# Uncertainty Analysis in Railway Asset Management using the Point Estimate Method

Dr.-Ing. Thorsten Neumann (DLR)

Partly based on work by: Dr.-Ing. René Schenkendorf, M.Sc. Beate Dutschk

**1st European Railway Asset Management Symposium**

University of Nottingham

27th & 28th March 2018


DLR

# Motivation

© Kurt Michel / pixelio.de

- **Degradation**
- **Failures**
- **Reliability**

**UNCERTAINTY**

- Detection
- Diagnosis
- Prognosis

**Modelling**

- Statistical models
- Physical models
- Hybrid models

© Thomas Max Müller / pixelio.de

**Preventive and condition-based maintenance**

© Peter Freitag / pixelio.de

# Models and Uncertainty Analysis

**(Variable) Input $X$**

- Data / Measurements
- Parameters

→

(Mathematical)
**Model**
$f$

- Formula-based
- Num. Simulation
- …

→

**Output $Y$**

- Current degradation state
- Detected failure mode
- Future health condition
- Estimated life time
- …

$\mathbb{P}^X$ ✓

Uncertainty propagation →

$\mathbb{P}^Y = ?$

<u>Sources of uncertainty in modelling</u>

- Data / Measurements ✓
- Model parameters ✓
- Model structure ✗

<u>Mathematical notation</u>

$f: \mathbb{R}^n \longrightarrow \mathbb{R} \qquad Y = f(X)$

where

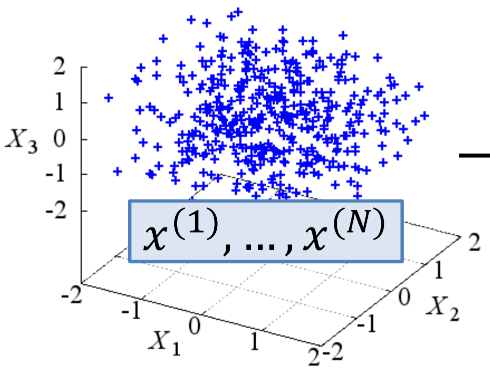
$X = (X_1, \ldots, X_n)^T$

($X_i$ stochastically independent)

# How to derive the output distribution?

$\mathbb{P}^X$ ✓ — Uncertainty propagation → $\mathbb{P}^Y = ?$

- Standard approach: **Monte Carlo simulation** (MC)




$\xrightarrow{f}$ $y^{(1)}, \dots, y^{(N)}$ $\rightsquigarrow$ $\widetilde{\mathbb{P}}^Y$ $\xrightarrow[N \to \infty]{}$ $\mathbb{P}^Y$

Random sample based on $\mathbb{P}^X$

Random sample based on unknown $\mathbb{P}^Y$

Numerical estimate of $\mathbb{P}^Y$ based on $y^{(1)}, \dots, y^{(N)}$

True $\mathbb{P}^Y$

- In general, **large samples** necessary!
- **Approximate results** only!

DLR

# Alternative approach: Point Estimate Method (PEM)

- Often it is **sufficient to know** some **basic statistics** of the output distribution

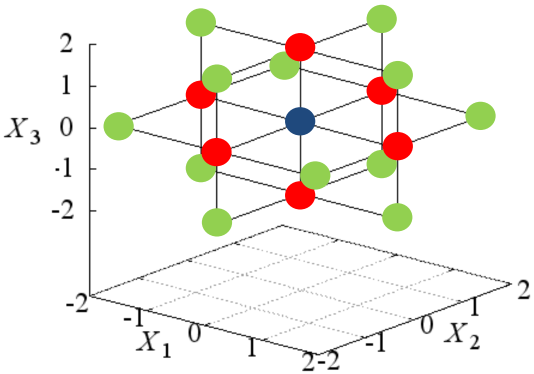
  Mean: $\mathbb{E}(Y)$ Variance: $\mathrm{Var}(Y)$ $k$-th moment: $\mathbb{E}(Y^k)$ …

