sigma_v^2 + \frac{1}{7,500^2(2-\sigma_v^2)} + \frac{8\sigma_t^2}{1,500^2} \end{pmatrix},$$

where  $(\mu_v, \sigma_v^2)$  are the parameters corresponding to the PDF of the health state of vehicle  $v$  and  $(\mu_t, \sigma_t^2)$  characterize the normal distribution of the number of operated cycles that occur during trip  $t$ . We designed  $\Delta$  so that a new or recently maintained train door would fail after approximately 1,500 cycles and that the rate of degradation increases with time. Finally, we assume that maintenance resets the parameters of a vehicle to  $\mu = 0.95$  and  $\sigma^2 = 0.0003$ .

## 4 The Resulting Instances

The instances resulting from the previously described construction procedure can be found on a server operated by our institute<sup>1</sup> and are completely anonymized. As already mentioned, the instances are based on the timetables of lines sharing the same vehicle type. The characteristics of the instances can be found in Table 1 containing the number of lines in the instance, the number of trips that need to be operated within one week, the number of considered locations, and the number of vehicles that may be used to operate the timetables. Figures 1b – 1g show the underlying rail networks of each instance. The objective for each of the instances is to determine rolling stock rotations and maintenance schedules, s.t. all trips of the given timetable are operated with minimum total costs. These costs consist of the operation costs of trips, deadhead trips, and maintenance activities, as well as the expected failure costs of the vehicles during trip operation. Here, the failure probability of a vehicle during a trip is determined by the health state of the vehicle after the operation of this trip.

## 5 Conclusion

Throughout this article, we described the construction of a test library for the RSRP-PdM. The instances of the library are based on genuine timetables of a German railroad company, publicly available real-world data, and were supplemented by meaningful assumptions. These assumptions include a non-linear degradation behavior and provide the opportunity to establish a maintenance regime scheduling the service tasks based on the failure probability of the vehicles, rather than setting thresholds for the health states. The instances vary in size and difficulty, and can be used to evaluate and compare algorithms for RSRP-PdM, filling the gap identified by [2].

<sup>1</sup> <https://cloud.zib.de/s/NZSWpyWL5YXS5ww>

Table 1: Number of lines, trips, locations, and vehicles of the resulting instances.

Instance	Lines	Trips	Locations	Vehicles
T1	2	566	8	6
T2	2	608	10	7
T3	1	636	15	16
T4	3	679	9	8
T5	5	813	16	14
T6	2	919	17	29
combined	15	4,221	52	80

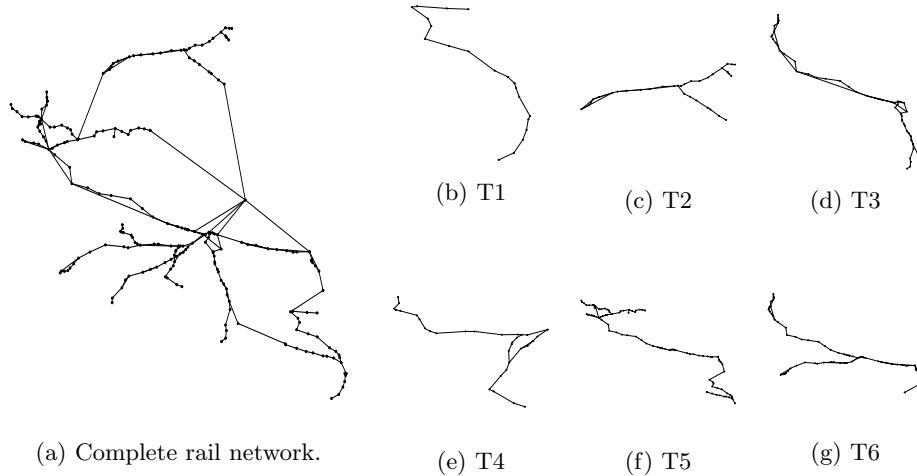


Fig. 1: The underlying rail network of the timetables (a), and the rail network restricted to the stations and routes used in each instance (b) – (g). Note that the straight lines possess the rail distance, but the intermediate stations have been omitted for the sake of simplicity since they are not stops of the considered trips.

**Acknowledgements** This work was supported by the innovation funding program Pro FIT (grant no. 10174564) funded by the State of Berlin and co-funded by the European Union.

## References

1. airliners.de: Kostenvergleich ICE/A320 - Bahn produziert billiger als jede Airline. <https://www.airliners.de/bahnfahren-aviation-management/51438>, accessed June 23, 2023
2. Bougacha, O.: Contribution to post-prognostic decision: a new framework based on prognostic/decision interaction. Ph.D. thesis, Université Bourgogne Franche-Comté (2020)
3. Bougacha, O., Varnier, C., Zerhouni, N.: Impact of decision horizon on post-prognostics maintenance and missions scheduling: a railways case study. *International Journal of Rail Transportation* **10**(4), 516–546 (2022). <https://doi.org/10.1080/23248378.2021.1940329>
4. BRouter contributors: BRouter: Advanced OSM Routing. <https://brouter.de>, accessed June 7, 2023
5. Heng, A., Zhang, S., Tan, A.C., Mathew, J.: Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical systems and signal processing* **23**(3), 724–739 (2009). <https://doi.org/10.1016/j.ymssp.2008.06.009>
6. Herr, N., Nicod, J., Varnier, C., Zerhouni, N., Cherif, M., Fnaiech, F.: Joint optimization of train assignment and predictive maintenance scheduling. In: *International Conference on Railway Operations Modelling and Analysis*. pp. 1–10 (2017)
7. Herr, N., Nicod, J.M., Varnier, C., Zerhouni, N., Dersin, P.: Predictive maintenance of moving systems. In: *2017 Prognostics and System Health Management Conference (PHM-Harbin)*. pp. 1–6. IEEE (2017). <https://doi.org/10.1109/PHM.2017.8079111>
8. Landesnahverkehrsgesellschaft Niedersachsen mbH: Kosten im SPNV. <https://www.lnvg.de/spnv/finanzierung/kosten-im-spnv>, accessed June 23, 2023
9. OpenStreetMap contributors: Planet dump retrieved from <https://planet.osm.org>. <https://www.openstreetmap.org> (2017), accessed June 7, 2023
10. Peters, R., Al-Sharif, L., Hammoudeh, A.T., Alniemi, E., Salman, A.: A systematic methodology for the generation of lift passengers under a poisson batch arrival process. In: *5th Symposium on Lift and Escalator Technologies*. vol. 5, pp. 177–186 (2015)
11. Reuther, M.: Mathematical optimization of rolling stock rotations. Ph.D. thesis, Technische Universität Berlin (Germany) (2017)
12. Ribeiro, R.P., Pereira, P., Gama, J.: Sequential anomalies: a study in the railway industry. *Machine Learning* **105**, 127–153 (2016). <https://doi.org/10.1007/s10994-016-5584-6>
13. Rokhforoz, P., Fink, O.: Hierarchical multi-agent predictive maintenance scheduling for trains using price-based approach. *Computers & Industrial Engineering* **159**, 107475 (2021). <https://doi.org/10.1016/j.cie.2021.107475>
14. Wu, M.J., Lai, Y.C.: Train-set assignment optimization with predictive maintenance. In: *RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA)*, Norrköping, Sweden, June 17th–20th, 2019. pp. 1131–1139. Linköping University Electronic Press (2019)