Stochastic modelling of track irregularities

G. Perrin, Denis Duhamel, Christian Soize, C. Fünfschilling

Track Irregularities Stochastic Modelling

G. Perrin[1,2,3], D. Duhamel[1], C. Soize[2], C. Funfschilling[3]
[1] University of Paris-Est,
Navier UMR 8205 CNRS, Marne-la-Vallée, France
[2]University of Paris-Est,
MSME UMR 8208 CNRS, Marne-la-Vallée, France
[3]Innovation and Research Department,
SNCF, Paris, France

Abstract

This paper presents a methodology to build representative railway track geometries thanks to a stochastic modelling. This modelling, which has to integrate the statistical and spatial variabilities and dependencies, is a key issue when using simulation for conception, maintenance or certification purposes, as the dynamic behaviour of the trains is mainly induced by the track geometry. The stochastic process theory is used, combining Karhunen-Loève and polynomial expansions. Through a practical example, this paper finally shows to what extent this methodology gives rise to new promising opportunities for the track geometry maintenance.

Keywords: stochastic modelling, track geometry, random process, polynomial chaos expansion.

1 Introduction

The dynamic behaviour of the trains is mainly induced by the track geometry.

Made up of straight lines and curves at its construction, the track geometry is gradually damaged and regularly subjected to maintenance operations during its lifecycle. The appearing irregularities may be different from one track to an other, from one country to an other, depending on the physical properties of the track substructures, on the traffic conditions (number, type of trains) and on the geographical locations (which can be correlated with weather conditions).

Hence, the train may be confronted to very different running conditions. In security or certification prospects, the dynamic behaviour has therefore to be analysed not only on a few track portions but on this whole realm of possibilities.

In reply to this concern, the measurement train IRIS 320 has been running continuously since 2007 over the French railway network, measuring and recording the track geometry of the main national lines. This measurement of the railway network quality and variability may then be implemented as an input in classical railway softwares to characterise the dynamic behaviour of each train that runs on the considered network. However it would be too time-consuming to simulate runs on the whole railway network, and it is difficult to find a portion of track that is representative of the network.

The work presented here therefore aims at building representative track geometries thanks to a stochastic modelling.

In this paper a parameterisation of the track geometry is presented at first. The formulation of the stochastic modelling is then described. The irregularity vector (gathering the four types of irregularity) is considered as a random field. According to the Karhunen-Loève expansion theory, the irregularity vector is then projected on a determinist orthonormal basis. At last, the projection coefficients, which are random values, are expanded on a polynomial basis.

Finally, the track stochastic modelling, which integrates the statistical and spatial variabilities and dependencies, allows to generate numerically, from a set of track measurements, as many realistic and representative portions of track as needed. These latter can be used in any determinist railway dynamic code to characterise the dynamic behaviour of the train. It could thus bring innovative technical answer to introduce numerical methods and treatments in the maintenance and certification processes.

2 Track irregularities modelling

In this part is formulated the track irregularities modelling.

2.1 Parameterising the track geometry

Let $\mathcal{R}_0 = (\boldsymbol{O}, \mathbf{X}_0, \mathbf{Y}_0, \mathbf{Z}_0)$ be the inertial reference frame. A railway track $\mathcal{T}$ is built up of two rails, which can be modelled as two parallel curves $R_l = \{\boldsymbol{M}_l(s), s \in [0, S]\}$ and $R_r = \{\boldsymbol{M}_r(s), s \in [0, S]\}$, where s is the curvilinear abscissa of the track of length S:

$$\mathcal{T} = R_l \times R_r. \tag{1}$$

Let's call E the rail gauge, and $C_m = \{\boldsymbol{O}_m(s), s \in [0, S]\}$ the track mean line so that:

$$\forall s \in [0, S]\,, \quad \boldsymbol{M}_r\boldsymbol{O}_m(s) = \frac{1}{2}\boldsymbol{M}_r\boldsymbol{M}_l(s). \tag{2}$$

The Frenet frame $(\boldsymbol{O}_m(s), \mathbf{T}(s), \mathbf{N}(s), \mathbf{B}(s))$ is also introduced as:

$$\boldsymbol{T}(s) = \frac{d\boldsymbol{OO}_m}{ds}(s), \tag{3}$$

$$\boldsymbol{N}(s) = \frac{\boldsymbol{M}_r\boldsymbol{M}_l}{\|\boldsymbol{M}_r\boldsymbol{M}_l\|}, \tag{4}$$

$$\boldsymbol{B}(s) = \boldsymbol{T}(s) \wedge \boldsymbol{N}(s). \tag{5}$$

The curvilinear inclination angle $\theta(s)$ is therefore defined as the angle between $\boldsymbol{N}(s)$ and the horizontal plane $(\boldsymbol{X}_0, \boldsymbol{Y}_0)$.

The irregularities appearing during the track lifecycle are of four types: vertical and horizontal alignment irregularities on the one hand, gauge and cross level irregularities on the other hand. These irregularities are characterised by a short wavelength evolution (between 3 and 150 meters) whereas the geometry of new tracks is characterised by long wavelengths.