- **Approximation scheme** (= core element of PEM):

$$\int_\Omega g(x)\,\mathrm{pdf}_X(x)\,dx \approx \boldsymbol{w_0}\, g(\mathbf{GF}[\mathbf{0}]) + \boldsymbol{w_1} \sum g(\mathbf{GF}[\pm\boldsymbol{\vartheta}]) + \cdots + \boldsymbol{w_m} \sum g\left(\mathbf{GF}\underbrace{[\pm\boldsymbol{\vartheta},\ldots,\pm\boldsymbol{\vartheta}]}_{m\text{ times}}\right)$$

with suitable **weights** $\boldsymbol{w_j}$ and **deterministic sample** points

- $\mathbf{GF}[\mathbf{0}]$
- $\mathbf{GF}[\pm\boldsymbol{\vartheta}]$
- $\mathbf{GF}[\pm\boldsymbol{\vartheta},\pm\boldsymbol{\vartheta}]$
- …




$\xrightarrow{g = f}$ $\mathbb{E}(Y)$

$\xrightarrow{g = f^2}$ $\mathbb{E}(Y^2)$

$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2$

# Standard scheme ($m = 2$)

$$\int_{\Omega} g(x)\,\mathrm{pdf}_X(x)\,dx \approx \boldsymbol{w_0}\, g(\mathbf{GF}[\mathbf{0}]) + \boldsymbol{w_1} \sum g(\mathbf{GF}[\pm\boldsymbol{\vartheta}]) + \boldsymbol{w_2} \sum g(\mathbf{GF}[\pm\boldsymbol{\vartheta}, \pm\boldsymbol{\vartheta}])$$

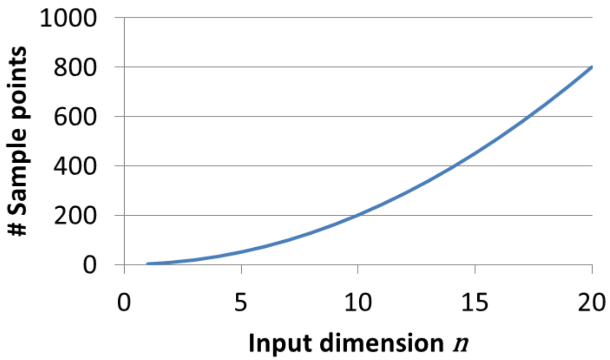
where

$$w_0 = 1 + \frac{n^2 - 7n}{18} \qquad w_1 = \frac{4-n}{18} \qquad w_2 = \frac{1}{36} \qquad \vartheta = \sqrt{3}$$

- Weights determined such that scheme is **exact for polynomials** $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ (degree $\leq 5$) in case of standard Gaussian inputs $X_i$ for $i = 1, \ldots, n$
- **Transformation** of the sample points in case of **other input distributions**

Number of sample points

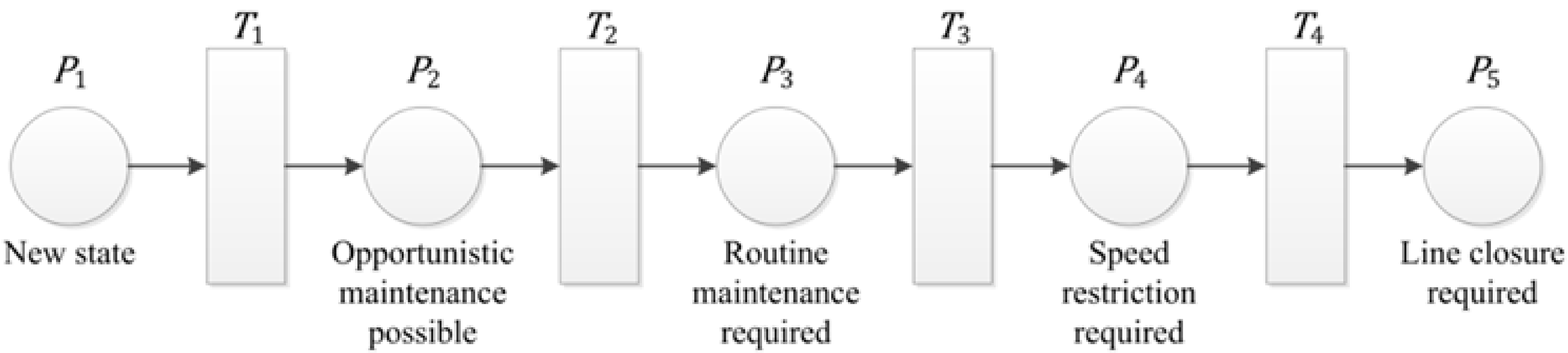
$$N = 2n^2 + 1$$




| $n$ | $N$ |
| --- | --- |
| 1 | (3) |
| 2 | 9 |
| 3 | 19 |
| 4 | 33 (25) |
| 5 | 51 |

