Strengthening the Rail Mode of Transport by Condition Based Preventive Maintenance

René Schenkendorf, German Aerospace Center
Jörn Groos, German Aerospace Center
Lars Johannes, German Aerospace Center

Available at: https://works.bepress.com/rene_schenkendorf/9/
Strengthening the Rail Mode of Transport by Condition Based Preventive Maintenance *

R. Schenkendorf * J.C. Groos * L. Johannes *

* Institute of Transportation Systems, German Aerospace Center (DLR e.V.), Lilienthalplatz 7, 38108 Braunschweig, Germany (e-mail: {rene.schenkendorf, joern.groos, lars.johannes}@dlr.de).

Abstract: In recent years, the demands on railroad infrastructure operators have been rising by means of profitability, availability, safety, and punctuality. Here, the condition based preventive maintenance aims at strengthening the rail mode of transport through an optimized scheduling of maintenance actions taking account of the actual infrastructure condition and its expected further degradation. Two crucial aspects of such a predictive maintenance strategy are 1) the reliable precise localization of faults within a widespread, distributed infrastructure system, and 2) the consideration of parameter uncertainties by the prediction of the degradation trend for the near future of the infrastructure under study. Both aspects are addressed in this contribution, the challenges are highlighted and some illustrative, preliminary results are shown.

1. INTRODUCTION

Ever wonder why your train is running late? One potential reason might be that the railroad infrastructure has been subject of unscheduled maintenance actions, i.e. repair of blocked switches or corrections of misaligned track sections. Though railroad infrastructure operators have always run preventive maintenance strategies, system components may fail leading to adverse effects by means of passenger comfort and maintenance costs. For instance, connecting trains are missed due to delays and sudden faults cause ad-hoc maintenance actions which by nature are extremely costly. A major reason therefor is the variety in material, age, and operating conditions of infrastructure components. By way of example, local over-stress leads to local, premature aging of the infrastructure, which is ignored by conventional, fixed-time interval maintenance strategies. Nevertheless, no noteworthy condition-based predictive maintenance takes place for railway infrastructure at present. Two main reasons are the lack of a sufficient condition monitoring as well as reliable degradation models (stochastic or deterministic) for the various individual components of the entire railway network. Nowadays, track inspection is typically done in intervals of several months with dedicated measurement trains. These inspections ensure a safe railway operation but provide in far most cases not a suitable data base to derive and parameterize reliable degradation models. Furthermore, degradation models have to take into account also numerous parameters describing significant time- and/or spatial-varying factors influencing track condition such as traffic density, weather or geology which are far often only vaguely known. To strengthen the railway mode by condition-based predictive maintenance two major challenges have to be solved: (i) A nearly continuous condition monitoring of the entire railway network. (ii) Degradation prediction by a model-based prognosis including the handling of the uncertainties arising from the difficult parameterization of the models in operational maintenance. In this manuscript contributions addressing these two major challenges are presented. A promising approach to operate a nearly continuous condition monitoring is the installation of measurement systems on regular in-service trains. Such a monitoring of the entire network by numerous moving sensor systems calls for advanced localization, georeferencing, and data management concepts. Only by this way, potential faults can be identified, assigned to previous records, and their progression can be analyzed algorithmically to obtain degradation models. In this manuscript, our approach of an advanced georeferencing is presented. Degradation models (deterministic as well as stochastic concepts) have to be parameterized once they are available. As important factors regarding the degradation are more often only vaguely known, those parameters/factors should be chosen with care which have the strongest impact onto the model outcome. Thus, a parameter sensitivity analysis should be mandatory and deserves a detailed explanation. In particular, it is shown how global parameter sensitivities can be calculated efficiently by combining Polynomial Chaos Expansion and Point Estimate Method principles.

2. LOCALIZATION

The comparison of failures detected at different inspections over time requires their localization. Failures at the rail
surface like squats (Fig. 1) occur separately and have therefore to be localized individually, others like rail corrugation extend over an area and are described by a localized start and end point. Failures of the rail alignment occur typically with wave lengths between 3 m up to 100 m. The association of locations (track ID and distance on track) to failures is called georeferencing.