Hence, a curvilinear parameterisation which suits this double scale property is proposed in this paper. The geometry of new tracks is characterised by the horizontal and vertical curvatures $c_H(s)$ and $c_V(s)$ and the cross level $c_L(s)$, which only depends on s, whereas four curvilinear fields are defined to represent the former described four track irregularities:

- $\alpha(s)$ and $\beta(s)$ for the horizontal and vertical alignment irregularities;
- $\delta(s)$ and $\epsilon(s)$ for the cross level and gauge irregularities.

It can be deduced that:

$$\boldsymbol{OM}_l = \boldsymbol{OO}_m + \Delta(s)\boldsymbol{N}\left(\theta(s)\right), \tag{6}$$

$$\boldsymbol{OM}_r = \boldsymbol{OO}_m - \Delta(s)\boldsymbol{N}\left(\theta(s)\right), \tag{7}$$

$$\Delta(s) = \frac{E + \epsilon(s)}{2}, \tag{8}$$

$$\sin\left(\theta(s)\right) = \frac{c_L(s) + \delta(s)}{E + \epsilon(s)}, \tag{9}$$

$$\mathbf{OO}_m(s) = \mathbf{OO}_{NT}(s) + \alpha(s)\mathbf{N}\left(\theta(s)\right) + \beta(s)\mathbf{B}\left(\theta(s)\right), \tag{10}$$

$$\mathbf{OO}_{NT}(s) = \boldsymbol{\mathcal{F}}\left(c_H, c_V, s\right), \tag{11}$$

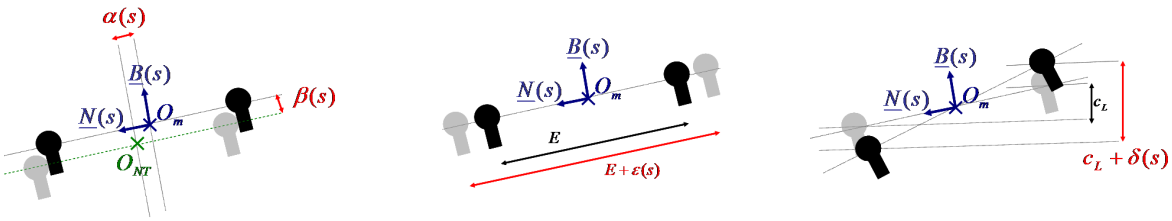

Figure 1: Parameterisation of the track irregularities

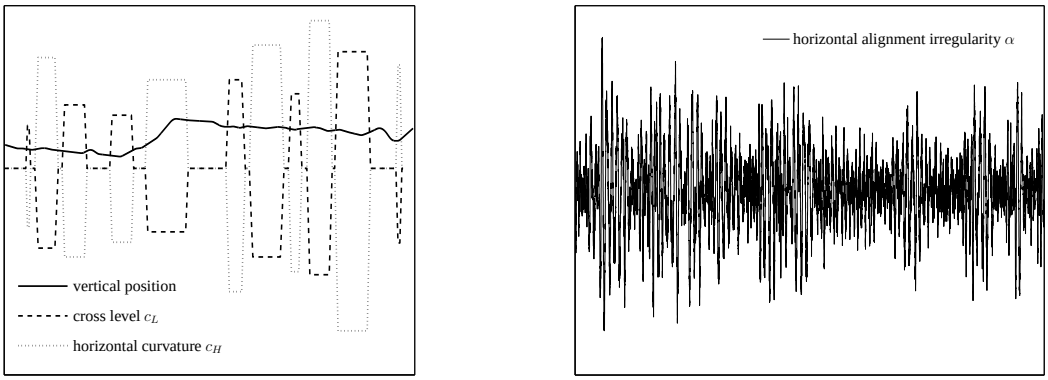

Figure 2: Long wavelengths track parameterisation (left), short wavelengths track parameterisation (right)

where $\mathbf{OO}_{NT}(s)$ is the mean line description of the new track, without any irregularities. In Figure 1 is represented the chosen parameterisation of the track, whereas in Figure 2 are plotted experimental measurements of this parameterisation.

2.2 Definition of the stochastic model

As it has been presented in introduction, a track $\mathcal{T}$ of length S contains several sources of uncertainty, that we decide to gather in an unknown random vector $\boldsymbol{\Xi}$. Consequently, $\mathcal{T}$, which depends on $\boldsymbol{\Xi}$, may be seen as a random field.

The goal of this paper is therefore to model the specific link between $\boldsymbol{\Xi}$ and $\mathcal{T}$ as:

$$\mathcal{M} : \boldsymbol{\Xi} \mapsto \mathcal{T} = \mathcal{M}(\boldsymbol{\Xi}) . \tag{12}$$

The length S can be seen as a description window of the model, and must therefore be chosen carefully. Indeed, it has to be large enough to take into account all the track irregularities wavelengths and correlations, without involving too many computational costs. The length S is thus derived from experimental measures analysis.

Let's be aware that two kinds of dependencies have to be taken into account in the model $\mathcal{M}$:

- the spatial dependencies, which characterise how two irregularities at two different abscissa of the same track are related;
- the statistical dependencies, which describe the link between two irregularities of two different tracks at the same abscissa.

