# Optimal design of rail level crossings and associated transition zones using adaptive surrogate-assisted optimization

Shang, Yue; Nogal, Maria; Teixeira, Rui; Wolfert, A. R.(Rogier) M.

**DOI**
10.1016/j.engstruct.2023.115740

**Publication date**
2023

**Document Version**
Final published version

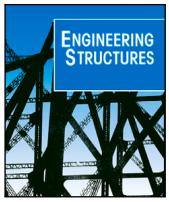
**Published in**
Engineering Structures

Engineering Structures 282 (2023) 115740

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct



# Optimal design of rail level crossings and associated transition zones using adaptive surrogate-assisted optimization ☆

Yue Shang [a,*], Maria Nogal [a], Rui Teixeira [b], A.R. (Rogier) M. Wolfert [a]

[a] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, The Netherlands
[b] School of Civil Engineering, University College Dublin, Richview Newstead, Belfield, Dublin 4, Ireland

ARTICLE INFO

*Keywords:*
Performance-based design
Track design optimization
Vehicle–track coupled dynamics
Railway transition zone
Level crossing
Surrogate-assisted optimization

ABSTRACT

Transition zones such as level crossing and bridge approaches are critical links in railway networks due to higher degradation rates and maintenance needs. In this context, parametric optimization has been applied to improve the design in transition zones; however, it requires a more computationally efficient tool to support repetitive function evaluations, since the involved vehicle–track dynamic simulations are becoming more expensive to evaluate. For this purpose, a surrogate-based simulation methodology is proposed to search for an optimal combination of parameters relevant to the geometry and elasticity of track structures. Specifically, the presented methodology integrates finite element (FE)-based modeling with surrogate-assisted optimization: (1) the FE model is developed to characterize the dynamic behavior of a level crossing under a moving vehicle; (2) the optimization problem is formulated upon this mechanical model by extending the expensive FE simulations to an adaptive surrogate modeling scheme. This integration facilitates efficient exploration of the track design space (thereby reducing the computational cost), and a reasonable balance can be achieved between solution quality and computational effort. The methodology is applied to a Dutch railway case. Results show that compared to a reference design, the optimized design significantly improves performance indicators relevant to wheel–rail contact forces and energy dissipation in the ballast layer. The solution brings great potential in achieving a more desirable vehicle–track interaction and improving the connecting performance between level crossings and transitions. The methodology is applicable to other railway structures and may also contribute to improvements in current track design practices.

## 1. Introduction

Transition zones near rigid structures (e.g., bridges and level crossings) are weak spots in a railway network due to higher degradation rates and maintenance demands. They often exhibit non-consistent track configurations (e.g., connections between ballast track and slab track) and variations in geotechnical foundations (e.g., embankment to a bridge). The first type is typically encountered in many level crossings, where a slab track, e.g., embedded rail system (ERS), is placed in the crossing area, and the ballast track forms the adjacent section (see Fig. 1). In ERS, the ballast is replaced by concrete slabs with channels, where the rails are placed and fixated by an elastic poured compound. The slabs provide an obstacle-free surface with the road pavement for crossing traffic. The elastic compound provides homogeneous continual support to the rail, differing from periodic sleeper support in the traditional ballast track. With design benefits such as noise reduction and savings in construction height and weight [1], ERS also has wide applications in bridges, tunnels, and tramlines [2–4].

Despite the design advantages, the ERS applied in those structures exhibits structural discontinuities with the adjacent ballast track, which are often characterized by (longitudinal) variations of track support stiffness. As a train passes through a transition zone, stiffness variations produce uneven track deflections under wheels, which disturbs the wheel–rail interface and amplifies the dynamic response. Since the ballast track (especially the ballast layer made up of granular materials) is susceptible to settlement while the slab track is designed for minimal settlement, the additional dynamic forces are associated with a change in the vertical wheel position, which over the load cycles will result in accumulation of differential settlement, i.e., track geometry degradation [5]. The geometry problem further causes a local disturbance to the wheel–rail interface and accelerates the degradation. Without interventions, the uneven settlement is often associated with the deterioration of track components, e.g., hanging sleepers and rail defects [6,7].

The underlying degradation mechanism has been extensively studied from theoretical and practical perspectives. Some theoretical

☆ This work was supported by the Chinese Scholarship Council, China under Grant [201907720116].
* Corresponding author.
*E-mail address:* y.shang@tudelft.nl (Y. Shang).

https://doi.org/10.1016/j.engstruct.2023.115740
Received 12 September 2022; Received in revised form 24 December 2022; Accepted 29 January 2023
Available online 20 February 2023


**Fig. 1.** An example level crossing with embedded rail system [17].

research characterizes track dynamics in transition zones as transition radiation, referring to the radiation emitted into the track (in the form of waves) when a train approaches and crosses an inhomogeneity in the railway track (i.e., a transition zone) [8]. The first study on transition radiation of elastic waves was carried out in [9], where an infinite string on a piecewise-homogeneous Winkler foundation is subjected to a moving point force/mass. Later in [10], a smooth transition was incorporated to study a smooth change of foundation stiffness on transient vibrations of a string. To account for the flexural rigidity, transition radiation in a beam resting on Winkler foundation was analyzed in [11–13]. More recently, studies in this field have been extended to consider a 2D continuum [14], nonlinear elastoplastic foundation [8], vehicle inertia [15], and sleeper periodicity [16]. In general, results show that the effect of stiffness variations increases with the train speed.

The above works mainly adopt analytical-numerical frequency-domain approaches to study transition radiation, which are computationally efficient and able to provide fundamental insight into the underlying mechanism. Another research stream develops time-domain models with different degrees of complexity to study case-specific transition zones, which are often supplemented by field instrumentation, e.g., bridge approaches [2,18,19] and culverts [20,21]. The time-domain models are mainly developed using the finite element method (FEM), with particular attention on addressing track geometry irregularities [2,22], track–soil interaction [23], and nonlinear behavior of the substructure [24,25]. To account for the effect of vehicle–track interactions (VTI), these works mainly model the moving vehicle as a multibody system instead of moving point force/mass in most theoretical research.

Various mitigation measures have been investigated and applied in transition zones to reduce the local dynamic amplification. As for the ballast-slab transition, efforts have been made to evaluate, e.g., optimal spacing between the junction and adjacent sleeper [26], varying baseplate pad stiffness and baseplate weight [27], and stepwise change of railpad and subgrade properties [28]. A comprehensive review of countermeasures for transition zones can be found in [6], where the guiding principle is smoothing the stiffness variation to achieve a gradual transition.

Those measures are often evaluated by modifying design parameters and repeated numerical simulations. The design parameters are relevant to the geometry and mechanical properties of track components, which form an essential input for the track behavior. The modification is time-consuming, especially when multiple parameters with high variability are involved. It is therefore often achieved by perturbing a single or two parameters at a time while fixing the remaining in a given value. However, a more systematic way of improving the track design is parametric optimization allowing a trade-off between a set of parameters. Combined with the dynamic simulations, the technique can facilitate the exploration of design space and search for an optimal combination of parameters on both sides of the track, e.g., ERS (in level crossings) and ballast track (in transition zones). The solution brings great potential in achieving a more desirable vehicle–track interaction and improving the connecting performance between different track forms.

Past works on railway design optimization have been applied to optimizing rail profiles [29–31] and track stiffness [31,32] in turnouts, railway alignments [33], rail corrugation [34], and vehicle suspension [35]. The most relevant work is provided in [36], where a multi-objective genetic algorithm (GA) was used to search for optimal solutions in a generic ballast-slab transition. From the operational viewpoint, track design optimization problems mostly require dynamic VTI simulations formulated by FEM and or multibody simulation. The computational time for running these models highly depends on the model complexity. One single run may take a few hours or even days. Evolutionary algorithms such as GA often operate with many individuals through many generations and require more than 1000 function evaluations to search for the optimum [37]. The total computational cost of optimizing a VTI model using GA may become intractable when considering, e.g., complex track configurations and nonlinear track behavior. Such conditions may certainly limit their applicability in track design practice. It is therefore more desirable to find a method that can search for better or optimal solutions while keeping the computational time affordable.

Surrogate modeling (also known as metamodeling, response surface methodology) can be an efficient way to solve such problems with expensive functions, since it replaces the original/true function with an approximation that is faster to evaluate. Surrogate models are constructed by querying the original model at given input points, and the data for training a surrogate is generated by evaluating the corresponding function values. Surrogate-assisted optimization (SAO) is more targeted than general surrogate modeling, where an adaptive sampling strategy is employed to refine the surrogate. It is achieved by exploiting the information from the existing surrogate and guiding the search to promising sampling areas in the design space [38]. In this way, the expensive objective function is evaluated only at carefully selected points, which in turn saves the computational budget (i.e., the number of function evaluations).

The present contribution focuses on the formulation and application of a surrogate-based adaptive modeling technique for optimizing geometry and elastic properties in level crossings and associated transitions, since this type of asset has not received much attention in the literature (except for an experimental study in [5]). A finite element (FE) model is developed to characterize the dynamic behavior of the ERS-type level crossings, where a numerically efficient approach for the VTI simulation is proposed combining COMSOL and MATLAB interface. The optimization problem is formulated upon this mechanical model by extending the FE simulations to an adaptive surrogate modeling scheme. This integration facilitates the efficient exploration of design space, and a reasonable balance can be achieved between the solution quality and computational effort. Note that this integrative methodology is applied to the level crossings, where a simplified vehicle–track model is incorporated for demonstration purposes. It is also applicable to other railway track structures and offers the possibility of incorporating a more complex model in terms of, e.g., track configurations, structural elements, modeling dimensions, and track–soil coupling effect.

Several response quantities that capture track dynamic amplification are proposed and compared as design criteria/objectives, which are transient responses generated from the FE model. Those showing higher solution quality and sensitivity to parametric variation are selected to formulate a multi-objective optimization (MOO) problem, which is solved by embedding an achievement scalarizing function (ASF) in the surrogate modeling scheme. The optimized design solutions are obtained by minimizing the proposed objectives. It is considered that if the transition zones are designed to reduce the transient responses (the amplitude) of the vehicle–track system, the risk of potential track degradation can also be reduced. This may further reduce

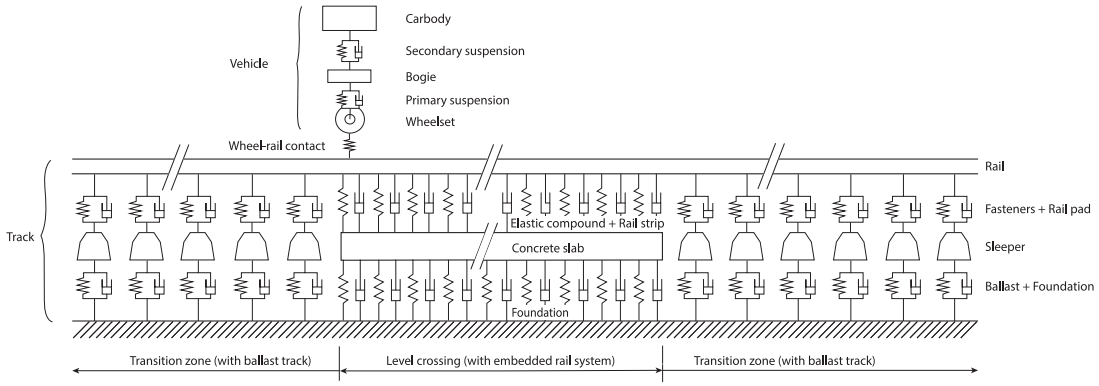



**Fig. 2.** Illustration of the vehicle–track interaction (VTI) model.

maintenance needs and the consequent impact on system life-cycle cost and network performance.

The paper is organized as follows. Section 2 presents four building blocks in the proposed methodology, including a method of simulating the vehicle–track coupling dynamics (Section 2.1), formulation of a general optimization problem for the level crossings (i.e., definition of design variable, objectives, and mathematical formulation) in Section 2.2, a surrogate-assisted optimization scheme (Section 2.3), and a method of integration (Section 2.4). Section 3 presents numerical examples to demonstrate the applicability of the proposed procedure. Section 4 discusses the quality of obtained solutions and effectiveness of the proposed objectives, using a reference design as a benchmark. Section 5 draws some conclusions and future research lines. Some specific mathematical formulations for the vehicle–track dynamic simulation are presented in Appendix.

## 2. Methodology

### 2.1. Modeling of vehicle–track interaction dynamics

The interaction between vehicle and track dynamics is essential, especially in transition zones. The track discontinuity induces an undesirable wheel–rail interaction that amplifies the track and vehicle response. In this case, accounting for the coupling dynamics of the systems can capture the variation in vertical momentum of the moving vehicle and the binding effect on track vibration. For this purpose, a VTI model is developed to simulate the dynamic response of a level crossing induced by a passing vehicle. Fig. 2 shows a schematic representation of the model, which is used as a basis for design alternatives comparison and is represented in a parametric way for optimization purposes. It consists of two subsystems, a vehicle modeled by multibody simulation and a track structure modeled by FEM. The two subsystems are coupled through wheel–rail contact to formulate an integrated time-dependent system. The model characterizes the vertical vehicle–track dynamics since it is typically pronounced in transition zones [6]. Symmetrical load distribution is assumed between the rails, where half of the track is considered in dynamic simulations.