![Fig. 1. Two snapshots of different failures at the rail surface (left: squat; right: corrugation)](image)

Nowadays, the localization of measurement trains is typically realized by the train’s odometer (instrument to indicate traveled distance by wheel turns) and the measurement protocol (manual/assisted documentation of start/end point of inspection and passage time of landmarks). The typical uncertainty of the longitudinal track position (distance on track) is up to dozens of meters. The analysis of the recorded data is usually done manually by attending responsible local asset managers with expert knowledge about the specific track. An automated and reliable analysis of the condition development of railroad tracks is hardly possible under these technical requirements, but a mandatory prerequisite for condition based preventive maintenance.

A crucial step is therefore the development of technologies for the automated, data-driven, and precise (track selective, uncertainties below 10 m) georeferencing of the condition information gathered by autonomous multi-sensor-systems on trains (e.g. [Lüddeke et al. (2012)]). Such systems will furthermore allow to use sensor systems mounted on regular in-line trains to monitor the track condition in a nearly continuous manner.

### 2.1 Multi-sensor concept

The sensors of the localization system presented here are a GNSS (Global Navigation Satellite System) receiver, a balise-antenna (balises are components located in the track bed to broadcast information such as speed limits to passing trains), an odometer and a speed sensor (e.g. a Doppler radar sensor). A digital map of the railroad network is the basis for position calculation. Furthermore, an inertial measurement unit (IMU) can be used to assist the localization for example by the detection of passed switches (e.g. [Rahmig and Lüddecke (2011)]). The trainborne sensors allow the estimation of an approximate position in near real-time at the train in the first step. Due to higher accuracies the final georeferencing of the gathered measurement data is done in post-processing. Track-selective localization is done by fusing the data of the mentioned sensors considering the geometry information about the rail network given by the digital map.

The Constant Turn Rate and Velocity model (CTRV), a common model for vehicle tracking, is used within an Extended Kalman Filter (EKF).

GNSS and the balise system are absolute sensors. Their data accuracy does not decrease over time/traveled distance but one or both of them may be not available at a specific track segment (e.g. due to shadowing effects or on poorly equipped secondary lines). The accuracy of the GNSS localizations is improved by the application of correction parameters and/or a differential evaluation in the post processing (GNSS Real Time Kinematics RTK). By this step uncertainties of less than 0.5 m can be reached under favorable conditions. The balise system consists of the balise-antenna mounted at the train and the balises installed in the track bed. Balises are stored in the digital map with their coordinates (uncertainties typically less than 0.1 m) and unique identification numbers. Both absolute sensor systems are capable to ensure a track-selective localization and complement each other. The localization principle is shown in figure 2.

The initialisation of the localization is done by an absolute sensor measurement (balise or GNSS RTK). Thereafter relative sensors in combination with the digital map is the basis for localization. As seen in figure 2 the localization accuracy (red line) gets more worse due to the limited accuracy of the relative sensors. By further detected balises or at GNSS RTK points the localization can be improved.

### 2.2 Estimated localization accuracies

The described localization method can be seen as a pre-localization, which can be significantly improved by a correlation analysis of the data of repeated measurements itself. The type of failure under analysis and the characteristics of the measurement system define the acceptable uncertainty of this pre-localization to ensure a reliable fine tuning by the correlation analysis. Assuming a failure of track alignment with a long periodicity of 5 up to 70 meters, so called long-wavelength failures, need a localization accuracy of about one quarter of the period to match correctly. In this case accuracy has to be better than 1.2 meters. The squat (Figure 1), a short-wave failure, may demand even smaller uncertainties.