The following methodology allows to well-distinguish these two kinds of dependencies, and therefore to facilitate the interpretation of the physics of the geometry.

2.3 Theoretical frame of the modelling

2.3.1 Notations

Let (Θ, A, P) be a probability space. Let $L_P^2(\Theta, \mathbb{R}^4)$ be the space of all the second-order random variables defined on (Θ, A, P) with values in $\mathbb{R}^4$, equipped with the inner product $\langle ., . \rangle$:

$$\langle \boldsymbol{U}, \boldsymbol{V} \rangle = \int_\Theta \boldsymbol{U}^T(\theta)\boldsymbol{V}(\theta)dP(\theta) = E\left(\boldsymbol{U}^T\boldsymbol{V}\right), \ \ \forall \boldsymbol{U}, \boldsymbol{V} \in L_P^2\left(\Theta, \mathbb{R}^4\right), \qquad (13)$$

where $E(.)$ is the mathematical expectation.

We consider track irregularities modelled by a second-order $\mathbb{R}^4$-valued stochastic process $\boldsymbol{X} = (\alpha, \beta, \delta, \epsilon)$, indexed by $s \in \Omega = [0, S]$, whose realisations are almost surely in the Hilbert space $L^2(\Omega, \mathbb{R}^4)$ equipped with the inner product $(., .)$:

$$(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \boldsymbol{u}^T(s)\boldsymbol{v}(s)ds, \ \ \forall \boldsymbol{u}, \boldsymbol{v} \in L^2(\Omega, \mathbb{R}^4). \qquad (14)$$

It is assumed that $\boldsymbol{X}$ is mean-square continuous.

It has to be noticed that gathering all the irregularities in the same vector, and adopting a vectorial approach certifies that the inner dependencies between different irregularity fields are accurately taken into account.

2.3.2 Karhunen-Loève expansion

Let $[R_{\boldsymbol{XX}}]$ be the autocorrelation matrix of the random field $\boldsymbol{X}$:

$$[R_{\boldsymbol{XX}}] : \ (s, s') \in \Omega^2 \mapsto [R_{\boldsymbol{XX}}(s, s')] = E\left(\boldsymbol{X}(s)\boldsymbol{X}^T(s')\right). \qquad (15)$$

Under the asumptions above, $[R_{\boldsymbol{XX}}]$ is continuous on $\Omega \times \Omega$ and can be written as:

$$[R_{\boldsymbol{X}\boldsymbol{X}}(s,s')] = \sum_{k\geq 1} \lambda_k \boldsymbol{u}^k(s)\ {\boldsymbol{u}^k}^T(s'), \tag{16}$$

where $\left(\lambda_k, \boldsymbol{u}^k\right)$ is an orthonormal basis of $L^2(\Omega, \mathbb{R}^4)$ solution of the Fredholm equation (see [1] for further details):

$$\forall k \geq 1, \quad \int_\Omega [R_{\boldsymbol{X}\boldsymbol{X}}(s,s')]\boldsymbol{u}^k(s')ds' = \lambda_k \boldsymbol{u}^k(s). \tag{17}$$

The eigenvalues λ_k are non-negative, and can be arranged in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \to 0. \tag{18}$$

The Karhunen-Loève expansion of the stochastic process $\boldsymbol{X}$ is then:

$$\boldsymbol{X}(s) = \sum_{k\geq 1} \sqrt{\lambda_k}\boldsymbol{u}^k(s)\eta_k, \tag{19}$$

with:

$$\eta_k = \frac{1}{\sqrt{\lambda_k}} \left(\boldsymbol{X}, \boldsymbol{u}^k\right). \tag{20}$$

Equations (15), (16), (19) imply:

$$E\left(\eta_k \eta_l\right) = \delta_{kl.} \tag{21}$$

Based on the eigenvalues decrease, $\boldsymbol{X}$ can thus be approximated as:

$$\boldsymbol{X}(s) \approx \sum_{k=1}^{N_x} \sqrt{\lambda_k}\boldsymbol{u}^k(s)\eta_k, \tag{22}$$

where N_x is related to a chosen value of the normalized mean-square error:

$$\epsilon_{KL}^2 = \frac{\sum_{k>N_x} \lambda_k}{\sum_{k\geq 1} \lambda_k}. \tag{23}$$

Under a matricial form, $\boldsymbol{X}$ can be rewritten as:

$$\boldsymbol{X}(s) = [Q(s)]\boldsymbol{\eta}, \tag{24}$$

with:

$$\boldsymbol{\eta} = \left(\eta_1, ..., \eta_{N_x}\right), \tag{25}$$

$$[Q(s)] = [u(s)][\lambda^{1/2}], \tag{26}$$

$$[u] = \left[\begin{array}{cccc} \boldsymbol{u}^1 & \boldsymbol{u}^2 & \cdots & \boldsymbol{u}^{N_x} \end{array} \right], \tag{27}$$

$$[\lambda^{1/2}] = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \sqrt{\lambda_{N_x}} \end{pmatrix}. \tag{28}$$

The condition (21) can then be rewritten as:

$$E\left(\boldsymbol{\eta}\boldsymbol{\eta}^T\right) = [I_{N_x}], \tag{29}$$

where $[I_{N_x}]$ is the N_x-dimension unitary matrix.