#### 2.1.1. Vehicle model

The vehicle system is represented by a quarter-car model traveling at a constant speed $v$, as shown in Fig. 2. It is treated as a multibody system including a carbody, bogie frame, and wheelset. The wheelset is connected to the bogie through the primary suspension, and the carbody and bogie are linked through the secondary suspension. Note that more advanced vehicle systems, i.e., a full-car system or a series of cars, can be implemented. Here, as the focus is to examine the surrogate-assisted optimization in the level crossing design, a vehicle system with three degree-of-freedom (DOFs) is considered in the coupling dynamics with the track.

The equations of motion of the vehicle model can be written as

$$\mathbf{M}_v\ddot{\mathbf{U}}_v + \mathbf{C}_v\dot{\mathbf{U}}_v + \mathbf{K}_v\mathbf{U}_v = \mathbf{F}_v \tag{1}$$

where $\mathbf{M}_v, \mathbf{C}_v$, and $\mathbf{K}_v$ denote, respectively, the mass, damping, and stiffness matrices of the vehicle. $\mathbf{U}_v, \dot{\mathbf{U}}_v$, and $\ddot{\mathbf{U}}_v$ denote, respectively, the displacement, velocity, and acceleration vectors of the vehicle. $\mathbf{F}_v$ is the force vector applied on the vehicle, which contains gravity loads and the wheel–rail contact force. The expression of the system matrices depends on what vehicle DOFs are considered in the simulation. For the current 3-DOF vehicle system, the detailed formulation of the system matrices is shown in Appendix.

#### 2.1.2. Track model

The track model consists of an ERS-based level crossing in the middle and transition zones on both sides in the ballast track form. The ballast track is represented by a two-layer discretely supported model; see Fig. 2. It consists of one rail meshed by Euler–Bernoulli beam elements, railpads as Kelvin–Voigt (KV) elements, sleepers as mass elements, and underlying ballast and foundation collectively as the KV elements. Each rail element has two nodes with 2 DOFs, vertical translation and rotation, at each node. The KV element consists of one linear spring and one linear damper placed in parallel, which is commonly used to represent viscoelastic materials in railway structures, such as railpads and ballast layer (e.g., [39]).

The embedded track comprises a rail, fastening, concrete slabs, and foundation layer. Fig. 3 presents a simplified ERS cross-section. The fastening is typically an elastic poured compound bonding the rail and a resilient rubber strip under the railbase to provide track elasticity and constrain the vertical rail deflection. Space-saving components can be used, and PVC tubes are for cable installation.

Previous works such as [40] analyzed the dynamic behavior of the embedded track, where a model with two beams (rail and slab) is compared against a more advanced model that accounts for the lateral flexibility of the slab, i.e., two beams (rails) and a flexible plate (slab). The comparison was made in terms of wave propagation and dynamic responses (e.g., vertical displacement of the rail and stresses in the concrete slab). Results showed that the former model can be employed for a quick and sufficiently accurate assessment of the dynamic behavior of the embedded track. The current work therefore adopts the simplified version, where the rail and slab are modeled by the Euler–Bernoulli beam elements, connected by parallel KV elements as the rail fastening and supported by viscoelastic (Winkler-type) foundation.

The equations of motion of the track model can be expressed as

$$\mathbf{M}_t\ddot{\mathbf{U}}_t + \mathbf{C}_t\dot{\mathbf{U}}_t + \mathbf{K}_t\mathbf{U}_t = \mathbf{F}_t \tag{2}$$

where $\mathbf{M}_t, \mathbf{C}_t$, and $\mathbf{K}_t$ denote, respectively, the mass, damping, and stiffness matrices of the track structure. $\mathbf{U}_t, \dot{\mathbf{U}}_t$, and $\ddot{\mathbf{U}}_t$ denote, respectively, the displacement, velocity, and acceleration vectors of the track. $\mathbf{F}_t$ is the force vector applied on the track by the running train.

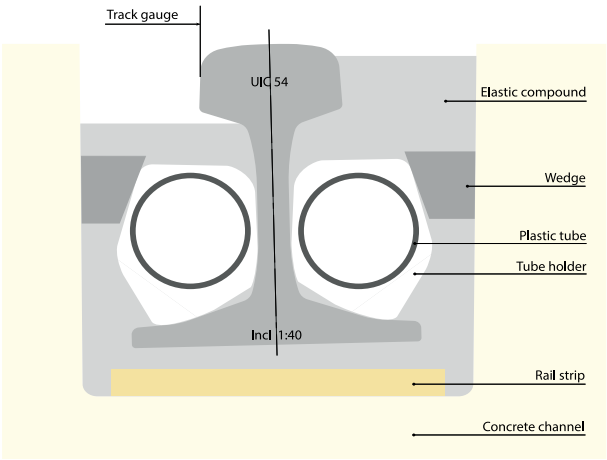



**Fig. 3.** Schematic representation of the ERS cross-section.

### 2.1.3. Coupling of vehicle and track model

Numerical simulations of the VTI dynamics are generally solved by two types of computational methods. One treats the vehicle and track structure as a unified system and formulates coupled system matrices, e.g., [41,42]. The other separates the two subsystems and solves individual equations of motion based on an iterative procedure, e.g., [43]. This paper adopts the former method that couples the vehicle and track to form an integrated time-dependent system. It results in a global system of equations that can be solved in a directional manner without the need for an iterative procedure.

Combining the equations of motion for the vehicle, Eq. (1), and track, Eq. (2), the unified formulation results in the following global system equations of motion.

$$\mathbf{M}_g\ddot{\mathbf{U}}_g + \mathbf{C}_g\dot{\mathbf{U}}_g + \mathbf{K}_g\mathbf{U}_g = \mathbf{F}_g \tag{3a}$$

where $\mathbf{M}_g$, $\mathbf{C}_g$, and $\mathbf{K}_g$ denote, respectively, the mass, damping, and stiffness matrices of the global integrated system. $\mathbf{U}_g$, $\dot{\mathbf{U}}_g$, and $\ddot{\mathbf{U}}_g$ are the displacement, velocity, and acceleration vectors of the global system, respectively. $\mathbf{F}_g$ is the global force vector. Their expressions are given by

$$\mathbf{M}_g = \begin{bmatrix} \mathbf{M}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_t \end{bmatrix},\ \mathbf{C}_g = \begin{bmatrix} \mathbf{C}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_t \end{bmatrix},\ \mathbf{K}_g = \begin{bmatrix} \mathbf{K}_{vv} & \mathbf{K}_{vt} \\ \mathbf{K}_{tv} & \mathbf{K}_{tt} \end{bmatrix},$$
$$\mathbf{U}_g = \begin{bmatrix} \mathbf{U}_v \\ \mathbf{U}_t \end{bmatrix},\ \mathbf{F}_g = \begin{bmatrix} \mathbf{F}_{vt} \\ \mathbf{F}_{tv} \end{bmatrix}. \tag{3b}$$

The wheel–rail contact is modeled by Hertz contact theory. The contact stiffness $k_w$ is linearized and calculated by $k_w = \frac{3}{2G}P^{1/3}$, where $P$ is the static wheel axle load and $G$ represents a contact constant [44]. As the wheel–rail contact is modeled by a Hertzian spring $k_w$, the vehicle and track is coupled through the stiffness matrices only, which is indicated through the entries of $\mathbf{K}_g$ in Eq. (3b). $\mathbf{K}_g$ is time-variant since the position of the train changes with time. The detailed formulation of $\mathbf{K}_g$ and $\mathbf{F}_g$ is given in Appendix.

## 2.2. Formulation of an optimization problem

A common guideline for improving the transition performance is to smooth the variation in vertical track stiffness, which is also specified in the European Standard 16432-2 [45]. According to this principle, special measures in transition zones have been developed to mitigate the degradation, as discussed in Section 1. Likewise, the design principle in this work is to obtain a gradual change of stiffness in the track, where relevant response quantities (in Section 2.2.2) defined to capture the local dynamic amplification are optimized or minimized.

The stiffness variation and differential settlement are the main causes that lead to the degradation in transition zones [25]. The latter generates irregularities in longitudinal level, which significantly influences the wheel–rail contact forces. As presented in Eqs. (A.12) and (A.13) (see Appendix), the irregularities can be incorporated in the current vehicle–track simulation, which can be inferred from track geometry measurement or simulated from predefined power spectral densities (PSDs) from relevant railway authorities. However, in this work, since the track geometry condition is location-specific, the transition zone is designed without irregularities in track geometry, approximately representing the situation when a line is open to traffic. If a prescribed track geometry were incorporated as input in the vehicle–track simulation, the design variables would be optimized to that specific irregularity profile. Here, it is more relevant to design a transition zone to mitigate the expected track degradation, and it can be achieved by optimizing a transition zone such that the impact of stiffness variations is minimized.

### 2.2.1. Definition of design variables

A *reference design* that follows a typical Dutch practice in level crossings is defined as a benchmark to assess the optimized design solutions. The parameter values and the justification for the chosen values are given in Section 3.1. In the crossing area with ERS, an alternative to the reference design is to install another type of rail strip (herein referred to as *Type II rail strip*). The rail strip is a resilient component fitting between the underside of the rail and slab channel, as depicted in Fig. 3. The component provides elasticity to the structure, improves the load distribution, and controls the rail deflection. It has predefined stiffness properties that can be chosen to meet specific system requirements.

For the ballast track in transition zones, the vertical track stiffness is influenced by the flexural stiffness of the rail and supporting stiffness contributed from track components below the rail, such as sleepers and ballast bed. The contribution from the substructure is also significant to the vertical track stiffness, especially in soft soil regions [21]. Modifying the stiffness of rail, sleepers, and ballast can cause problems in track stability and resistance [46]. An adequate solution is introducing elastic components in the track structure, such as railpads, under sleeper pads, and under ballast mat. The railpad stiffness is the main characteristic parameter of these elements and comes with a wide range of values [46]. Stiff pads such as standard Dutch pads FC9 reduce rail deflection and vibration, while soft pads can mitigate the damage in sleepers and reduce vibration in sleepers and ballast.

Moreover, sleeper parameters such as size, spacing, and material type have a considerable effect on track dynamic behavior, where effort has been made to evaluate the variation of these parameters on transition performance, e.g., [6,26,47–49]. Referring to the Dutch practice, two types of concrete sleepers are considered in the design space, i.e., a commonly used NS90 sleeper and the other option with larger weight and stability (herein referred to as *strengthened sleepers*). The strengthened type is mainly used in tight curves or in connection with level crossings, switches, and other structures to ensure gradual track stiffness changes. According to Prorail (Dutch railways) system specifications, at least 5 to 8 strengthened sleepers should be placed next to the ERS-type level crossings. And the sleeper bay right next to the concrete slab should be reduced to 0.4 m (for comparison, the standard sleeper spacing is 0.6 m). However, questions remain in approach zones regarding the optimal number of strengthened sleepers being installed next to the junction and the distances between centers of those neighboring sleepers.

The above design parameters are considered influential to the dynamic performance of the level crossings. In this paper, they are collected as four major types of design variables in formulating the optimization problems, which are displayed in Table 1.

**Table 1**
Types of track design variables, **x**.

| Track type | Component | Variables |
|---|---|---|
| Ballast | Railpad | Stiffness, $(x_{ri}, i = 1, 2, \ldots)$ |
| | Sleeper | Number of strengthened sleeper, $(x_n)$ |
| | | Sleeper spans, $(x_{sj}, j = 1, 2, \ldots)$ |
| ERS | Rail strip | Length of placing Type II strip, $(x_l)$ |

#### 2.2.2. Assessment criteria of track performance

Track degradation is generally reflected at two levels: the one related to the wheel–rail interface and the other concerning the supporting elements below the rail, such as sleepers and ballast bed.

At the wheel–rail interface, undesirable contact degrades both rail and wheel profiles, and contact properties influence the degradation rates [50]. The track degradation at the contact level is commonly revealed in rail defects, such as rail corrugation and rolling contact fatigue (RCF). These degradation modes represent short-wave components of track irregularities, which significantly affect the magnitude of wheel–rail interaction forces and are of relevance to driving safety and vehicle stability [51].

At the lower supporting level, the ballast bed and underlying substructure are important contributors to the deterioration in track geometry. Due to the sliding and breakage of granular particles, the ballast layer presents a progressive deformation with the passage of trains. The layers below the ballast also cause plastic deformations due to consolidation and cyclic loading, which further contributes to the development of the accumulated settlement. On the other side, however, the slab track in level crossings or bridge structures is designed for minimal settlement, resulting in a differential settlement between the two track forms.