As mentioned before the absolute sensors are responsible to keep accuracy on a sufficient level. Thus the distance in between the available absolute sensor measurements is of interest. The one sigma accuracies of the different sensors are given in the following listing. Its assumed that the data is normal distributed.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Accuracy (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNSS RTK</td>
<td>0.02 - 0.20</td>
</tr>
<tr>
<td>balise</td>
<td>0.20</td>
</tr>
<tr>
<td>odometer</td>
<td>0.4% (of covered</td>
</tr>
<tr>
<td></td>
<td>distance)</td>
</tr>
<tr>
<td>speed</td>
<td>0.8% (of covered</td>
</tr>
<tr>
<td></td>
<td>distance)</td>
</tr>
</tbody>
</table>

The error estimation of the two relative sensors and the combination of both over the traveled distance is shown in Figure 3. Due to the scaling factor the error level raises up with distance. The error level of the combined solution is little bit less than the single errors of each sensor.
2.3 Testing

For testing of the presented concept the multi-sensor system is implemented in the two-way rail and road testing vehicle RailDriVE® of the DLR (Fig. 4) and realized as a portable version for upcoming measurement campaigns on regular rail vehicles in 2015. An example for the detectability of rail corrugation by means of increased vibrations of the rail vehicle is shown in (Fig. 5). The measurement campaigns on regular rail vehicles within a complex railway network will yield information about the localization accuracies which can be reached in practice. These boundary conditions significantly influence for which failure types a reliable condition based preventive maintenance is feasible.

3. PROGNOSTICS

Assuming suitable degradation models and condition monitoring are available, the remaining useful life (RUL) of technical devices (e.g. switch engines, signals) or critical track defects (e.g. track settlement/misalignment, rail roughness) of rail segments under study can be prognosticated. Subsequently, ongoing maintenance actions can be optimally re-scheduled. The prognosticated infrastructure performance indices, however, might be affected by various uncertain factors, as: (i) imprecise parameters of the applied degradation model (e.g. soil/rail quality and soil moisture), (ii) vaguely known operating conditions (e.g. future temperature and load stresses/tonnages), and (iii) empirically chosen tuning/nuisance parameters of pre- and post-processing steps (e.g. noise filtering and feature selection calculations). In practice, the impact of these uncertain factors onto the model result (e.g. RUL) varies strongly. Some of them can be changed by order of magnitude without any significant model response variation, whereas slight displacements of other quantities might have a serious impact. In sort, the model response can be sensitive or insensitive with regard to the mentioned uncertainties.

The sensitive group of uncertain factors has to be treated with great care, the insensitive group, however, can be considered as deterministic identities and kept at fixed values.

A systematic quantification and classification of the vaguely known factors can be conducted by a sensitivity analysis. In what follows, a motivation for the so-called Distribution Based Global Sensitivity Analysis (DB-GSA) is given. In particular, it is shown how the DB-GSA can be calculated at low computational costs by combining Polynomial Chaos Expansion (PCE) and Point Estimate Method (PEM) principles. This seems to be a reasonable choice as PCE provides workable meta-models which can be efficiently parameterized by PEM-approximated multi-dimensional integrals.
The effect of uncertain factors onto the model response can be described generically by the following mapping problem:

\[ y = g(x) \]  

(1)

The model response, \( y \in \mathbb{R}^n \), expresses the prognosticated infrastructure performance index (e.g. track settlement) and the uncertain factors (e.g. soil stiffness, load stresses, speed, temperature, etc.) are summarized by the random vector \( x \in \mathbb{R}^m \), respectively. Moreover, \( g(\cdot) \) might represent an empirical or a mechanistic degradation model [Yousefkia et al. (2014)].

For non-linear problems, this local approach is only valid for slight uncertainties (i.e. the variance of \( x_i \) is low) and if \( g(\cdot) \) is analytical and differentiable (i.e. no black-box model). In many applications, however, this is not the case and derivative-free GSA concepts have to be put in operation to ensure meaningful inferences.

Frequently, the Global Sensitivity Analysis is based on so-called Sobol’ Indices (SI-GSA) which address the scatter of the uncertain factors explicitly. The key idea of SI-GSA is to identify the variance contribution that each factor, \( x_i \), adds to the variance of the model response. For more details see [Sobol’ (1993), Sobol’ (2001), Saltelli et al. (2005)] and references therein.