Let's note that the more correlated the process is, the smaller the needed number of terms to achieve a given error is. Moreover, thanks to the Karhunen-Loève expansion, spatial and statistical correlation are clearly separated. Whereas $[u]$ emphasizes the predominant track irregularity spatial shapes, $\boldsymbol{\eta}$ gathers all the statistical variability. In order to fully describe $\mathcal{M}$, the statistical content of $\boldsymbol{\eta}$, and more specially its joint probability density function (PDF) p_η has to be focused on.

2.4 Computation of the joint PDF of $\boldsymbol{\eta}$

2.4.1 Gathering of ν^{exp} realisations of $\boldsymbol{\eta}$

As it has been showed in introduction, the measurement train IRIS 320 has recently given access to huge data bases, and motivated statistical analysis of the network. In the following, it is considered that all the measurements at a given period can be seen as a set $\mathcal{S}^{exp}$ of ν^{exp} track portions of same length S:

$$\mathcal{S}^{exp} = \left\{\mathcal{T}^{exp,i},\ 1 \leq i \leq \nu^{exp}\right\}. \tag{30}$$

In reference to the stochastic description given in section 2.2, each element $\mathcal{T}^{exp,i}$ in $\mathcal{S}^{exp}$ is then regarded as an independent realisation of the random process $\mathcal{T}$. More precisely, the corresponding set $\mathcal{X}^{exp} = \{\boldsymbol{x}^{exp,i},\ 1 \leq i \leq \nu^{exp}\}$ is defined such that $\boldsymbol{x}^{exp,i}$ refers to the measured track irregularities of $\mathcal{T}^{exp,i}$, and is such that $\boldsymbol{x}^{exp,1}$,...,$\boldsymbol{x}^{exp,\nu^{exp}}$ are ν^{exp} independant realisations of random field $\boldsymbol{X}$.

Consequently, the autocorrelation matrix as well as ν^{exp} realisations $\{\boldsymbol{\eta}^{exp,i},\ 1 \leq i \leq \nu^{exp}\}$ of random vector $\boldsymbol{\eta}$ can be assessed as:

$$[R_{\boldsymbol{X}\boldsymbol{X}}(s,s')] \approx \frac{1}{\nu^{exp}} \sum_{i=1}^{\nu^{exp}} \boldsymbol{x}^{exp,i}(s)\boldsymbol{x}^{exp,i}(s')^T, \tag{31}$$

$$\forall 1 \leq i \leq \nu^{exp},\ \forall 1 \leq k \leq N_x,\ \ \eta_k^{exp,i} = \frac{\left(\boldsymbol{x}^{exp,i}, \boldsymbol{u}^k\right)}{\sqrt{\lambda_k}}. \tag{32}$$

2.4.2 Polynomial chaos expansion of random vector $\boldsymbol{\eta}$

From equation (21), random variables η_1,...,η_{N_x} are orthogonal, but are generally not independent. Hence, the joint PDF p_η is difficult to characterise in high dimension (N_x large) using non-parametric statistical estimators from ν^{exp} realisations of the vector $\boldsymbol{\eta}$. The maximum entropy principle, discussed in [2], has been introduced to build *a priori* stochastic model, by focusing on the only usable and available information. Very relevant when dealing with small dimension problems with even small avaiblable data, this method is limited in high dimension. More recently, polynomial chaos expansion methods have underlined very promising results in high-dimension (see [3]). Based on the projections of the considered random vector $\boldsymbol{\eta}$ on known and chosen orthonormal polynomial basis $\{\psi_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathcal{A}_p\}$, these methods aim at building a chaos representation that makes explicit the link between the vector of interest and an other random vector $\boldsymbol{\xi}$ of known dimension and joint PDF:

$$\boldsymbol{\eta} \approx \boldsymbol{\eta}^{chaos}(N) = \sum_{\boldsymbol{\alpha} \in \boldsymbol{\mathcal{A}}_p} \boldsymbol{y}^{(\boldsymbol{\alpha})} \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{33}$$

$$\boldsymbol{\xi} \mapsto \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = H_{\alpha_1}(\xi_1) \otimes ... \otimes H_{\alpha_{N_g}}(\xi_{N_g}), \tag{34}$$

$$\boldsymbol{\mathcal{A}}_p = \left\{ \boldsymbol{\alpha} = \left(\alpha_1, ..., \alpha_{N_g}\right) \mid |\boldsymbol{\alpha}| = \sum_{i=1}^{N_g} \alpha_i \leq p \right\}, \tag{35}$$

where $\boldsymbol{\xi}$ is a N_g-dimensional normalized gaussian random vector, $N = (N_g + p)!/\left(N_g! p!\right)$ is the dimension of $\boldsymbol{\mathcal{A}}_p$, and $x \mapsto H_{\alpha_p}(x)$ is the normalized Hermite polynomial of degree α_p. By carrying out a different index ordering, equation (33) can be rewritten as:

$$\boldsymbol{\eta} \approx \sum_{j=1}^{N} \boldsymbol{y}^j \psi_j(\boldsymbol{\xi}) = [y] \boldsymbol{\Psi}(\boldsymbol{\xi}), \tag{36}$$

in which:

$$[y] = \left[\boldsymbol{y}^1\ \cdots\ \boldsymbol{y}^N\right]. \tag{37}$$

It can be noticed that condition (21) now implies that:

$$[y] \in \tilde{\mathcal{O}} = \left\{ [b] \in \mathbb{M}_{N_x,N}(\mathbb{R}) \mid [b][b]^T = [I_{N_x}] \right\}. \tag{38}$$