# Example of use: Track degradation

- Simple **Petri net model** for track degradation (without maintenance):




*(cf. Andrews, 2012)*

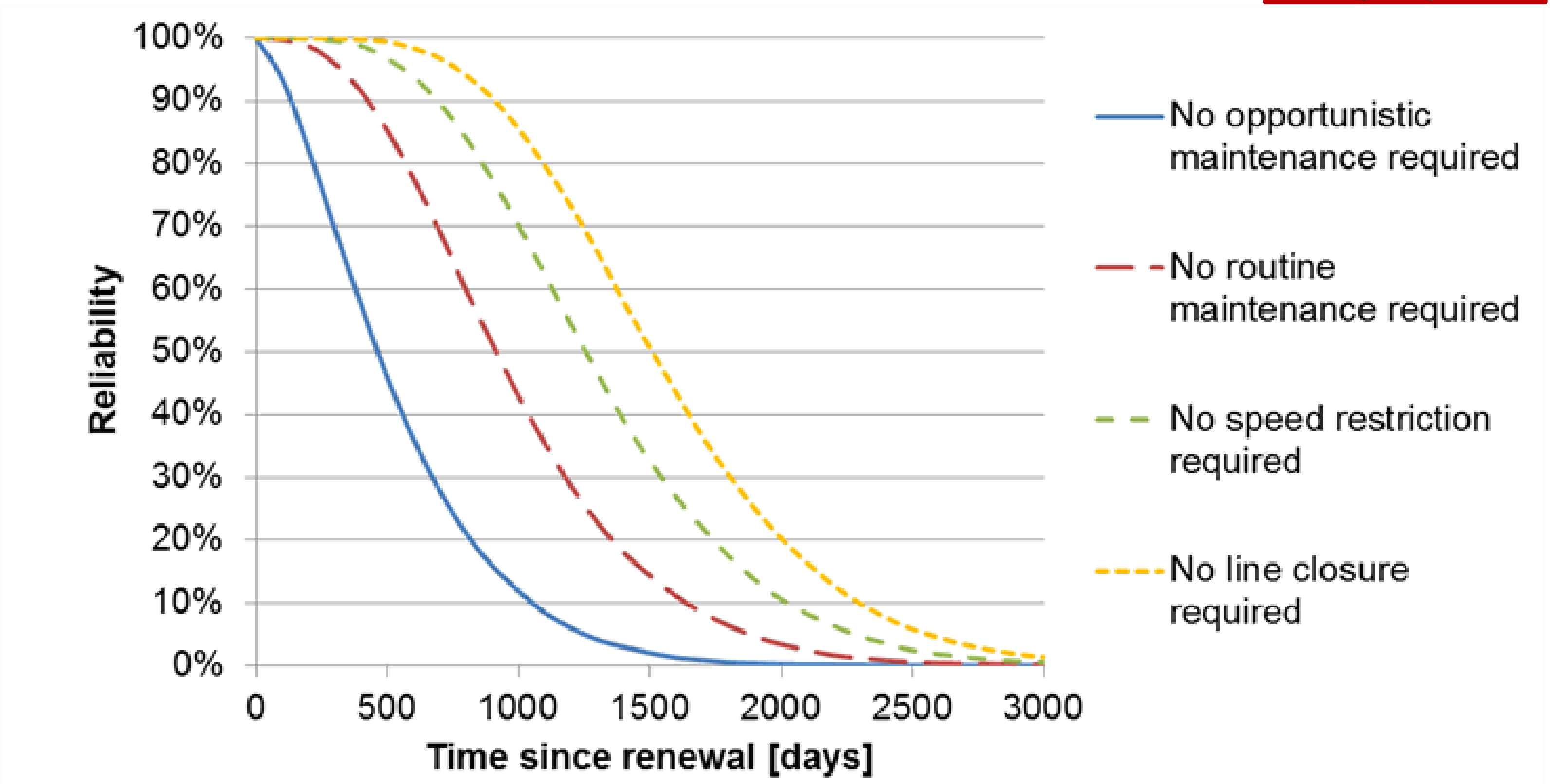
- Stochastic **transitions**: $T_r \sim \mathcal{W}(\beta_r, \eta_r)$
  (i.e., Weibull distributed duration [in days] before jumping to the next "health state")