Based on the typical features, bi-level criteria are defined to assess the sensitivity of a track design to the expected level of performance/degradation: (1) those relevant to the wheel–rail contact are selected as the first-level assessment criteria, where the magnitude of dynamic contact forces is representative. A larger magnitude of wheel–rail vertical forces $F$ implies more dynamic amplification in the track structure induced by the passing vehicle, and consequently, reducing the amount of $F$ represents damage mitigation at the wheel–rail contact and reduction in potential rail defects; (2) the second-level measure concerns the damage to the ballast layer as it is a significant influencer in track geometry degradation. The mechanical energy dissipated in the track ballast is selected as an indicator to assess the sensitivity of a track design to the expected damage in the substructure, which is proposed in [47] and further elaborated in [16,52]. For one wheel passage, the energy dissipated by the ballast damping in the $i$th sleeper support is given by [47],

$$E_i = \int_{-\infty}^{\infty} c_{b,i} v_{s,i}^2(t)\, dt \tag{4}$$

where $c_{b,i}$ is the ballast damping under the $i$th sleeper; $v_{s,i}(t)$ is the velocity response of the $i$th sleeper in the time domain. The higher the energy dissipated into the substructure, the stronger the degradation can be expected. Accordingly, reducing the amount of energy dissipated in the ballast layer represents an important aim for damage reduction in the ballast and therefore in overall track geometry.

The response quantities, i.e., $F$ and $E$, are generated from VTI dynamic simulations and presented in time series. To better capture the features embedded in the responses, the simulation data are further processed by two statistical metrics, i.e., the root mean square ($rms$) and maximum-to-minimum ($max$) value. The former captures the spatial variability of dynamic amplification over the influenced track section. The latter represents the maximum difference in a response in the influenced area. The evaluation of the metrics results in four design objectives, i.e., $F_{rms}$, $F_{max}$, $E_{rms}$ and $E_{max}$, which are treated as dynamic benchmarks for track design comparison and parametric optimization.

As previously mentioned, the ballast and substructure are modeled by the Kelvin–Voigt elements. The damping parameter of the ballast layer is considered constant, referring to characteristics of the typical Dutch track, as elaborated in Section 3.1. Although this representation can describe the elastic resistance and damping provided by the layers to support the sleepers during the train passage, the most appropriate way of quantifying the dissipation in those layers requires a 3D representation accounting for, e.g., track–soil interaction and nonlinear material behavior. However, the focus of this work is parametric optimization, and the energy dissipation (Eq. (4)) is to quantify the susceptibility of a track design to the expected degradation, where the assessment of the effect of parametric variations on the dissipated energy is relative (i.e., to compare various track design solutions). It is considered sufficient given the aim of the current work. For further applications, absolute quantification of the long-term degradation and the associated practicality such as the implication to the maintenance level can be incorporated into the present methodology, which is outside the scope of this work.

#### 2.2.3. The optimization problem

The design variables described in Section 2.2.1 can be collected in a design vector $\mathbf{x} = \left[x_{ri}, x_n, x_{sj}, x_l\right]'$ with $i, j \in \mathbb{N}$, which can be varied to improve the performance characteristics of the track structure. For instance, in the case of three railpads ($i = 3$) and four sleeper spans ($j = 4$), it yields a $3+1+4+1 = 9$ dimensional optimization problem. To minimize the dynamic amplification at the ballast-to-ERS transition, a general single-objective optimization problem can be formalized in the following form.

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & x_{ri} \in \mathbb{R} : x_{ri}^l \le x_{ri} \le x_{ri}^u,\ i \in \mathbb{N} \\ & x_n \in \mathbb{Z} : 0 \le x_n \le x_n^u \\ & x_{sj} \in \mathbb{R} : x_{sj}^l \le x_{sj} \le x_{sj}^u,\ j \in \mathbb{N} \\ & x_l \in \mathbb{R} : 0 \le x_l \le x_l^u \end{aligned} \tag{5}$$

where the superscripts $l$ and $u$ represent the lower and upper bound of each design variable, respectively; $f(\mathbf{x})$ is an objective function defined by each of the performance criteria described in Section 2.2.2. It represents a true function that maps a given input design vector $\mathbf{x}$ and a function value $f(\mathbf{x})$ calculated from vehicle–track dynamic simulations.

It is expected that the bi-level objectives, i.e., $F$ and $E$, are conflicting since some design variables by nature have the opposite effect on the objectives. For example, stiff railpads contribute to a reduction in noise and vibration from wheel–rail contact. At the same time, soft pads allow for a lower effect of loads transmitted to underlayers and thus reduce vibration in sleepers and ballast particles [46]. For this reason, it is necessary to perform simultaneous minimization of the objectives, where a multi-objective optimization (MOO) problem should be formulated to search for the best optimal compromise solution.

The statistical metrics (i.e., $rms$ and $max$) used to evaluate the dynamic responses are compared based on the results obtained from the single objective problems: those showing higher solution quality and sensitivity to the parametric variation are chosen for the MOO problem formulation.

Without loss of generality, a MOO problem with $k\,(\ge 2)$ objective functions can be expressed as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{F}(\mathbf{x}) = \left(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})\right), \\ \text{s.t.} \quad & \mathbf{x} \in \Omega \end{aligned} \tag{6}$$

where $f_i(\mathbf{x})$ is the $i$th objective and $\mathbf{F} : \Omega \to \Lambda$ maps the design variables $(\mathbf{x} = x_1, \ldots, x_n)$ to vectors $(\mathbf{y} = y_1, \ldots, y_k)$ in the objective function space $\Lambda$.

When dealing with MOO in surrogate optimization, the most common way is to build a surrogate for each objective function, which

is however time-consuming considering the training time for multiple surrogates [53]. Another way is to convert a MOO problem to a single objective problem using an achievement scalarizing function (ASF). A single surrogate is built on the scalarizing function and applied to search for optimal solutions. ASFs, introduced by [54], serve to map $k$ objective functions to a scalar, which is a priori preference articulation in MOO problems. Certain properties of ASFs guarantee Pareto optimality of the solutions obtained from a scalarizing problem (See [55] for a detailed description). This approach reduces the computational complexity as only one surrogate is built and one infilling criterion (see Section 2.3.2) is used in the optimization workflow.

The current work adopts the ASF-based mono surrogate approach to deal with the expensive MOO problem, given the simplicity and computational efficiency. Specifically, the (conflicting) objectives selected from the single-objective simulation round is scalarized into one global function by an ASF of augmented Chebyshev type in a form [56],

$$g\left(\mathbf{x},\mathbf{f}^T\right)=\max_{i\in\{1,\ldots,k\}}\left\{\kappa_i\left(f_i\left(x\right)-f_i^r\right)\right\}+\rho\sum_{i\in\{1,\ldots,k\}}\kappa_i\left(f_i\left(x\right)-f_i^r\right) \tag{7}$$

where $f_i\,(i=1,\ldots,k)$ are performance objectives selected from the single-objective optimization problems; $\rho>0$ is an arbitrary small parameter and $\mathbf{f}^T=\left[f_1^r,\ldots,f_k^r\right]$ is a vector that defines a reference point. $\kappa_i$ are non-negative normalization coefficients. The idea of this function is to minimize the deviations from the reference objective, and any reasonable or desired point in the objective space specified by the decision maker can be considered as a reference objective [57]. The second term of the function guarantees that all the objectives play a role, not only the one more deviated from the reference value. Through the scalarization, a MOO problem for optimizing track dynamic amplification can be formulated by minimizing Eq. (7), given constraints defined in Eq. (5).

### 2.3. Surrogate-assisted optimization

#### 2.3.1. A surrogate model: radial basis function interpolation

Various types of surrogate models have been applied to support engineering tasks, such as Kriging [39,58], radial basis function [59], and neural networks [60,61]. Specifically, in the railway field, Kriging models were trained to approximate the relationship between track parameters and frequency response function (FRF) features [39], and neural networks were used to predict responses of a vehicle-bridge system [60,61], which are all developed based on FE numerical simulations.

The surrogate models can be either interpolating or non-interpolating, and parametric or nonparametric [62]. For the current VTI model, each simulation run is deterministic. Moreover, while training a surrogate for such mechanical problems, it is often reasonable to assume the true objective function can be evaluated precisely or with a minor approximation error at sampled input points [63]. For this purpose, radial basis function (RBF) interpolation is used to approximate the solutions of the VTI simulation, considering its powerful convergence properties and easily adjustable smoothness [64]. Specifically, a cubic RBF interpolation is employed as it showed a competitive performance profile compared to other surrogates in Ref. [62].

An RBF interpolant is defined as

$$\hat{f}(\mathbf{x})=\sum_{k=1}^{n}\lambda_k\phi\left(\left\|\mathbf{x}-\mathbf{x}_k\right\|\right)+p\left(\mathbf{x}\right) \tag{8}$$

where in the first term, $\mathbf{x}_k, k=1,\ldots,n$, denotes the points that have been evaluated by the true objective function; $\|\cdot\|$ is the Euclidean norm; $\lambda_k\in\mathbb{R}, k=1,\ldots,n$ are coefficients; $\phi(\cdot)$ is a radial basis function and many function forms are available as described in Ref. [65]. The cubic RBF interpolant uses the cubic function $\phi(r)=r^3$ (where $r=\left\|\mathbf{x}-\mathbf{x}_k\right\|$).

The second term in Eq. (8) represents a polynomial tail whose order depends on the chosen RBF [38]. The general form is defined as $p\left(\mathbf{x}\right)=\sum_{l=1}^{m}\beta_l p_l\left(\mathbf{x}\right)$, with $m$ denoting the order of the basis $p_l(\cdot)$ and $\beta_l$ as the coefficients. For the cubic RBF, it should be at least a linear polynomial, and it becomes $p\left(\mathbf{x}\right)=a+\mathbf{b}^T\mathbf{x}$ with coefficients $a\in\mathbb{R}$ and $\mathbf{b}=\left[b_1,\ldots,b_d\right]^T\in\mathbb{R}$ [66]. The coefficients $\lambda_k$, $a$ and $\mathbf{b}$ are determined by solving the following linear system of equations

$$\begin{bmatrix}\boldsymbol{\Phi} & \mathbf{P}\\ \mathbf{P}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\lambda\\ \beta\end{bmatrix}=\begin{bmatrix}\hat{\mathbf{f}}\\ 0\end{bmatrix} \tag{9}$$

where $\boldsymbol{\Phi}_{kv}=\phi\left(\left\|\mathbf{x}_k-\mathbf{x}_v\right\|\right),\ k,v=1,\ldots,n$, and

$$\mathbf{P}=\begin{bmatrix}\mathbf{x}_1^T & 1\\ \mathbf{x}_2^T & 1\\ \vdots & \vdots\\ \mathbf{x}_n^T & 1\end{bmatrix},\quad \lambda=\begin{bmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{bmatrix}\quad \boldsymbol{\beta}=\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_d\\ a\end{bmatrix},\quad \hat{\mathbf{f}}=\begin{bmatrix}\hat{f}\left(\mathbf{x}_1\right)\\ \hat{f}\left(\mathbf{x}_2\right)\\ \vdots\\ \hat{f}\left(\mathbf{x}_n\right)\end{bmatrix} \tag{10}$$

The matrix in Eq. (9) is invertible if and only if rank $(\mathbf{P})=d+1$ [38,64].

#### 2.3.2. An efficient infilling scheme

Once a surrogate model is built, optimal solutions can be searched using surrogate function values. The original function $f(\mathbf{x})$ in Eq. (5) is replaced by a surrogate model $\hat{f}(\mathbf{x})$, which shifts the problem to minimize the function $\hat{f}(\mathbf{x})$. However, the estimated function values (i.e., $\hat{f}(\mathbf{x})$) obtained from exploring the surrogate are subject to the model accuracy. Strategies are required that improve the surrogate model accuracy while guiding the search to promising areas of the design space [59]. This type of strategy typically balances exploration and exploitation. Exploration samples the regions far from any optimum searched before, thereby having high uncertainty. Exploitation concerns the search in local (promising) areas with the hope of improving the current optimum.

The key idea of such strategies is to pick the next function evaluation point based on the surrogate predictions and a measure of the error in this model [65]. It is essentially a way of refining the surrogate model while guiding the search towards the optimum within a relatively limited number of function evaluations (i.e., to keep the total computational time manageable). The process of adding the next evaluation point based on this 'strategic' sampling is the so-called infilling scheme. The chosen point for the next function evaluation generally refers to infilling/adaptive point, and the measures to determine the point are known as infilling criteria (also acquisition/merit functions).

The adaptive surrogate modeling has been applied in many mechanical systems with various engineering purposes, e.g., design optimization for vehicle crashworthiness [67], aerodynamic shape [59], and wing typology [68], reliability analysis of a hydrokinetic turbine blade and a hysteretic oscillator [69], and material parameter identification of a specimen bending system [70]. However, the application in the railway field, especially for structural design, needs more exploration. The technique is well-suited to expensive simulations where the budget of function evaluations is limited to 20–200 [71]. It is therefore a promising tool to deal with optimization problems in railway track design involving intensive calculations of vehicle–track dynamic simulations.