2.4.3 Identification of the polynomial chaos expansion coefficients

Based on the maximum likelihood principle, and the ν^{exp} independant realisations $\left\{\boldsymbol{\eta}^{exp,i},\ 1 \leq i \leq \nu^{exp}\right\}$ of $\boldsymbol{\eta}$, the optimal polynomial chaos expansion coefficients gathered in matrix $[y^{opt}]$ may be found as the result of a maximization problem (see [4] for further details):

$$[y^{opt}] = \arg\max_{[y]\in\tilde{\mathcal{O}}} \mathcal{L}\left([y]\right), \tag{39}$$

where $\mathcal{L}$ is the log-likelihood:

$$\mathcal{L}\left([y]\right) = \sum_{i=1}^{\nu^{exp}} \ln p_{\eta^{chaos}(N)}\left(\boldsymbol{\eta}^{exp,i}, [y]\right). \tag{40}$$

Using the kernel estimation method, the PDF $p_{\eta^{chaos}(N)}$ of $\boldsymbol{\eta}^{chaos}(N)$ can be estimated from ν^{chaos} independent realisations $\left\{\boldsymbol{\xi}(\theta_p),\ 1 \leq p \leq \nu^{chaos}\right\}$ of $\boldsymbol{\xi}$ for any $[y]$ in $\tilde{\mathcal{O}}$ and $\boldsymbol{x}$ in $\mathbb{R}^{N_x}$:

$$p_{\eta^{chaos}(N)}\left(\boldsymbol{x}, [y]\right) \approx \frac{1}{(2\pi)^{N_x/2}\,\nu^{chaos}\prod_{k=1}^{N_x} h_k} \sum_{p=1}^{\nu^{chaos}} exp\left(-\frac{1}{2}\sum_{k=1}^{N_x}\left(\frac{x_k - \eta_k^{chaos}(N,\theta_p)}{h_k}\right)^2\right), \tag{41}$$

$$\boldsymbol{\eta}^{chaos}(N,\theta_p) = [y]\boldsymbol{\Psi}\left(\boldsymbol{\xi}(\theta_p)\right), \tag{42}$$

where $\boldsymbol{h} = (h_1, \cdots, h_{N_x})$ is the optimal bandwith vector with respect to the AMISE criteria (see [5]) of the Kernel smoothing estimation of $p_{\eta^{chaos}(N)}$. Equations (40) and (41) yield:

$$\mathcal{L}\left([y]\right) \approx \hat{\mathcal{L}}\left([y]\right) = \hat{\mathcal{L}}_C + \hat{\mathcal{L}}_V, \tag{43}$$

$$\hat{\mathcal{L}}_C = -\nu^{exp} \ln\left((2\pi)^{N_x/2}\,\nu^{chaos}\prod_{k=1}^{N_x} h_k\right), \tag{44}$$

$$\hat{\mathcal{L}}_V = \sum_{i=1}^{\nu^{exp}} \ln\left(\sum_{p=1}^{\nu^{chaos}} \exp\left(-\frac{1}{2}\sum_{k=1}^{N_x}\left(\frac{\eta_k^{exp,i} - \eta_k^{chaos}(N,\theta_p)}{h_k}\right)^2\right)\right). \tag{45}$$

As $\hat{\mathcal{L}}$ is non concave, random maximization algorithms are used to estimate $[y^{opt}]$.

2.5 Post-processing of the track modelling

Finally, once $[y^{opt}]$ has been computed, it can be deduced from equations (24) and (36):

$$\forall\,(s,\theta)\in\Omega\times\Theta,\ \boldsymbol{X}(s,\theta)=[Q(s)][y^{opt}]\boldsymbol{\Psi}\left(\boldsymbol{\xi}(\theta)\right). \tag{46}$$

By focusing only on the track irregularities, and by identifying respectively $\mathcal{T}$ and Ξ to $\boldsymbol{X}$ and $\boldsymbol{\xi}$, the stochastic model $\mathcal{M}$ of equation (12) becomes:

$$\mathcal{M}:\ \boldsymbol{\xi}\mapsto\mathcal{T}=\mathcal{M}\left(\boldsymbol{\xi}\right)=[Q][y^{opt}]\boldsymbol{\Psi}\left(\boldsymbol{\xi}\right). \tag{47}$$

The parameters N_g and p being chosen thanks to convergence studies, $[Q]$ and $[y^{opt}]$ being computed, any independant realisation of $\boldsymbol{\xi}$ leads to a representative and realistic realisation of the irregularity vector of a potential track portion of length S. Hence, any statistical post-treatment can be carried out on the stochastic model $\mathcal{M}$.