In general, SI-GSA approaches the variance contribution exclusively. Effects caused by changes in the mean and higher moments (e.g. skewness or kurtosis) are ignored. For instance, a factor which leads to a strong offset/bias in the model response but contributes less to the response variance might be important for the model performance, too. Thus, so-called moment-independent importance measures have been recently introduced in the field of GSA [Bor gonovo (2007)]. Here, the analysis aims at quantifying the impact of a factor, \( x_i \), onto the probability density distribution (PDF) of the model response as shown in Fig. 6. That means, the difference between the unconditional PDF, \( p df(y) \), and a conditional PDF, \( p df(y|x_i) \), is determined according to

\[ s(x_i) = \int \Omega \left| p df(y) - p df(y|x_i) \right| dy \]  

(2)

As \( x_i \), however, is actually a random variable the expected value of \( s(x_i) \) has to be analysed.
large-scale applications analyzing various track segments, \( g(\cdot) \), this might be prohibitive due to limited cpu-power. A remedy is to evaluate handy surrogate functions, \( \hat{g}(\cdot) \), instead. Here, the Polynomial Chaos Expansion comes into play and is applied to bypass potential CPU-intensive processes as indicated in Fig. 7. The basics of PCE and its efficient parameterization by the Point Estimate Method are described in what follows.

In uncertainty analysis, PCE has become quite popular in the last years. PCE aims at representing the model response by a weighted superposition of deliberately chosen basis functions [Maitre and Knio (2010)], \( \Psi_i(\cdot) \), according to

\[
y = g(x) = \sum_{i=0}^{\infty} a_i \Psi_i(x)
\]

For practical applications the expansion is implemented in truncated form \( (l_{pce} << \infty) \) as

\[
\hat{y} = \sum_{i=0}^{l_{pce}} a_i \Psi_i(x)
\]

Here, the unknown coefficients, \( a_i \), can be derived by

\[
a_i = \frac{\int_\Omega g(x) \Psi_i \, df(x) \, dx}{\int_\Omega \Psi_i^2 \, df(x) \, dx}
\]

To keep the computational load low the integrals can be evaluated by the Point Estimate Method, which approximates the integrand by monomials of predefined degree [Schenkendorf (2014)]. In detail, by combining PCE with PEM an overall number of \( 2n_i^2 + 1 \) function evaluations of \( g(\cdot) \) is to be expected to provide reasonable results. Obviously, this corresponds to a significant reduction in comparison to standard numerical integration methods, e.g. Gaussian Quadrature or Monte Carlo simulations [Maitre and Knio (2010)]. Subsequently, the surrogate model response, \( \hat{y} \), can be derived at low computational costs, i.e. Monte Carlo simulations propagating the uncertain factors onto the model response are feasible. Hence, from a large number of samples of the resulting model responses, \( \hat{y}_n \), the associated PDF can be derived applying a Gaussian Kernel density approximation [Bishop (2008)].

In the same way, the conditional PDFs, \( pdf(y|x_i) \), can be determined to provide the basics of Eq.(2). Finally, the actual importance measure of an uncertain factor, \( x_i \), can be approximated by PEM, too. Hence, DB-GSA can be efficiently applied even for advanced and complex degradation models and maintenance optimization routines, respectively.

### 3.2 In-silico example

In a next step, an appropriate degradation model has to be derived and adapted, respectively. A process which is subject of ongoing work. Thus, the proposed concept of combining the Polynomial Chaos Expansion with the Point Estimate Method to efficiently apply DB-GSA is demonstrated by the following non-linear degradation function

\[
y(t) = g(x, t) = x_1 e^{-x_2(e^{-x_3}t)}
\]