An efficient balancing strategy is developed by Regis and Shoemaker [72], where an algorithmic framework called Metric Stochastic Response Surface (MSRS) is introduced for global optimization of expensive functions. It executes an adaptive learning process where at each iteration, (1) the adaptive point is chosen based on a merit function from a sequence of random candidate points, and (2) the true objective function value of the adaptive point is evaluated and used to update the surrogate model. This process continues until a stopping criterion is satisfied.

The merit function proposed in [72] is a weighted combination of function values from the current surrogate (*response surface criterion*) and distances to previously evaluated points (*distance criterion*). At each iteration, a candidate point set $\Omega_c$ is generated randomly by adding

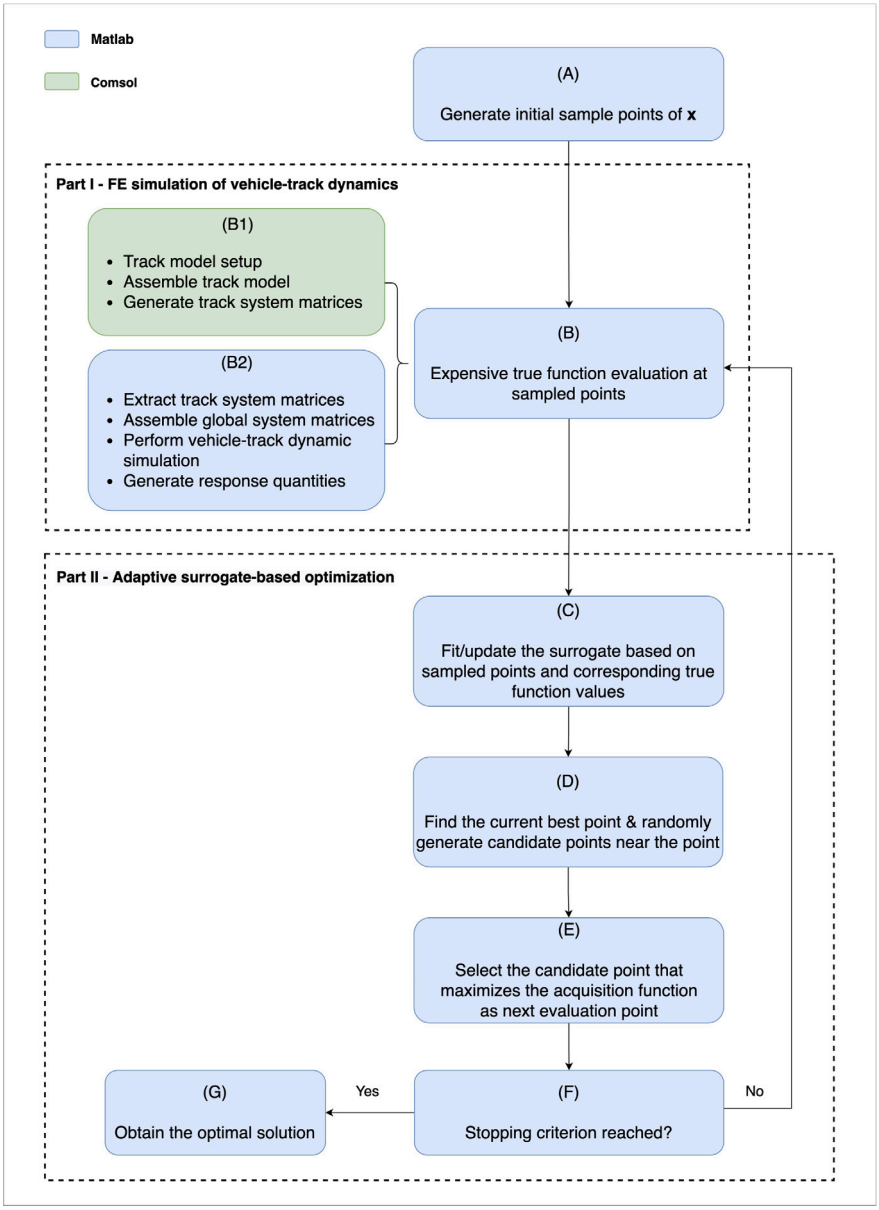



**Fig. 4.** Illustrative flowchart of the simulation methodology integrating (a) part I: FE simulation of vehicle–track dynamics; and (b) part II: adaptive surrogate-based optimization.

perturbations to the best point found so far $x_{\text{best}}$ (see [72] for a detailed discussion on the random perturbations). Each candidate point in $\Omega_c$ is given a score by evaluating the corresponding merit function, and the one with the lowest score is selected as the adaptive point.

Define $x_k (k = 1, \ldots, n)$ as $n$ previously evaluated points in $\Omega_k$ and $x_c^v (v = 1, \ldots, t)$ as $t$ candidate points in $\Omega_c$. For each candidate point $x_c^v$, the merit function is expressed as

$$u\left(x_c^v\right) = \omega \frac{\hat{f}\left(x_c^v\right) - \hat{f}_{\min}}{\hat{f}_{\max} - \hat{f}_{\min}} + (1-\omega)\frac{d_{\max} - d\left(x_c^v\right)}{d_{\max} - d_{\min}} \tag{11}$$

where on the right-hand side, the first term corresponds to the response surface criterion and the second refers to the distance criterion. The parameter $\omega$ is a weight with $0 < \omega < 1$ that balances minimizing the surrogate value (*response surface criterion*) while exploring the space to improve the model accuracy (*distance criterion*). $\hat{f}\left(x_c^v\right)$ is the current surrogate model evaluated at point $x_c^v$, $\hat{f}_{\min} = \min\left(\hat{f}\left(x_c^v\right)\right)$ and $\hat{f}_{\max} = \max\left(\hat{f}\left(x_c^v\right)\right)$; $d\left(x_c^v\right) = \min\left(d_{kv}\right)$ with $d_{kv}$ being the distance from an evaluated point $x_k$ to a candidate point $x_c^v$, $d_{\min} = \min\left(d\left(x_c^v\right)\right)$ and $d_{\max} = \max\left(d\left(x_c^v\right)\right)$.

The MSRS framework in Ref. [72] is developed for continuous optimization problems. When dealing with integrality constraints, as often encountered in engineering problems, the points generated from random perturbations should be guaranteed to satisfy the integrality constraints. This is done by relaxing the integer condition while sampling and then rounding the values of the integer variables in the obtained points. A similar treatment can be found in Müller [38]. In parallel, a Branch-and-Bound approach can be used to supplement the random perturbations and search for the minimum of the merit function. This approach has been embedded in other surrogate-based optimization frameworks to deal with mixed-integer problems [68,73]. The idea here is to enhance exploration of the space: the RBF is a continuous fit, and if only evaluating the points from the random perturbations (by rounding the numbers), it may miss regions of the space that account for relevant information. Therefore, the sampling is supplemented by branching and searching locally in the RBF function. In this case, the candidate point set $\Omega_c$ is enhanced by the Branch-and-Bound approach (in addition to the random perturbations), where the point with the lowest merit function score is chosen for the next iteration in the sequential enrichment of the surrogate.

### 2.4. Integrative simulation methodology

The methodology for optimizing geometric and elastic properties of level crossings involves an integration of the VTI dynamic simulations (in Section 2.1) and adaptive surrogate-based technique (in Section 2.3). As illustrated in Fig. 4, the proposed integrated simulation methodology has an iterative procedure that is synthesized as follows.

First, $n$ initial distinct sample points of the design vector $\mathbf{x}$ are generated from a space-filling design, e.g., Latin hypercube sequence (Step A). FE simulations of vehicle–track dynamics are evaluated at the sampled points, and the corresponding function values are obtained (Step B). A detailed description of Step B is provided in the following paragraphs. In Part II, the surrogate-based optimization workflow consists of two phases, namely, the surrogate construction and adaptive learning phase: (1) initially, a surrogate of the objective function is constructed by interpolating a cubic RBF interpolant, Eq. (8), through the $n$ evaluated points (Step C - at the first iteration). (2) Then, in the adaptive learning phase, the surrogate function is updated (Step C - at successive iterations), where the next point of evaluation is selected according to the merit function, Eq. (11), from a sequence of random candidate points (Step D and E). The solution of this new point is evaluated with the expensive/true function, i.e., FE simulations in Step B. This process continues until a stopping criterion is reached (Step G). Here, the maximum number of function evaluations is selected as the stopping condition, as emphasized in Ref. [38,74,75], where in many surrogate-assisted optimization problems, termination depends more upon the computational budget than on a measure of convergence due to the high computational cost of the function evaluation.

Numerical solutions for the VTI models are often developed by self-programmed codes or FE software. The software is more flexible in generating track models with complex structural configurations but is often computationally intensive. Although introducing the surrogate-based approaches reduces the number of function evaluations in the optimization workflow, repetitive calls of the VTI simulation are still required to train a surrogate. To further relieve the computational burden while easing the generation of the track model, a combined simulation approach is proposed to model the vehicle–track coupling dynamics. The approach of using a general FE software COMSOL and MATLAB interface has been demonstrated in a baseline case by Shang et al. [76], where a general beam model subject to a moving vehicle is simulated and verified by a benchmark case coded in MATLAB. Here the methodology is extended to the level crossing case.

Specifically, in Step B (Part I of the integrated methodology), COMSOL is used to establish a ballast-ERS transition, which characterizes the track used in a typical heavy-duty level crossing. The track system matrices $\mathbf{M}_t, \mathbf{C}_t$, and $\mathbf{K}_t$ in Eq. (2) are generated in COMSOL and further exported to MATLAB through Livelink interface. The software used in the steps is highlighted in Fig. 4. The vehicle system matrices $\mathbf{M}_v, \mathbf{C}_v$, and $\mathbf{K}_v$ in Eq. (1) are formulated and coupled with the extracted track matrices in MATLAB to form global system matrices (Eq. (3b)). The coupled equations of motion (Eq. (3a)) that govern the vehicle–track dynamics are solved in the time domain using Newmark-$\beta$ integration scheme. Post-processing of the numerical results is realized in MATLAB. Response quantities given a parameter set are also generated accordingly.

The simulation approach that couples COMSOL and MATLAB to model the vehicle–track dynamics is validated against the result in [77], where a similar track form is considered, i.e., a connection between a floating slab track (FST) and a ballast track. The FST is similar to the embedded track, which consists of concrete slabs supported on resilient elements [78]. In [77], the floating slabs are modeled in discrete precast sections and the rails are supported periodically. The simulation setting for the track and vehicle system is modified according to the case in [77], and response quantities are generated for comparison. Fig. 5 presents the wheel and rail displacement at the contact point when a vehicle moves from the FST to ballast track, showing that the patterns agree well with those generated from the original work (see Figure 3 in [77]). From the FST to the ballast track, the current simulation result indicates the averaged rail and wheel displacement (for both axles) decreases by about 2.8 and 2.7 mm, respectively. In the reference case, 2.8 mm was reported for the change in the rail displacement, and 2.6 mm was for the wheel displacement. Therefore, it is considered the current methodology to simulate the vehicle–track dynamics can be properly used as a basis for track parametric optimization.

**Table 2**
Vehicle parameters.

| Parameter | Expression | Value |
|---|---|---|
| Carbody mass | $m_c$ | 56378 kg |
| Bogie mass | $m_b$ | 3772 kg |
| Wheelset mass | $m_w$ | 1819 kg |
| Primary suspension stiffness | $k_{s1}$ | 2445 kN/m |
| Secondary suspension stiffness | $k_{s2}$ | 2227 kN/m |
| Primary suspension damping | $c_{s1}$ | 2 kN s/m |
| Secondary suspension damping | $c_{s2}$ | 50.1 kN s/m |
| Static wheel load | $P$ | 100 kN |
| Contact constant | $G$ | $5.13 \times 10^{-8}$m/N$^{\frac{2}{3}}$ |
| Velocity | $v$ | 140 km/h |

## 3. Numerical study

### 3.1. Characterization of the VTI model

Numerical examples are presented in this section to demonstrate the applicability of the proposed procedure. Vehicle parameters used in the model refer to VIRM trains, which are double-deck trains operated by NS (Dutch Railways). Table 2 lists parameter values of a full-car system according to VIRM technical parameters. A conversion is made in the computation to reduce a full car to the current 3-DOF model. Moreover, following the half track, the vehicle parameters are reduced further to half.

Track parameters correspond to characteristics of the typical Dutch track, which are based on a full-track scale and presented in Table 3. Rail properties conform to nominal values of UIC54 rail, which are consistent in both track forms. The rail is pinned at both ends, which implies reflections of the vibration will occur at the ends. However, the boundary effect can be limited or negligible when the track length is sufficient [42]. For this purpose, the total track length in the demonstration cases is extended to 134 m, including a ballast track section (100 m, on the left), an ERS-based level crossing (12 m, middle), and the other ballast track section (22 m, right), as shown in Fig. 2.

Sleepers provide periodic rail support in the ballast track. The commonly accepted spacing between centers of adjacent sleepers is 0.6 m [79], which is considered here with six rail elements per sleeper bay. Railpads and sleepers refer to Dutch standard components in the ballast track, i.e., FC9 4.5-mm cork-rubber pad and NS90 concrete sleeper. The properties are determined referring to Prorail system specifications and relevant works in Ref. [80–82].