3 Application

In this part, the previously described methodology is applied step by step to the computation of realistic track portions, whose irregularities are representative of the global quality of the network. In a certification prospect, these representative tracks could be used to numerically homologate a train.

This study being confidential, very few normalized values are presented.

3.1 Step 1: computation of the matrix-valued autocorrelation function $[R_{\boldsymbol{XX}}]$

For this study, track irregularity measurements on around $\nu^{exp}=1850$ portions of same length S have been gathered.

From equation (31), $[R_{\boldsymbol{XX}}]$ was evaluated. As an illustration, in Figure 3 is represented $[R_{\boldsymbol{XX}}(s,s')]_{11}$.

3.2 Step 2: Karhunen-Loève expansion

The solutions $(\boldsymbol{u},\lambda)$ of the Fredholm equation (17) were then computed thanks to a Finite Element approach. Indeed, we define $\Omega_h=\{s_1=0,s_2,...,s_{N_S}=S\}$ and eigenvectors $\boldsymbol{u}$ are projected on linear shape functions:

$$\boldsymbol{u}(s)=\sum_{n=1}^{N_S}\boldsymbol{b}^n h_n(s)=[H]\boldsymbol{d}(s), \tag{48}$$

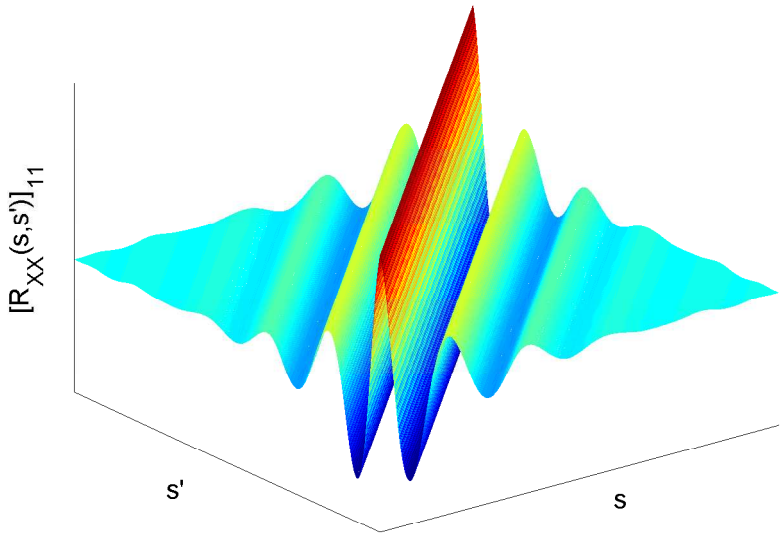

Figure 3: Representation of $(s, s') \mapsto [R_{\boldsymbol{XX}}(s, s')]_{11}$

$$\boldsymbol{d}(s) = \left(b_1^1, b_1^2, ..., b_1^{N_S}, b_2^1, ..., b_4^{N_S}\right), \tag{49}$$

$$[H] = \begin{bmatrix} h_1 & \dots & h_{N_S} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & h_1 & \dots & h_{N_S} & 0 & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 0 & h_1 & \dots & h_{N_S} & 0 & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & h_1 & \dots & h_{N_S} \end{bmatrix}, \tag{50}$$

where $\{s \mapsto h_n(s), 1 \leq n \leq N_S\}$ are unidimensional linear shape functions. Solutions of (17) are finally looked for from the classical eigenvalue problem:

$$([K] - \lambda[M])\, \boldsymbol{D} = \boldsymbol{0}, \tag{51}$$

where:

$$[K] = \int_\Omega \int_\Omega [H(s)]^T [R_{\boldsymbol{XX}}(s, s')][H(s')] ds' ds, \tag{52}$$

$$[M] = \int_\Omega [H(s)]^T [H(s)] ds. \tag{53}$$

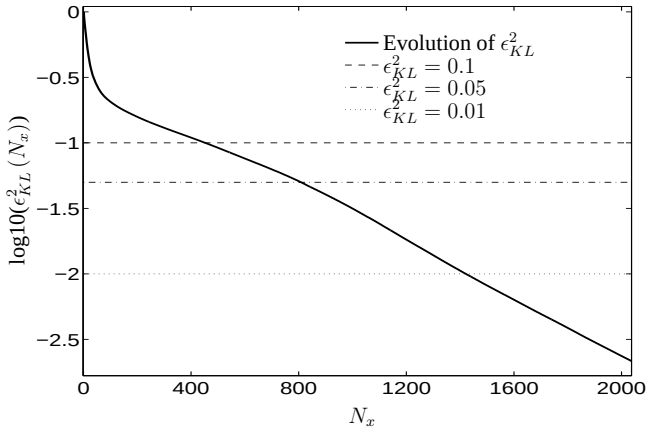

Figure 4: Characterisation of the Mean-Square Error ϵ^2_{KL} with respect to the truncation parameter N_x

Given acceptable values of truncation for the mean-square error ϵ^2_{KL} of equation (23) (whose evolution is represented in Figure 4), the truncation parameter N_x of equation (22) is assessed:

- $\epsilon^2_{KL} = 0.1 \leftrightarrow N_x = 452$;
- $\epsilon^2_{KL} = 0.05 \leftrightarrow N_x = 807$;
- $\epsilon^2_{KL} = 0.01 \leftrightarrow N_x = 1423$.