As many track degradation and asset maintenance models include exponential terms [Quiroga and Schnieder (2012); Andrews et al. (2014) and references therein] this is a reasonable choice, at least for illustration purposes. In detail, the uncertain factors, \( x \in \mathbb{R}^3 \), are described by Gaussian distributions according to

\[
\begin{align*}
x_1 &\sim \mathcal{N}(5, 1) \\
x_2 &\sim \mathcal{N}(2, 1) \\
x_3 &\sim \mathcal{N}(3, 1)
\end{align*}
\]

In a first step, the approximation power of the proposed workflow is addressed. In Fig.(8) reconstructed PDFs of the model output by evaluating the original function \( g(\cdot) \) and, respectively, its surrogate \( \hat{g}(\cdot) \) are shown. In both cases 10,000 Monte Carlo simulations are applied according to \( x \). The PDFs of the model responses, \( y \& \hat{y} \), are derived by the Gaussian Kernel density approximation. In this case, there is no loss of information using \( \hat{g}(\cdot) \) instead of \( g(\cdot) \). In subsequent steps, the conditional PDFs are solely based on \( \hat{g}(\cdot) \) and used to quantify the importance measure (Eq.(4)). Hence, the impact of the uncertain factors onto the model response can be derived as illustrated in Fig.(9). In this particular case, the factor \( x_2 \) contributes at most at the very beginning of the degradation process, whereas \( x_1 \) dominates the long-term progression and should be known as precisely as possible to ensure reliable prognostics. All calculations are exclusively based on the surrogate function \( \hat{g}(\cdot) \). In case of more complex, CPU-intensive degradation models and optimization routines this might be of utmost relevance in sensitivity analysis.
A remedy is to evaluate handy surrogate functions, Fig. 7. To avoid potential CPU-intensive processes (dark gray) a handy, surrogate process (light gray) might be evaluated instead. In this regard, the Polynomial Chaos Expansion (PCE) is applied.

To keep the computational load low the integrals can be evaluated, e.g. optimization step \( x_i \) or noise filtering \( y \). The actual importance measure of an uncertain factor, \( g \), can be approximated by PEM, too. Hence, DB-GSA is featured as it addresses the characteristic of stochastic model for railway track asset management. The rail mode of transport will only benefit from a holistic approach, a research field which is subject of ongoing work at the Institute of Transportation Systems, DLR.

As demonstrated, the concept of a condition based preventive maintenance to strengthen the rail mode of transport is faced to a number of challenging problems. Two of them, the precise localization of track records and the prognostic of infrastructure degradation, have been explained in more detail. In this context, some general difficulties have been shown and potential solutions approaches have been highlighted by illustrative, preliminary results. In particular, DB-GSA is featured as it addresses the characteristic of the entire PDF in sensitivity analyzing resulting in credible sensitive indexes. By utilizing PCE in combination with PEM, these valuable indexes can be derived efficiently. Continual effort is required to calibrate and to validate the proposed concepts individually. At the end of the day, however, a condition based maintenance strategy can only be successful when addressing economical [Böhm (2013)] and operator staff [Naumann et al. (2013)] aspects as well. The rail mode of transport will only benefit from a holistic approach, a research field which is subject of ongoing work at the Institute of Transportation Systems, DLR.

4. CONCLUSION

As demonstrated, the concept of a condition based preventive maintenance to strengthen the rail mode of transport is faced to a number of challenging problems. Two of them, the precise localization of track records and the prognostic of infrastructure degradation, have been explained in more detail. In this context, some general difficulties have been shown and potential solutions approaches have been highlighted by illustrative, preliminary results. In particular, DB-GSA is featured as it addresses the characteristic of the entire PDF in sensitivity analyzing resulting in credible sensitive indexes. By utilizing PCE in combination with PEM, these valuable indexes can be derived efficiently. Continual effort is required to calibrate and to validate the proposed concepts individually. At the end of the day, however, a condition based maintenance strategy can only be successful when addressing economical [Böhm (2013)] and operator staff [Naumann et al. (2013)] aspects as well. The rail mode of transport will only benefit from a holistic approach, a research field which is subject of ongoing work at the Institute of Transportation Systems, DLR.

REFERENCES