The embedded track in crossings corresponds to the standard solution for heavy-duty (ERS-type) level crossings. It consists of several concrete slabs, each with a length of 6 or 9 m. By combining different slabs, level crossings can be installed with variations in length. The slab is lowered onto a ditch filled with mixed granulate and stabilized sand layer. As the ability to adjust the track geometry after the construction is limited, the requirement for the substructure is generally very high for slab track [81]. To reinforce the substructure layers, measures such as geogrids and injection mortar should be applied. Accordingly, in this work, a stiff substructure underneath the slabs is applied; the parameter values are chosen according to Ref. [80], which are collected in Table 3. The geometry and mechanical properties of the concrete slab correspond to the actual design of the ERS-type level crossings. Stiffness

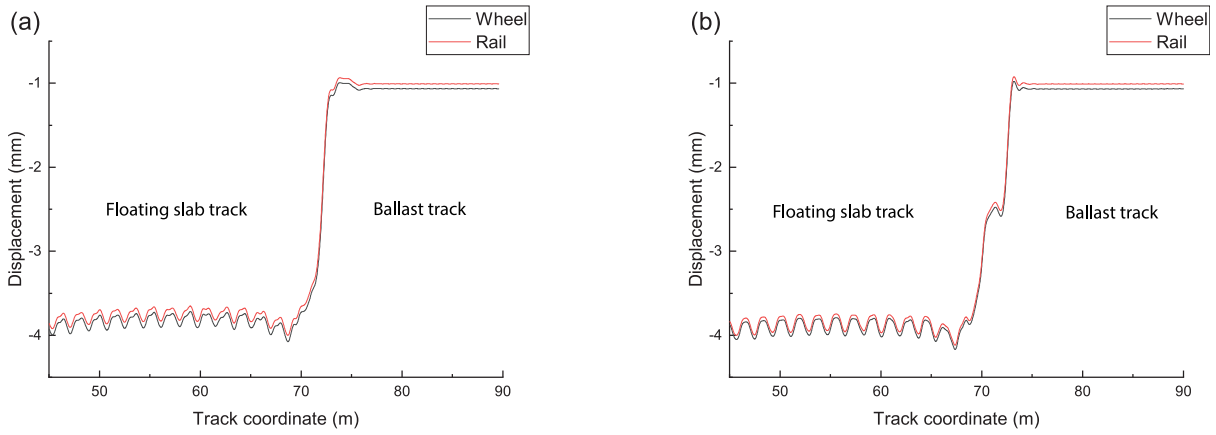



**Fig. 5.** Wheel and rail displacement of the front axle (a) and rear axle (b) when a vehicle moves from the FST to ballast track.

**Table 3**
Track parameters.

| Track Type | Component | Parameter | Value |
|---|---|---|---|
| Ballast/ERS | Rail (54E1) | Young's modulus | 210 GPa |
| | | Poisson's ratio | 0.3 |
| | | Density | 7800 kg/m$^3$ |
| | | Moment of inertia | $2.337 \times 10^{-5}$ m$^4$ |
| | | Area of cross section | $6.977 \times 10^{-3}$ m$^2$ |
| Ballast | Railpad (FC9) | Stiffness | 1120 MN/m |
| | | Damping | 120 kN s/m |
| | Sleeper (NS90) | Mass | 251 kg |
| | | Support spacing | 0.6 m |
| | Ballast/subgrade | Stiffness | 45 MN/m |
| | | Damping | 96 kN s/m |
| ERS | Elastic compound incl. rail strip | Stiffness | 54 MN/m/m |
| | | Damping ratio | 0.2 |
| | Slab | Young's modulus | 31 GPa |
| | | Poisson's ratio | 0.3 |
| | | Density | 2500 kg/m$^3$ |
| | | Width | 2.37 m (bottom); 2.23 m (top) |
| | | Length | 6 m ×2 |
| | | Height | 0.58 m |
| | Subgrade | Stiffness | 500 MN/m/m |
| | | Damping | 20 kN s/m |

and damping properties of elastic fastening are gathered from product specifications calibrated based on laboratory experiments.

Note that the material properties of the track are different when they are measured by either quasi-static or dynamic loading tests. The material properties are commonly referred to as static properties when they are measured by quasi-static loading tests; those measured by dynamic loading tests are referred to as dynamic properties [83]. A comprehensive review regarding the testing methodology for dynamic characteristics of track components can be found in [84]. The static material properties are adopted in most previous works to simulate the vehicle–track dynamics [83], which is also the case in some works of railway design optimization (e.g., [36]). However, a proper dynamic simulation requires dynamic properties as input.

The present work considers the effect of dynamic material properties on the vehicle–track dynamic simulation, where the chosen values for specific track components follow the suggestions concluded from the comparative study in [83]. The most relevant material properties to dynamic excitations in the track are the stiffness of elastic elements and the modulus of elasticity of concrete. The dynamic stiffness of railpads significantly increases the dynamic impact factor (DIF) based on the wheel–rail contact force, while the dynamic modulus of elasticity has less effect compared with the former [83]. In this work, dynamic stiffness values of railpads (FC9) and elastic compounds in ERS are adopted in the simulation, which are also the main design variables (therefore are of significance) in the optimization problems.

It is also worth mentioning that the properties of railpads are temperature-dependent and frequency-dependent, and it is also sensitive to preload and aging [80,82,83]. In railway practice, the railpads are commonly simplified as the spring and dashpot elements, and constant values are used to describe the viscoelasticity characteristics. The current work adopts this representation, while more advanced models that describe the sensitivity of railpads to those factors (such as [82]) can be incorporated for a more accurate representation of the railpad behavior. However, this would require more computational effort, especially when combined with the simulation involving vehicle–track interaction dynamics.

### 3.2. Case study: The Dutch level crossing design

Single-objective optimization problems are formulated for demonstration, where performance criteria, i.e., $F_{rms}$, $F_{max}$, $E_{rms}$ and $E_{max}$, are compared as dynamic benchmarks for track parametric optimization. Each measure is tested against the reference design (as defined in Table 3) to evaluate its effectiveness concerning the solution quality and sensitivity to the parametric variation. It is expected that the vehicle velocity highly influences optimization results as studies have shown that track response and expected degradation are sensitive to the train speed (e.g., [47,83]). Here as the main purpose is to demonstrate the applicability of the integrative simulation approach, the velocity considered in the examples is defined as 140 km/h, referring to the

**Table 4**
Track design variables and corresponding range of definition.

| Track type | Component | Variables | Unit | Range of definition |
|---|---|---|---|---|
| Ballast | Railpad | Stiffness, $(x_{ri}, i = 1,2,3)$ | MN/m | $x_{ri} \in \mathbb{R} : x_{ri} \in [50, 1200], i = 1,2,3$ |
| | Sleeper | Number of strengthened sleeper, $(x_n)$ | – | $x_n \in \mathbb{Z} : x_n \in [0, 20]$ |
| | | Sleeper spans, $(x_{sj}, j = 1,2,3)$ | – | $x_{sj} \in \mathbb{Z} : x_{sj} \in [3, 7], j = 1,2,3$ |
| ERS | Rail strip | Length of placing Type II strip, $(x_l)$ | – | $x_l \in \mathbb{Z} : x_l \in [0, 120]$ |

standard speed of conventional passenger trains (e.g., VIRM trains) in the Dutch railway lines.

Based on the general definition in Section 2.2.1, the variables and corresponding range of definitions used in the numerical examples are presented in Table 4. The first type of variables considers the stiffness distribution of railpads placed right next to the level crossing. Three pads are included in the optimization, which results in three design variables, $x_{ri}, i = 1,2,3$. The upper and lower bounds refer to the scope in Ref. [85], where railpads are classified according to their vertical stiffness $k_r$ (in unit: MN/m): very soft ($k_r = 100$), soft ($k_r = 200$), stiff ($k_r = 400$), and very stiff ($k_r = 800$). The current range of definitions covers the suggested values and also includes the consideration of the FC9 railpad (used in the reference design; see Table 4) to provide reasonable design space. With the stiffness change, the railpad damping is also varied in the search process, which is scaled linearly with the stiffness values.

The second type of variables is related to sleeper parameters. The number of strengthened sleepers $(x_n)$ applied in an approach is limited to 20, forming an approximately 12m-long track section. It is considered sufficient for the typical length of a level crossing approach, as in practice, it is prescribed to use $5 \sim 8$ strengthened sleepers in the approach to ERS-type level crossings; also Wang et al. [86] based on experimental analysis of bridge approaches suggested 4.5 m as the upper limit for the length of the studied transition.

Another variable related to sleepers is spacing $(x_{sj}, j = 1,2,3)$, which concerns distances between centers of three adjacent sleepers placed next to the junction. Due to an insufficient (continuous) depth of ballast, there is an inherent operational discontinuity of tamping works at the ballast-to-slab connections, which may lead to a group of sleepers never mechanically maintained [5]. Apart from the requirement for the ballast, the operability of tamping machines also specifies distances between sleepers. To avoid further disturbance to the regular tamping works, the number of sleepers is limited to three in this work. As shown in Table 4, the range of variables $(x_{sj})$ is adapted to the FE-based simulation environment, where the variables are discrete values rather than continuous (see Eq. (5)) to align the optimization with the discretization of the FEM. The value of $x_{sj}$ implies the number of finite elements, and each represents a 0.1m-long rail element, where the lower bound ($x_{sj} = 3$) refers to the smallest sleeper span (0.3 m), and the upper bound ($x_{sj} = 7$) defines the largest span as 0.7 m.

The last type of variables concerns the length of Type II rail strip in the embedded rail channel. Again the range of definition is adapted to the FE simulation environment. The lower bound ($x_l = 0$) implies that Type II strip is not applied in the channel. The upper bound ($x_l = 120$, with each element in a length of 0.1 m) corresponds to the full length (12 m) of the example level crossing, meaning Type II strip is applied in full.

The general design vector $\mathbf{x}$, given in Eq. (5), is reduced to $\mathbf{x} = \left[x_{r1}, x_{r2}, x_{r3}, x_n, x_{s1}, x_{s2}, x_{s3}, x_l\right]^T$ in the studied case. In single-objective problems, the general objective function $f(\mathbf{x})$ is defined by each of the performance measures, i.e., $F_{rms}(\mathbf{x}), F_{max}(\mathbf{x}), E_{rms}(\mathbf{x})$ and $E_{max}(\mathbf{x})$.

As previously mentioned, termination depends more upon the computational budget in the surrogate-assisted optimization problems [74,75]. The present work specifies a fixed budget of 200 function evaluations in the single objective problems to assess the effectiveness of the proposed objectives and identify appropriate ones for the follow-up search process. The statistical metrics, namely, *rms* and *max* values, are compared, and those showing higher solution quality and sensitivity to the parametric variation are chosen to formulate the MOO problem. A fixed budget of 400 function evaluations is assigned to the MOO problem to guarantee that the desired improvement level can be achieved.

In the MOO example, the scalarized global function, Eq. (7), contains two objectives, i.e., $f_i$ $(i = 1,2)$. The parameter $\rho = 0.05$. The normalization vector is calculated by $\kappa_i = \frac{1}{f_i^{max} - f_i^r}$ $(f_i^r \le f_i \le f_i^{max}, i = 1,2)$, with $f_i^r$ and $f_i^{max}$ being the reference (ideal) point and worst value obtainable for objective $i$, respectively. The ideal point defines the desired improvement level for each objective. Objectives are set of equal importance and the desired level is defined as a 20% improvement from the current reference design. The worst value, $f_i^{max}$, is determined according to the results simulated from the single-objective problems.

## 4. Results and discussion

### 4.1. Single-objective optimization

The detailed processes of minimizing force ($F$)-related and energy ($E$)-related measures are presented in Figs. 6 and 7, respectively, where the role of surrogate modeling and adaptive learning can be observed. All black triangles and dots represent sample points. Each point corresponds to a specific design solution $\mathbf{x}$ with eight variables ($\mathbf{x} = \left[x_{r1}, x_{r2}, x_{r3}, x_n, x_{s1}, x_{s2}, x_{s3}, x_l\right]^T$), which is evaluated by the objective or true function (i.e., the FE model). The objective function values shown in Figs. 6 and 7 (the vertical axes) represent response quantities simulated from the FE model, which is dependent on $\mathbf{x}$ (the design alternatives).

In the surrogate construction phase, the first sequence of 20 points is used to build a cubic RBF, which are indicated by a pink diamond (as the initial point) and 19 black triangles (the random samples generated from the experimental design). The pink diamond, as noted by the legend 'Initial Samples' in the figures, represents the point(s) that is(are) specified beforehand, which in the studied case is the reference design. It means that all the optimization problems share the same starting point for comparison purposes. Besides, the size of the random samples generated in the vicinity of the starting point is determined referring to Regis and Shoemaker [72]. It highlights the measure of $\min(2d, 10)$ with $d$ being the problem dimensions to start a simulated annealing algorithm. The studied case has dimensions $d = 8$, and the sample size is considered sufficient for the initial surrogate construction.