N_x can be directly compared to the truncation value $4N_S \times 4N_S$ due to the finite element approximation. We verify the inequality $N_x << 4N_S \times 4N_S$, which justifies the importance of the Karhunen-Loève expansion in term of reduction efficiency.

From equation (32), the ν^{exp} realisations $\{\boldsymbol{\eta}^{exp,i},\ 1 \leq i \leq \nu^{exp}\}$ of $\boldsymbol{\eta}$ are computed. The PDF of $\boldsymbol{\eta}$ can thus be estimated and analysed. For instance, in Figure 5, kernel smoothing estimations of the PDF of η_1, η_2 and η_3 are compared to the normal distribution. Marginal distributions of $\boldsymbol{\eta}$ being non-gaussian, the random process $\boldsymbol{X}$ is non gaussian. The joint PDF of $\boldsymbol{\eta}$ needs therefore to be properly characterised.

3.3 Step 3: polynomial chaos expansion

In agreement with the mathematical frame of section 2.4, $\boldsymbol{\eta}$ is expanded on a known polynomial basis of parameters N_g and p:

$$\boldsymbol{\eta} \approx \boldsymbol{\eta}^{chaos}(N_g, p) = \sum_{j=1}^{(N_g+p)!/(N_g!p!)} \boldsymbol{y}^{j,opt}\Psi_j\left(\xi_1, ..., \xi_{N_g}\right) = [y^{opt}]\boldsymbol{\Psi}\left(\xi_1, ..., \xi_{N_g}\right). \tag{54}$$

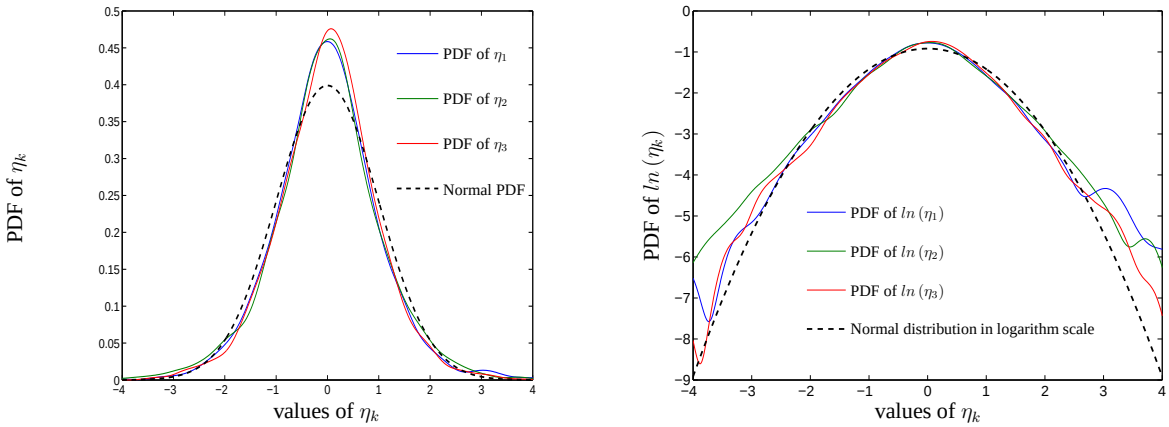

Figure 5: Estimation of the PDF of three elements of $\boldsymbol{\eta}$.

The values of N_g et p stems from convergence analysis. As described in [4], for each component $\eta_k^{chaos}(N_g, p)$, the L^1-log error function err_k is introduced:

$$\forall 1 \leq k \leq N_x, \;\; err_k(N_g, p) = \int_{BI_k} |log_{10}\left(p_{\eta_k^{exp}}(x)\right) - log_{10}\left(p_{\eta_k^{chaos}}(x)\right)| dx, \quad (55)$$

where:

- BI_k is the bounded domain which is adapted to the values of η_k^{exp};
- $p_{\eta_k^{exp}}$ and $p_{\eta_k^{chaos}}$ are the PDF of η_k^{exp} and η_k^{chaos} respectively.

For instance, evolution of $err_1(N_g, p)$ is represented in Figure 6. It can be noticed that the choice of (N_g, p) is not easy: the higher the values of (N_g, p) are, the more complex the polynomial basis is, the more accurate the projection should be, but unfortunately the more difficult and less precise the identification is.

The multidimensional error function $err(N_g, p)$ is deduced from the unidimensional L^1-log error function to evaluate the final N_g and p:

$$err(N_g, p) = \sum_{k=1}^{N_x} err_k(N_g, p) \quad (56)$$

3.4 Step 4: realisation of representative track irregularities

Once the polynomial projection matrix $[y^{opt}]$ of equation (54) has been computed, equation (47) allows to generate a representative track geometry from any realisation of $\boldsymbol{\xi} = \left(\xi_1, \cdots, \xi_{N_g}\right)$. In Figure 7 are plotted two realisations of the stochastic process $\boldsymbol{X}$. In order to be clearer, the graphs of each component of $\boldsymbol{X}$, whose mean values are equal to zero, have been translated.