| | $T_1$ | $T_2$ | $T_3$ | $T_4$ |
|---|---|---|---|---|
| $\beta_r$ (shape) | 1.5 | 1.5 | 1.6 | 1.7 |
| $\eta_r$ (scale) | 600 | 500 | 370 | 280 |

# Example of use: Track degradation

Estimated **track reliability** based on MC simulation

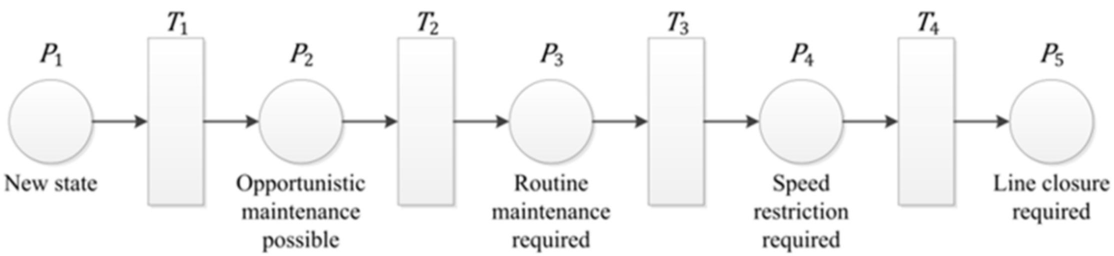
**10,000**
Sample points




# Example of use: Track degradation

- Duration [in days] before reaching state $P_{s+1}$ after renewal or new construction:

$$\tilde{T}_s := \sum_{r=1}^{s} T_r$$




| Analytical solution | $\mathbb{E}(\cdot)$ | $\sigma(\cdot)$ |
|---|---|---|
| $\tilde{T}_1$ | 541.6 | 367.8 |
| $\tilde{T}_2$ | 993.0 | 478.7 |
| $\tilde{T}_3$ | 1324.8 | 523.7 |
| $\tilde{T}_4$ | 1574.6 | 545.1 |

| $\Delta_{\text{rel}}$ (MC) | $\mathbb{E}(\cdot)$ | $\sigma(\cdot)$ |
|---|---|---|
| $\tilde{T}_1$ | –0.68% | +0.50% |
| $\tilde{T}_2$ | –1.00% | –0.03% |
| $\tilde{T}_3$ | –0.56% | +0.50% |
| $\tilde{T}_4$ | –0.49% | +0.67% |