In the adaptive learning phase, i.e., after evaluation number 20, samples are generated for surrogate updating and optimum search. Both black dots ('Adaptive Samples') and triangles ('Random Samples') are shown in this phase. The adaptive samples are those found through random perturbations, and the random samples are searched by the Branch-and-Bound approach, as discussed in Section 2.3.2. All these points are evaluated by the FE model, and the thick green line keeps track of the best value found among all evaluated points. The current problems are formulated as minimization so that the best value implies the lowest objective function value (i.e., the minimum of the $F$- and $E$-related response quantities).

Specifically, Fig. 6(a) shows a noticeable drop in the best value near evaluation number 50. Then the best point slowly drops in value and becomes stable after evaluation number 130. In Fig. 6(b), a rapid change in value is captured at the initial surrogate construction phase, whereas it shows little improvement afterward. Fig. 7(a) shows a trend similar to Fig. 6(a), where the search process becomes stable around the evaluation number $130 \sim 140$.

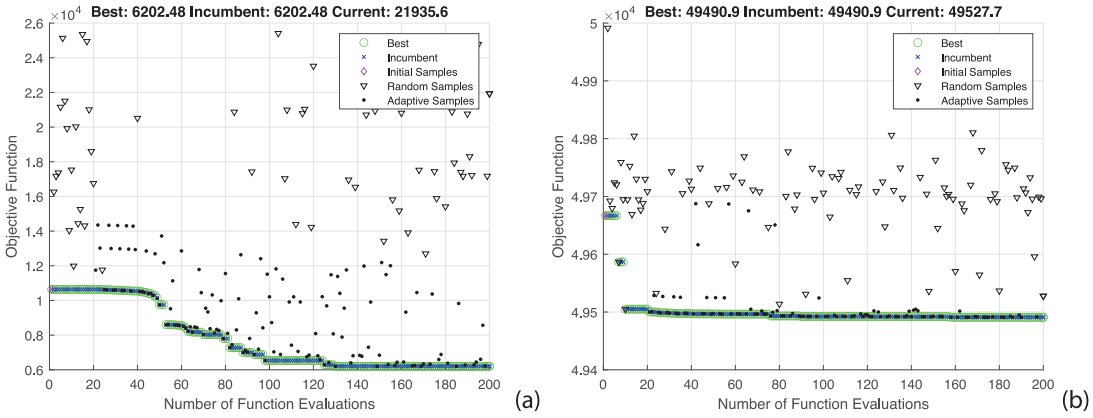



**Fig. 6.** Search process of optimizing objective $F_{max}$ (a) and $F_{rms}$ (b).

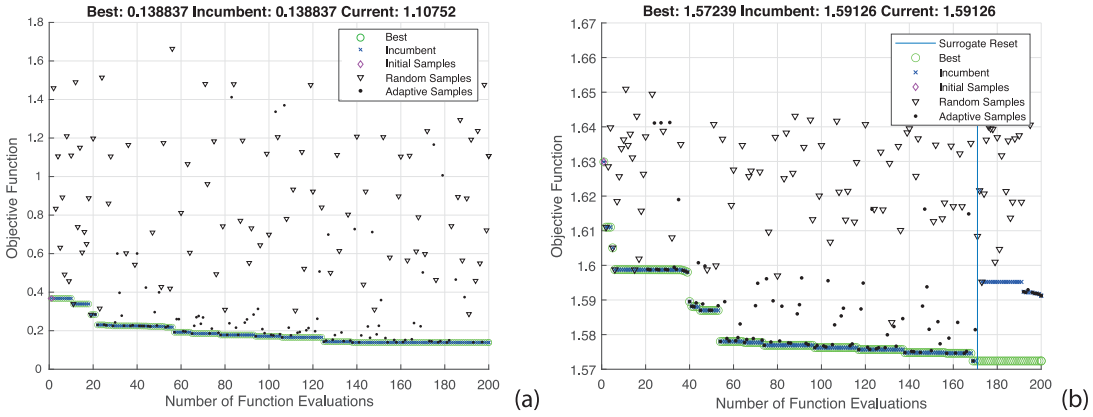



**Fig. 7.** Search process of optimizing objective $E_{max}$ (a) and $E_{rms}$ (b).

The solver may get stuck in a local minimum, which is measured by a distance-based tolerance parameter. In this case, the surrogate will be reset and returned to a new construction phase. While it may seem better to retain the past optimization trajectory in the next run, Regis and Shoemaker [72] suggested otherwise, based on the computational experience. This is because the past trajectory may bias the selection of candidate points towards the previously found optimum. The incumbent value, represented by the blue cross in the figures, is for tracking the lowest objective function value since the recent surrogate reset. In the process of optimizing the objective $E_{rms}$, a surrogate reset has occurred, which is indicated in Fig. 7(b) by a vertical straight line. It means that the evaluated points are tightly clustered around the best point, and a reset is needed to avoid getting trapped in the local optimum. For comparison purposes, further function evaluations are not assigned, where all the single-objective problems presented share the same starting point and computational budget. Besides, by comparing the scales of objective values, it can be observed from both figures that the *rms*-based values are less sensitive to changes in parameters, implying the statistical metric *rms* may not be a suitable indicator for optimization purposes.

Apart from investigating the search process, the solutions of the single-objective optimization and reference case are collected in Tables 5 and 6. Table 5 indicates the force ($F$)-related measures in general suggest higher railpad stiffness than the energy ($E$)-related measures, especially for $F_{rms}$, where the optimized values of variable $x_{ri}, (i = 1, 2, 3)$ almost reach the predefined upper bound. This can be explained by the fact that stiffer railpads contribute to reducing vibration from the wheel–rail contact (quantified by the $F$-related measures); however, this may lead to a higher effect of loads transmitted to underlayers, thereby causing vibration in sleepers and ballast (quantified by the $E$-related measures).

**Table 5**
Design solutions from single-objective optimization.

| Design variables | Solution (reference) | Solution (objective functions) | | | |
|---|---|---|---|---|---|
| | | $F_{rms}$ | $F_{max}$ | $E_{rms}$ | $E_{max}$ |
| $x_{r1}$ [MN/m] | 1120 | 1199 | 103 | 59 | 50 |
| $x_{r2}$ [MN/m] | 1120 | 1199 | 50 | 62 | 77 |
| $x_{r3}$ [MN/m] | 1120 | 1199 | 72 | 50 | 50 |
| $x_n$ | 8 | 0 | 4 | 0 | 1 |
| $x_{s1}$ [# of 0.1-m FE] | 6 | 5 | 6 | 6 | 7 |
| $x_{s2}$ [# of 0.1-m FE] | 6 | 7 | 6 | 3 | 5 |
| $x_{s3}$ [# of 0.1-m FE] | 4 | 4 | 4 | 7 | 4 |
| $x_l$ [# of 0.1-m FE] | 0 | 120 | 0 | 0 | 0 |

Besides, compared with the reference design, all the solutions suggest a limited number of strengthened sleepers (variable $x_n$) used in the transition. As for the spacing (variable $x_{si}, i = 1, 2, 3$), except for $E_{rms}$, the other objectives all suggest the smallest distance ($x_{s3}$) between the structural interface and adjacent sleeper. The reduction in support distances implies an increase in local track stiffness [26]. Therefore, the result indicates that a relatively high support stiffness should be assigned near the structural interface to ensure a smooth transition.

In terms of the last variable $x_l$, the minimization of $F_{rms}$ results in a solution of applying Type II rail strip along the full length of the embedded track. It is questionable as it was found in the reference case that the ballast track with the FC9 pads is stiffer than the ERS; applying the softer Type II strip in the ERS may cause more dynamic amplification in the structure. This finding can be justified by the consistent solutions (i.e., $x_l = 0$) obtained from optimizing the other objectives, as listed in Table 5.

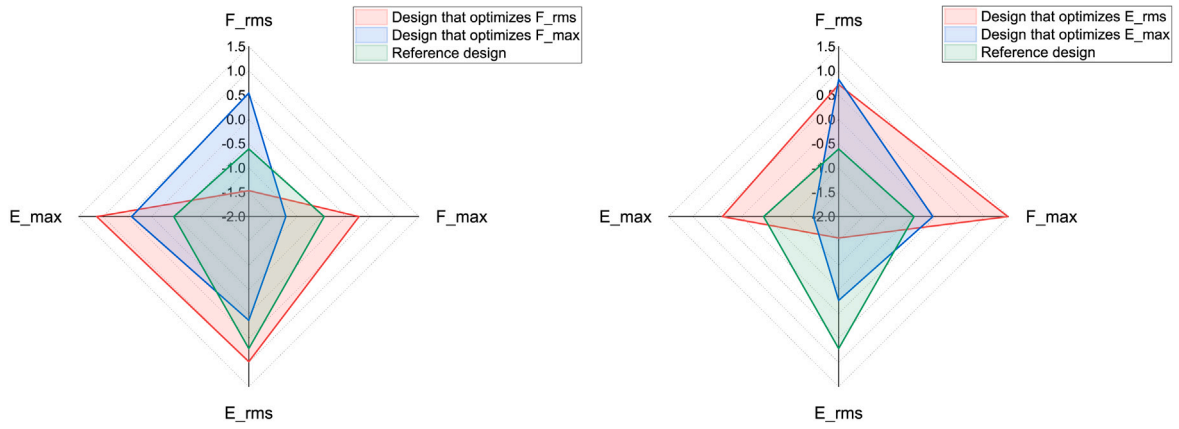



**Fig. 8.** Comparison of metric *rms* and *max* in terms of solution quality. For representative purposes, the values are normalized with respect to objective values obtained from single-objective problems.

**Table 6**
Objective values in single-objective optimization. Minimum (optimal) values are highlighted for each optimization problem.

| | $F_{rms}$ [N] | $F_{max}$ [N] | $E_{rms}$ [N·m] | $E_{max}$ [N·m] |
|---|---|---|---|---|
| Design that optimizes $F_{rms}$ | **49491** | 14683 | 1.6367 | 0.7222 |
| Design that optimizes $F_{max}$ | 49903 | **6263** | 1.6153 | 0.5620 |
| Design that optimizes $E_{rms}$ | 49939 | 21508 | **1.5724** | 0.5575 |
| Design that optimizes $E_{max}$ | 49961 | 12814 | 1.6049 | **0.1388** |
| Reference design | 49667 | 10636 | 1.6299 | 0.3676 |

Table 6 maps the performance of the optimal designs compared to the reference design. It can be seen that the *rms*-based values, listed in column $F_{rms}$ and $E_{rms}$, are less sensitive to parametric variation, which is in line with the findings in Figs. 6 and 7. Moreover, no single solution exists that simultaneously optimizes each objective, and the conflicting nature of the objectives $F_{max}$ and $E_{max}$ can be captured from this table. The design that optimizes $F_{max}$ (second row) has associated very large values of $E_{max}$. On the contrary, the design that optimizes $E_{max}$ (fourth row) has associated very large values of $F_{max}$.

To visually evaluate the performance of *rms* and *max* metrics, Fig. 8 compares the performance of the designs obtained from optimizing $F$-related measures (a) and $E$-related measures (b), where the data are normalized beforehand based on Table 6 considering different scales used in the objectives. It can be seen from Fig. 8(a) that, based on the reference design, optimizing $F_{max}$ results in a reduction of two objective values, i.e., $F_{max}$ and $E_{rms}$, whereas taking $F_{rms}$ as the objective contributes only to improving itself. Likewise, Fig. 8(b) indicates that optimizing $E_{max}$ can effectively improve $E_{max}$ and $E_{rms}$ without degrading other objectives too much, as observed in the case of optimizing $E_{rms}$. It can be concluded from this that the *max* metric in general performs better than the other in terms of the solution quality and sensitivity to the parameter change and therefore is more suitable for parametric optimization.

### 4.2. Multi-objective optimization

The evaluation of the single-objective problems in the previous section indicate that the metric *max* performs better than the other in terms of solution quality and sensitivity to the parametric variation. Fig. 9 compares normalized objective values calculated from minimizing $F_{max}$ and $E_{max}$. The reference design is used as a benchmark, where the conflicting nature between the two objectives can be seen, suggesting simultaneous optimization of the objectives.

The scalarized global function (Eq. (7)) is formulated by two objectives ($F_{max}$ and $E_{max}$). The process of minimizing the global function is presented in Fig. 10, which shows that the search process becomes

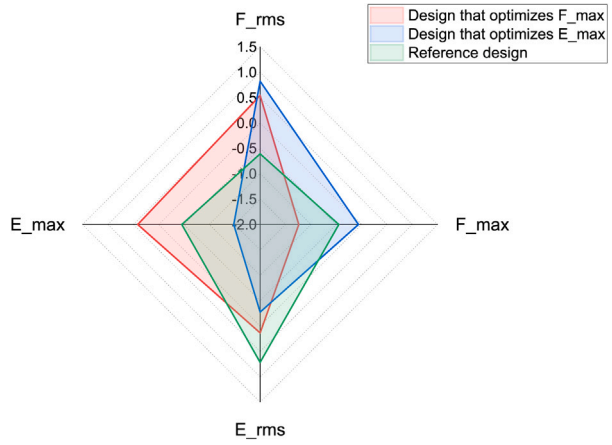



**Fig. 9.** Comparison of objective $F_{max}$ and $E_{max}$ in terms of solution quality (normalized values; see also Fig. 8).