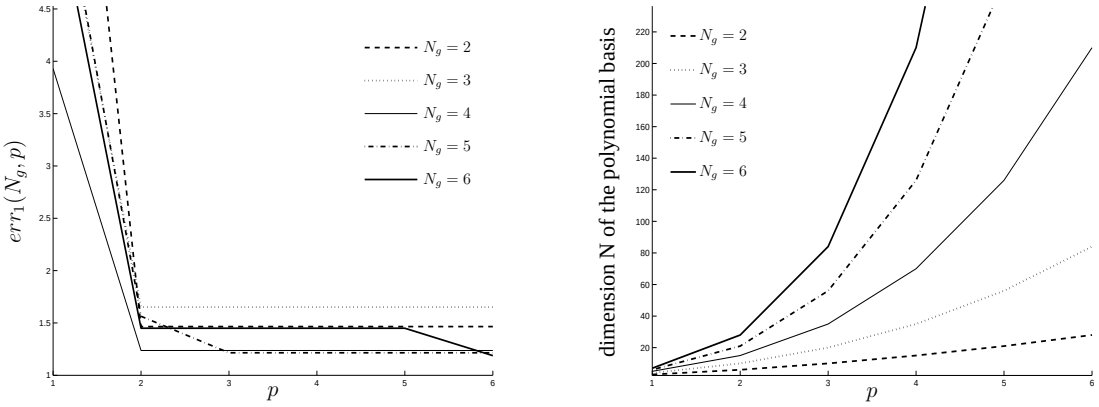

Figure 6: Graphs of $p \mapsto err_1(N_g, p)$ (left); increase of the polynomial basis dimension with respect to N_g and p.

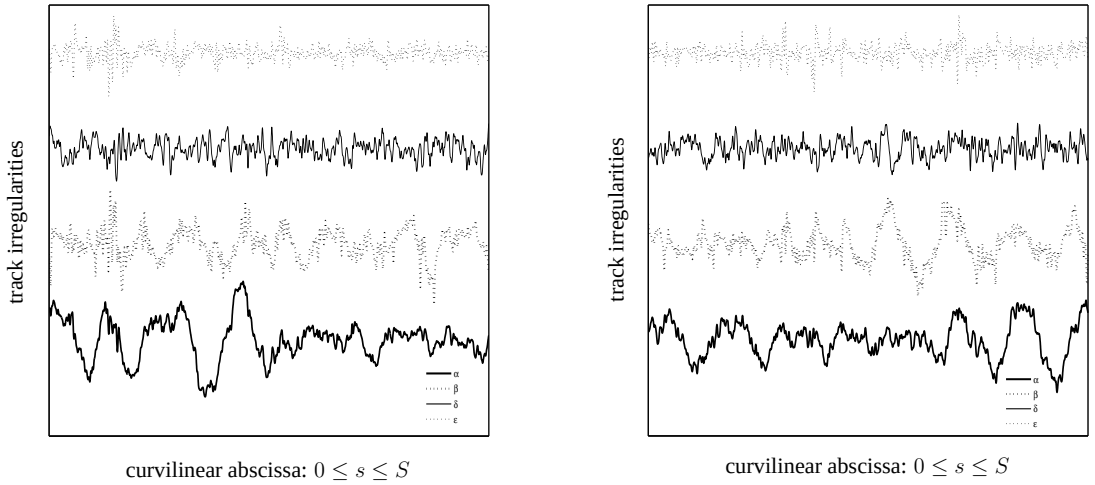

Figure 7: Two particular realisations of the track irregularities process.

4 Conclusions and prospects

At a time when the numerical power and the mechanical simulation algorithms precision keep increasing, the introduction of the simulation in the railway maintenance and certification would represent an important progress. The numerical characterisation of the track geometry is therefore bound to play a key role in this evolution.

From a sample of track measurements, a complete methodology to generate realistic and representative track geometries has been described in this paper.

Coupled with any railway software without requiring an access to the sources codes, these track geometries should allow to characterise the dynamic behaviour without simulating runs on the whole network.

At last, the influence of the evolution of the vertical and horizontal curvatures on the track irregularities could enrich this study.

References

[1] O.P. Le Maître and O.M. Knio. *Spectral Methods for Uncertainty Quantification*. Springer, 2010.

[2] C. Soize. Maximum entropy approach for modeling random uncertainties in transient elastodynamics. *Journal of the Acoustical Society of America*, 109:1979–1996, 2001.

[3] R. Ghanem and P.D. Spanos. Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics*, Transactions of teh ASME 57:197–202, 1990.

[4] C. Soize. Identification of high-dimension polynomial chaos expansions with random coefficients for non-gausian tensor-valued random fields using partial and limited experimental data. *Computer Methods in Applied Mechanics and Engineering*, 199:2150–2164, 2010.

[5] A. Dutfoy. Reference guide. Tutorial, Open TURNS version 0.11.3, 2008.