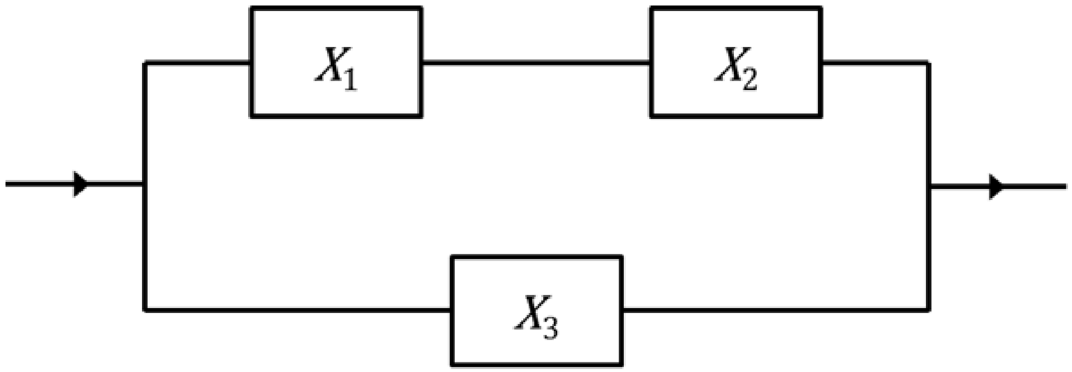
**10,000**
Sample points

| $\Delta_{\text{rel}}$ (PEM) | $\mathbb{E}(\cdot)$ | $\sigma(\cdot)$ |
|---|---|---|
| $\tilde{T}_1$ | +0.00% | –0.03% |
| $\tilde{T}_2$ | +0.00% | –0.03% |
| $\tilde{T}_3$ | +0.01% | –0.05% |
| $\tilde{T}_4$ | +0.01% | –0.07% |

**25 (!)**
Sample points

# Example of use: System reliability

- **Composite system** with three (stochastically) **independent components**:




- Exponential **life distributions**: $X_i \sim \mathrm{Exp}(\beta_i)$ → **Reliability**: $R_{X_i}(t) = \exp(-\frac{t}{\beta_i})$

- **System reliability** (= probability that the system "survives" until time $t$):

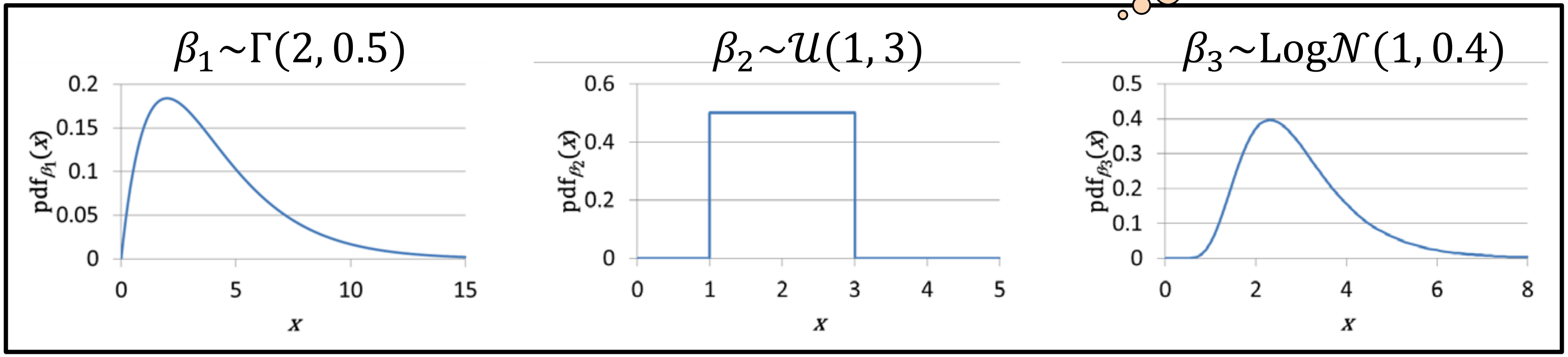
$$R_{\mathrm{sys}}(t) = R_{X_1}(t)R_{X_2}(t) + R_{X_3}(t) - R_{X_1}(t)\,R_{X_2}(t)\,R_{X_3}(t)$$

# Example of use: System reliability

$$R_{\mathrm{sys}}(t) = \exp\left(-\frac{t}{\beta_1} - \frac{t}{\beta_2}\right) + \exp\left(-\frac{t}{\beta_3}\right) - \exp\left(-\frac{t}{\beta_1} - \frac{t}{\beta_2} - \frac{t}{\beta_3}\right)$$

- Parameters $\beta_i$ **fixed → Deterministic function**!
- But, $\beta_i$ **uncertain → Random variable** for each $t$!

Exemplary assumptions!