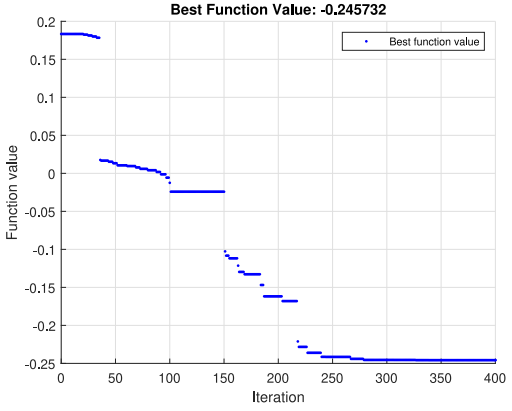



**Fig. 10.** Process of simultaneous optimization of $F_{max}$ and $E_{max}$.

stable after evaluation number 280. As previously mentioned, 400 function evaluations are assigned to the MOO problem, where an additional computational budget is allocated compared with the single-objective cases. The reason is that the main purpose of the single-objective problems is to assess the effectiveness of the candidate objectives and identify appropriate ones for the follow-up search. In contrast, the MOO problem is formulated to search for the optimum while guaranteeing that the desired improvement level can be achieved.

**Table 7**
Design solutions and objective values for different optimization problems.

| Objectives | Design vector $\mathbf{x}$ $[x_{r1}, x_{r2}, x_{r3}, x_n, x_{s1}, x_{s2}, x_{s3}, x_l]'$ | $F_{max}$ [N] | $E_{max}$ [N·m] |
|---|---|---|---|
| Min $F_{max}$ | $[103, 50, 72, 4, 6, 6, 4, 0]'$ | 6202.5 | 0.5620 |
| Min $E_{max}$ | $[50, 77, 50, 1, 7, 5, 4, 0]'$ | 12814 | **0.1388** |
| Reference design | $[1199, 1199, 1199, 8, 6, 6, 4, 0]'$ | 10636 | 0.3676 |
| MOO ($F_{max}$, $E_{max}$) | $[214, 155, 50, 5, 6, 7, 3, 0]'$ | **5662.6** | 0.1919 |

As mentioned in Eq. (7), the ideal point $f_i^r (i = 1, 2)$ is the desired improvement level, which is defined as a 20% improvement from the reference design. This represents that the target values of $F_{max}$ and $E_{max}$ are 8508.5 N and 0.2941 N·m, respectively. Table 7 lists the obtained design solutions and corresponding objective values from single-objective and MOO problems. It can be observed that the optimized function values from the MOO problem (i.e., 5662.6 N for $F_{max}$ and 0.1919 N·m for $E_{max}$) are below the targets, which means the optimization reaches the desired improvement level. The result shows 46.8% ($F_{max}$) and 47.8% ($E_{max}$) improvement from the reference design, which is beyond the target improvement. The value of $F_{max}$ is even better than the single-objective problem, demonstrating how complex and nonlinear the objective function (simulated from the FE model) is. The solver may get trapped in regions containing a local optimum, and some searches may provide better solutions (like the current MOO case).

As for the design solutions, it can be seen in Table 7 that the major difference between the reference design and MOO-based design lies in the distribution of the railpad stiffness. Compared with the reference, the solution indicates that applying soft pads in the vicinity of the structural interface contributes to a reduction in both objectives. It is reasonable, according to experience (literature and practice), that track geometry degradation (typically the ballast settlement) is the major concern in the connecting area. Vibration isolation and attenuation of dynamic loads are typically demanding, and soft pads are helpful from this perspective. It also agrees well with the suggestion in Ref. [16,87], where soft pads are recommended in general. Meanwhile, the solution suggests a decreasing trend in pad stiffness where the softest pad should be placed right next to the interface. It is the weakest spot that is susceptible to differential settlement.

Moreover, compared with the reference, the use of fewer strengthened sleepers in the transition, from 8 to 5, is shown preferable since more strengthened ones may make the ballast track even stiffer. However, the solution still suggests a few numbers, which might be helpful in vibration isolation for the underlayers (typically the ballast) due to the larger size and improved stability. For comparison, Fig. 11 shows a profile of energy dissipation under each sleeper in the MOO-based and reference scenarios. In total, 20 sleepers are presented, corresponding to the upper bound of variable $X_n$. Sleeper number 1 is the one right next to the structural interface. It can be seen that near the interface, the energy dissipated in MOO is lower than that in the reference case, indicating an expected reduction in ballast degradation and the consequent impact on local track geometry.

The sleeper span suggested by MOO does not change much with the reference design. A shorter distance is generally recommended between the structural interface and the first sleeper to ensure a smooth transition in ballast-to-ERS connections. The last variable indicates that the current Type I rail strip is preferable to the alternative as the former is stiffer and can balance the arrangement of track support stiffness between different track forms.

Note that the optimized results are influenced by the default parameter setting (Table 4). For instance, the railpad stiffness comes in a wide range of values. The stiff FC9 type is considered in the default setting, which makes the ballast track stiffer than the ERS structure. In this case, the solver guides the search towards the default (i.e., Type I) rail strip used in the ERS to guarantee a smooth stiffness transition.

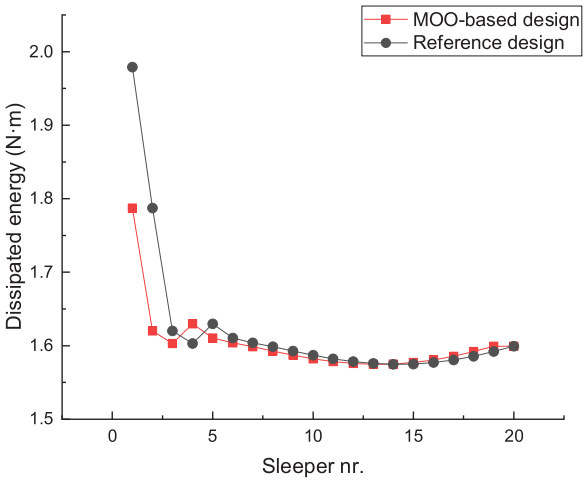



**Fig. 11.** Comparison of optimized and reference design in terms of energy dissipation in track substructure.

The use of other pads in ballast track may however make the ERS structure stiffer than the other, and this may result in different design solutions, e.g., probably Type II rail strip will be a preferred option. The case discussed is for demonstration purposes, and the default parameter setting refers to the Dutch standard practice.

## 5. Conclusions

Optimization on the railway track and or vehicle systems often requires dynamic simulations of vehicle–track coupling dynamics, as by nature the two systems are interactive during the train passage. Numerical models for simulating the coupling dynamics have become more complex and accurate, but the computational time for running these models has not necessarily been reduced. It makes it difficult to complete engineering tasks that rely on these models, such as design space exploration and optimization.

The present work screens the issue in design optimization of railway transition zones, as they are one of the weakest links in railway network infrastructure due to higher degradation rates and maintenance needs. Specifically, a vehicle–track interaction (VTI) model is developed to simulate the dynamic behavior of a level crossing, where an embedded rail structure is installed in the crossing with connections to the conventional ballast track. Based on this, an integrative simulation methodology is proposed that embeds the VTI model in the adaptive modeling scheme to facilitate the efficient exploration of the design space. The goal is to optimize the local track performance by varying geometry and elastic properties in track structures. The mitigation of track degradation may further reduce maintenance needs and the consequent impact on, e.g., system life-cycle cost and network performance.

Four different objectives are proposed to capture the track structural performance from wheel–rail contact and lower supporting level. Several optimization problems are formulated accordingly, where single-objective problems are for comparing the effectiveness of the candidate objectives, and the MOO problem is defined to search for the optimal compromise solution. Compared with the reference design, the MOO-based solution shows significant improvement in the most relevant objectives (46.8% in $F_{max}$ and 47.8% in $E_{max}$). It can be considered the best-known solution to support relevant decision-making in track preliminary design. While 400 function evaluations are assigned to obtain the solution, more computational resources can be allocated for the proposed methodology to search for (possibly) better solutions. Besides the final MOO solution, evaluating the other solutions is meant

to inform and improve the current transition design practice, especially in areas using the embedded rail structures.

The duration of optimization is based on the number of function evaluations. Considering the (general) required function evaluations (> 1000) in evolutionary algorithms, the present methodology shows a good balance between the quality of the solutions and computational time (a few hundred function evaluations). This would allow models with a greater computational expense to be used, e.g., 3D track models, which also points out the limitations of the current work. The ballast and underlying substructure are simplified as the Winkler-type foundation, where the nonlinear structural and material behavior are not considered. Also, a 3D model for the embedded rail structure might be preferable as it provides homogeneous continual rail support, where wheel loads are distributed more evenly and that influences the vehicle–track dynamics in all three directions.

The present work focuses on formulating surrogate-based adaptive modeling in railway track design, where the methodology is demonstrated in a simplified vehicle–track model. It offers the possibility of adding a more complex model in terms of, e.g., track configurations, structural elements, and track–soil coupling effect, in the methodology, which is currently under development. A single simulation run of those models would be more computationally intensive. However, considering the total computational costs (used for repetitive calls of function evaluations), more significant savings can be expected. The reason is that the current methodology requires fewer function evaluations than the common optimization methods used in the railway field (e.g., genetic algorithms), where the former is well-suited to deal with (expensive) simulation-based mixed-integer optimization problems.

## CRediT authorship contribution statement

**Yue Shang:** Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft. **Maria Nogal:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Rui Teixeira:** Conceptualization, Methodology, Writing – review & editing. **A.R. (Rogier) M. Wolfert:** Conceptualization, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The authors do not have permission to share data.

## Appendix. Formulation of vehicle–track system matrices

The system matrices of a 3-DOF vehicle system are given by

$$\mathbf{M}_v = \mathrm{diag}\begin{bmatrix} m_c & m_b & m_w \end{bmatrix} \tag{A.1}$$

$$\mathbf{C}_v = \begin{bmatrix} c_{s2} & -c_{s2} & 0 \\ -c_{s2} & c_{s2}+2c_{s1} & -c_{s1} \\ 0 & -c_{s1} & c_{s1} \end{bmatrix} \tag{A.2}$$

$$\mathbf{K}_v = \begin{bmatrix} k_{s2} & -k_{s2} & 0 \\ -k_{s2} & k_{s2}+2k_{s1} & -k_{s1} \\ 0 & -k_{s1} & k_{s1} \end{bmatrix} \tag{A.3}$$

where $m_c$, $m_b$, and $m_w$ denote the mass of the carbody, bogie, and wheelset. $c_{s1}$ and $c_{s2}$ represent damping properties of the primary and secondary suspension; $k_{s1}$ and $k_{s2}$ denote stiffness of the corresponding suspension systems.

As for the coupled equations of motion (Eqs. (3a) and (3b)), the details of formulating $\mathbf{K}_g$ and $\mathbf{F}_g$ are shown below.