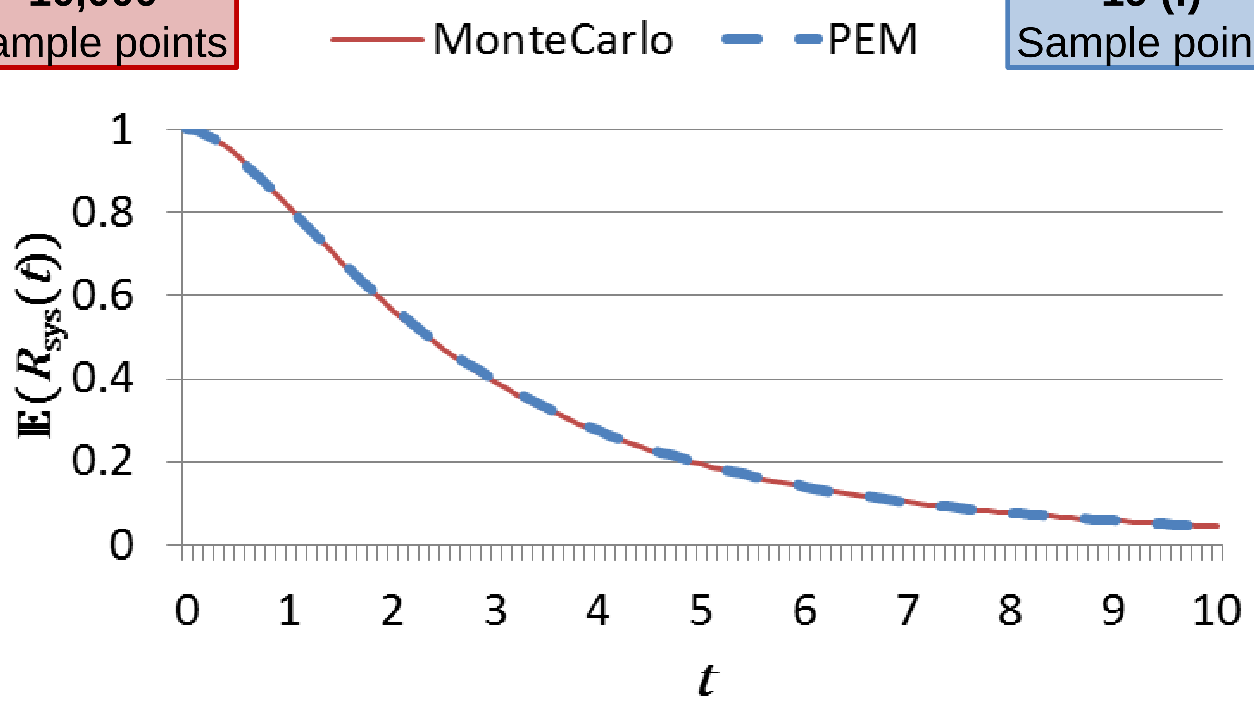

- Apply **PEM** (in comparison to MC) for estimating $\mathbb{E}(R_{\mathrm{sys}}(t))$ and $\mathrm{Var}(R_{\mathrm{sys}}(t))$ depending on $t$

# Example of use: System reliability

**Expected system reliability** based on PEM and MC simulation

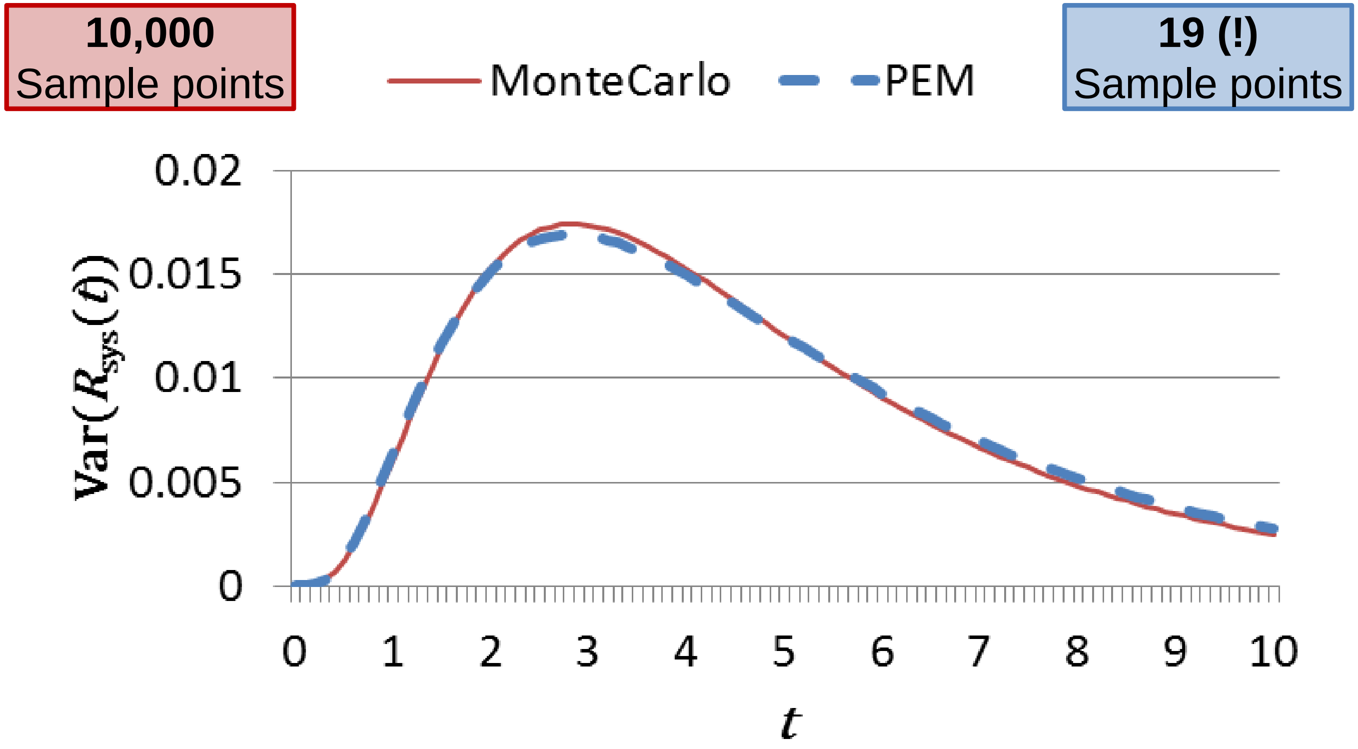
**10,000**
Sample points

**19 (!)**
Sample points



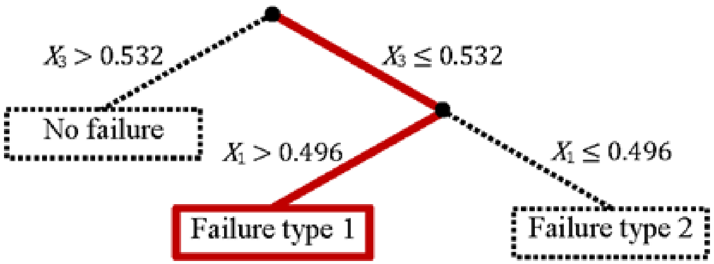
# Example of use: System reliability

**Variance** of the **system reliability** based on PEM and MC simulation

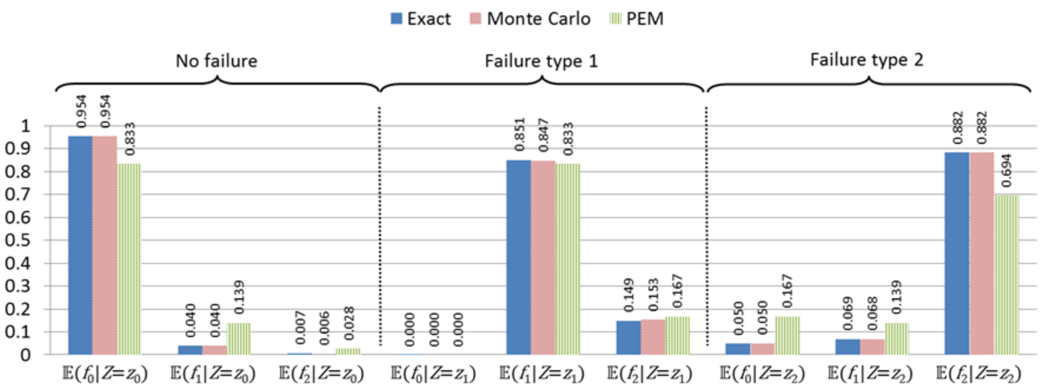


# Example of use: Failure diagnosis using decision trees

- Decision tree:




- Results:




- Further information: see **supplementary material**!

FOR PERSONAL USE ONLY!

**Draft manuscript**
(submitted to Proc IMechE Part O: J Risk and Reliability)

Analyzing uncertainties in model response using the point estimate method: applications from railway asset management

Thorsten Neumann[1]. Beate Dutschk[2]. René Schenkendorf[3].

[1] German Aerospace Center (DLR), Institute of Transportation Systems, Berlin, Germany
[2] Karlsruhe Institute of Technology (KIT), Institute of Industrial Information Technology, Karlsruhe, Germany
[3] Technical University Braunschweig, Institute of Energy and Process Systems Engineering, Braunschweig, Germany

Abstract

Predicting current and future states of rail infrastructure based on existing data and measurements is essential for optimal maintenance and operation of railway systems. Statistical and/or other complex models provide helpful tools for detecting failures and extrapolating current states into the future. This, however, inherently gives rise to uncertainties in the model response that must be analyzed carefully to avoid misleading results and conclusions. Commonly, Monte Carlo (MC) simulations are used for such analyses which often require a large number of sample points to be evaluated for convergence. Moreover, even if quite close to the exact distributions, the MC approach necessarily provides approximate results only. In contrast to that, the present contribution reviews an alternative way of computing important statistical quantities of the model response. That is the so-called point estimate method (PEM) which can be shown to be exact under certain constraints and which usually (i.e., depending on the number of input variables) works with only a few specific sample points. Thus, this method helps to reduce the computational load for model evaluation considerably in the case of complex models or in large-scale applications. Based on three more or less academic examples from the wide field of railway asset management, the performance of the PEM is demonstrated: i) track degradation, ii) reliability analysis of composite systems and iii) failure detection/identification using decision trees. Advantages as well as limitations of the PEM in comparison to common MC simulations are discussed.

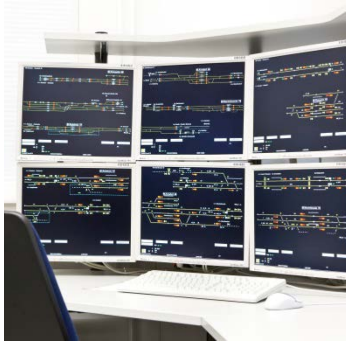
Keywords

Uncertainty propagation analysis, reliability, asset management, prognostics, health management, point estimate method

# Summary: Assets and drawbacks of PEM

- ⇧ **Flexible** approach (**Easy to apply**)
- ⇧ **Reduction** of **computing time** because of small samples (compared to MC)
- ⇧ **Exact results** (for “polynomial models” with given maximum degree)
- ⇧ Applicable to **various types of distributions** (by transforming sample points)
- ⇧ **Reproducible** results (deterministic approach)
- ⇧ Approximate calculation of the **full output distribution possible** by combining PEM with further approaches (e.g., polynomial chaos expansion)

- ⇩ **Basic statistics** of the output distribution **only** (when using the original PEM)
- ⇩ **Stochastically independent** inputs required
- ⇩ **No general guarantee concerning** the **accuracy** of the results (i.e., no convergence as $N \to \infty$)



# Thanks for your attention!

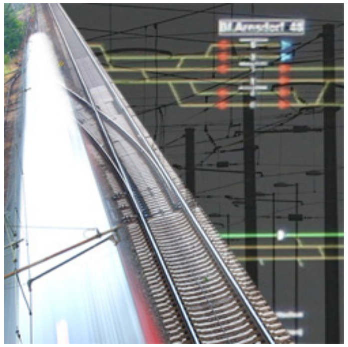
Thorsten.Neumann@DLR.de
+49 (0)30 67055-208

**German Aerospace Center (DLR)**

**Institute of Transportation Systems**

Rutherfordstr. 2
12489 Berlin
Germany

http://www.dlr.de/ts/




DLR