The vehicle stiffness matrix $\mathbf{K}_{vv}$ can be expressed as

$$\mathbf{K}_{vv} = \mathbf{K}_v + \mathbf{K}_v' \tag{A.4}$$

where $\mathbf{K}_v$ is the stiffness matrix of the vehicle itself, as shown in Eq. (A.3). $\mathbf{K}_v'$ is the stiffness matrix of the vehicle induced by the wheel–rail contact. For the current 3-DOF vehicle, it is written as

$$\mathbf{K}_v' = \mathrm{diag}\begin{bmatrix} 0 & 0 & k_w \end{bmatrix} \tag{A.5}$$

$\mathbf{K}_{tt}$ is the track overall stiffness matrix and is written as

$$\mathbf{K}_{tt} = \mathbf{K}_t + \mathbf{K}_t' \tag{A.6}$$

with

$$\mathbf{K}_t' = k_w \cdot \mathbf{N}^{\mathrm{T}} \cdot \mathbf{N} \tag{A.7}$$

where $\mathbf{K}_t$ represents the stiffness matrix of the track itself; $\mathbf{K}_t'$ is related to the rail displacement under the wheel, representing a portion of the track stiffness matrix induced by the vehicle. $\mathbf{N}$ is a location vector that defines the correspondence between the vehicle position and rail element in contact. For one wheel–rail contact point, the vector $\mathbf{N}$ is given by

$$\mathbf{N} = \begin{bmatrix} 0 & \dots & 0 & N_{1j} & N_{2j} & N_{3j} & N_{4j} & 0 & \dots & 0 \end{bmatrix}'_{n\times 1} \tag{A.8a}$$

with $n$ denoting the total number of DOFs of the track structure. The non-zero entries of $\mathbf{N}$ represent the rail element in contact with the wheel, where Hermitian shape functions for the Euler–Bernoulli beam are applied. By letting $\xi_j \in [0, l_j]$ be the local coordinate of rail element $j$ with length $l_j$, the non-zero entries of $\mathbf{N}$ can be expressed as

$$N_{1j}(\xi_j) = 1 - 3\left(\frac{\xi_j}{l_j}\right)^2 + 2\left(\frac{\xi_j}{l_j}\right)^3 \tag{A.8b}$$

$$N_{2j}(\xi_j) = \xi_j\left(1 - \frac{\xi_j}{l_j}\right)^3 \tag{A.8c}$$

$$N_{3j}(\xi_j) = 3\left(\frac{\xi_j}{l_j}\right)^2 - 2\left(\frac{\xi_j}{l_j}\right)^3 \tag{A.8d}$$

$$N_{4j}(\xi_j) = \frac{\xi_j^2}{l_j}\left(\frac{\xi_j}{l_j} - 1\right) \tag{A.8e}$$

$\mathbf{K}_{tv}$ and $\mathbf{K}_{vt}$ are coupling matrices induced by the wheel–rail interaction and given by

$$\mathbf{K}_{tv} = \begin{bmatrix} 0 & 0 & -k_w\mathbf{N} \end{bmatrix}_{n\times 3}, \quad \mathbf{K}_{vt} = \mathbf{K}_{vt}' \tag{A.9}$$

At the right-hand side of Eq. (3a), the global force vector $\mathbf{F}_g$ is formulated by two coupled load vectors, i.e., $\mathbf{F}_{vt}$ and $\mathbf{F}_{tv}$. $\mathbf{F}_{vt}$ is the coupled load vector of the vehicle and is given by

$$\mathbf{F}_{vt} = \mathbf{F}_{vg} + \mathbf{F}_{vr} \tag{A.10}$$

with $\mathbf{F}_{vg}$ denoting the load vector induced by the vehicle gravity and $\mathbf{F}_{vr}$ being the load vector induced by the track geometry irregularity. They can be calculated by

$$\mathbf{F}_{vg} = \begin{bmatrix} m_c g & m_b g & m_w g \end{bmatrix}' \tag{A.11}$$

$$\mathbf{F}_{vr} = \begin{bmatrix} 0 & 0 & k_w r(x) \end{bmatrix}' \tag{A.12}$$

where $g$ is gravitational acceleration, i.e., 9.8 m/s$^2$. $r(x)$ denotes the track irregularity in vertical profile, which is location specific.

$\mathbf{F}_{tv}$ is the coupled load vector of the track and is written as

$$\mathbf{F}_{tv} = -k_w r(x) \cdot \mathbf{N} \tag{A.13}$$

## References


[1] Yang Z, Zhang P, Wang L. Wheel-rail impact at an insulated rail joint in an embedded rail system. Eng Struct 2021;246:113026.
[2] Wang H, Markine V. Dynamic behaviour of the track in transitions zones considering the differential settlement. J Sound Vib 2019;459:114863.
[3] Stančík V, Ryjáček P, Vokáč M. Thermal and load rate-dependent interaction between embedded rail system and bridge. Proc Inst Mech Eng F 2019;233(3):326–36.
[4] Ling L, Han J, Xiao X, Jin X. Dynamic behavior of an embedded rail track coupled with a tram vehicle. J Vib Control 2017;23(14):2355–72.
[5] Le Pen L, Watson G, Powrie W, Yeo G, Weston P, Roberts C. The behaviour of railway level crossings: Insights through field monitoring. Transp Geotech 2014;1(4):201–13.
[6] Indraratna B, Sajjad MB, Ngo T, Correia AG, Kelly R. Improved performance of ballasted tracks at transition zones: A review of experimental and modelling approaches. Transp Geotech 2019;21:100260.
[7] Paixão A, Fortunato E, Calçada R. The effect of differential settlements on the dynamic response of the train–track system: A numerical study. Eng Struct 2015;88:216–24.
[8] Fărăgău AB, Metrikine AV, van Dalen KN. Transition radiation in a piecewise-linear and infinite one-dimensional structure–a Laplace transform method. Nonlinear Dynam 2019;98(4):2435–61.
[9] Vesnitskii A, Metrikin A. Transition radiation in one-dimensional elastic systems. J Appl Mech Tech Phys 1992;33(2):202–7.
[10] Metrikine A, Wolfert ARM, Dieterman H. Transition radiation in an elastically supported string. Abrupt and smooth variations of the support stiffness. Wave Motion 1998;27(4):291–305.
[11] Vesnitskii A, Metrikin A. Transition radiation in mechanics. Phys-Usp 1996;39(10):983.
[12] Dimitrovová Z, Varandas J. Critical velocity of a load moving on a beam with a sudden change of foundation stiffness: Applications to high-speed trains. Comput Struct 2009;87(19–20):1224–32.
[13] Ang KK, Dai J. Response analysis of high-speed rail system accounting for abrupt change of foundation stiffness. J Sound Vib 2013;332(12):2954–70.
[14] van Dalen KN, Tsouvalas A, Metrikine AV, Hoving JS. Transition radiation excited by a surface load that moves over the interface of two elastic layers. Int J Solids Struct 2015;73:99–112.
[15] Fărăgău AB, Mazilu T, Metrikine AV, Lu T, van Dalen KN. Transition radiation in an infinite one-dimensional structure interacting with a moving oscillator—the Green's function method. J Sound Vib 2021;492:115804.
[16] Sadri M, Lu T, Steenbergen M. Railway track degradation: The contribution of a spatially variant support stiffness-local variation. J Sound Vib 2019;455:203–20.
[17] Shang Y, van den Boomen M, de Man A, Wolfert ARM. Reliability-based life cycle costing analysis for embedded rails in level crossings. Proc Inst Mech Eng F 2019;234(8):821–33.
[18] Paixão A, Fortunato E, Calçada R. Transition zones to railway bridges: track measurements and numerical modelling. Eng Struct 2014;80:435–43.
[19] Mishra D, Boler H, Tutumluer E, Hou W, Hyslip JP. Deformation and dynamic load amplification trends at railroad bridge approaches: effects caused by high-speed passenger trains. Transp Res Rec 2017;2607(1):43–53.
[20] Varandas J, Paixão A, Fortunato E. A study on the dynamic train-track interaction over cut-fill transitions on buried culverts. Comput Struct 2017;189:49–61.
[21] Zuada Coelho B, Hicks MA. Numerical analysis of railway transition zones in soft soil. Proc Inst Mech Eng F 2016;230(6):1601–13.
[22] Milne D, Harkness J, Le Pen L, Powrie W. The influence of variation in track level and support system stiffness over longer lengths of track for track performance and vehicle track interaction. Veh Syst Dyn 2021;59(2):245–68.
[23] Galvín P, Romero A, Domínguez J. Fully three-dimensional analysis of high-speed train–track–soil-structure dynamic interaction. J Sound Vib 2010;329(24):5147–63.
[24] Mishra D, Qian Y, Huang H, Tutumluer E. An integrated approach to dynamic analysis of railroad track transitions behavior. Transp Geotech 2014;1(4):188–200.
[25] Varandas JN, Hölscher P, Silva MA. Settlement of ballasted track under traffic loading: application to transition zones. Proc Inst Mech Eng F 2014;228(3):242–59.
[26] Sañudo R, Cerrada M, Alonso B, dell'Olio L. Analysis of the influence of support positions in transition zones. A numerical analysis. Constr Build Mater 2017;145:207–17.
[27] Ngamkhanong C, Ming QY, Li T, Kaewunruen S. Dynamic train-track interactions over railway track stiffness transition zones using baseplate fastening systems. Eng Fail Anal 2020;118:104866.
[28] Zakeri J-A, Ghorbani V. Investigation on dynamic behavior of railway track in transition zone. J Mech Sci Technol 2011;25(2):287–92.
[29] Wan C, Markine V, Shevtsov I. Improvement of vehicle–turnout interaction by optimising the shape of crossing nose. Veh Syst Dyn 2014;52(11):1517–40.
[30] Pålsson BA, Nielsen JC. Track gauge optimisation of railway switches using a genetic algorithm. Veh Syst Dyn 2012;50(sup1):365–87.
[31] Nicklisch D, Kassa E, Nielsen J, Ekh M, Iwnicki S. Geometry and stiffness optimization for switches and crossings, and simulation of material degradation. Proc Inst Mech Eng F 2010;224(4):279–92.
[32] Wan C, Markine V, Shevtsov I. Optimisation of the elastic track properties of turnout crossings. Proc Inst Mech Eng F 2016;230(2):360–73.
[33] Li W, Pu H, Schonfeld P, Zhang H, Zheng X. Methodology for optimizing constrained 3-dimensional railway alignments in mountainous terrain. Transp Res C 2016;68:549–65.
[34] Oyarzabal O, Gomez J, Santamaria J, Vadillo E. Dynamic optimization of track components to minimize rail corrugation. J Sound Vib 2009;319(3–5):904–17.
[35] Alkhatib R, Jazar GN, Golnaraghi MF. Optimal design of passive linear suspension using genetic algorithm. J Sound Vib 2004;275(3–5):665–91.
[36] Aggestam E, Nielsen JC. Multi-objective optimisation of transition zones between slab track and ballasted track using a genetic algorithm. J Sound Vib 2019;446:91–112.
[37] Namura N, Shimoyama K, Obayashi S. Expected improvement of penalty-based boundary intersection for expensive multiobjective optimization. IEEE Trans Evol Comput 2017;21(6):898–913.
[38] Müller J. MISO: mixed-integer surrogate optimization framework. Opt Eng 2016;17(1):177–203.
[39] Shen C, Dollevoet R, Li Z. Fast and robust identification of railway track stiffness from simple field measurement. Mech Syst Signal Process 2021;152:107431.
[40] Shamalta M, Metrikine A. Comparison of the dynamic response of one-and two-dimensional models for an embedded railway track to a moving load. Heron-Engl. Ed. 2002;47(4):243–62.
[41] Xia H, Zhang N. Dynamic analysis of railway bridge under high-speed trains. Comput Struct 2005;83(23–24):1891–901.
[42] Aggestam E, Nielsen JC, Bolmsvik R. Simulation of vertical dynamic vehicle–track interaction using a two-dimensional slab track model. Veh Syst Dyn 2018;56(11):1633–57.
[43] Lei X, Wu S, Zhang B. Dynamic analysis of the high speed train and slab track nonlinear coupling system with the cross iteration algorithm. J Nonlinear Dynam 2016;2016.
[44] Lei X. Model for vertical dynamic analysis of the vehicle-track coupling system. In: High speed railway track dynamics. Springer; 2017, p. 161–99.
[45] EN 16432-2. Railway applications - Ballastless track systems - Part 2: System design, subsystems and components. Standard, European Committee for Standardization; 2017.
[46] Sol-Sánchez M, Moreno-Navarro F, Rubio-Gámez MC. The use of elastic elements in railway tracks: A state of the art review. Constr Build Mater 2015;75:293–305.
[47] Sadri M, Steenbergen M. Effects of railway track design on the expected degradation: Parametric study on energy dissipation. J Sound Vib 2018;419:281–301.
[48] Shahraki M, Warnakulasooriya C, Witt KJ. Numerical study of transition zone between ballasted and ballastless railway track. Transp Geotech 2015;3:58–67.
[49] Zakeri JA, Xia H. Sensitivity analysis of track parameters on train-track dynamic interaction. J Mech Sci Technol 2008;22(7):1299–304.
[50] Sedghi M, Kauppila O, Bergquist B, Vanhatalo E, Kulahci M. A taxonomy of railway track maintenance planning and scheduling: A review and research trends. Reliab Eng Syst Saf 2021;215:107827.
[51] Steenbergen M, Metrikine A, Esveld C. Assessment of design parameters of a slab track railway system from a dynamic viewpoint. J Sound Vib 2007;306(1–2):361–71.
[52] Sadri M, Lu T, Steenbergen M. Railway track degradation: The contribution of a spatially variant support stiffness-Global variation. J Sound Vib 2020;464:114992.
[53] Chugh T. Scalarizing functions in Bayesian multiobjective optimization. In: 2020 IEEE congress on evolutionary computation. IEEE; 2020, p. 1–8.
[54] Wierzbicki AP. The use of reference objectives in multiobjective optimization. In: Multiple criteria decision making theory and application. Springer; 1980, p. 468–86.
[55] Miettinen K. Nonlinear multiobjective optimization, vol. 12. Springer Science & Business Media; 2012.
[56] Zhang X-D. A matrix algebra approach to artificial intelligence. Springer; 2020.
[57] Chiandussi G, Codegone M, Ferrero S, Varesio FE. Comparison of multi-objective optimization methodologies for engineering applications. Comput Math Appl 2012;63(5):912–42.
[58] Cheng Q, Wang S, Liu Z, Yuan Y. Surrogate-based simulation optimization approach for day-to-day dynamics model calibration with real data. Transp Res C 2019;105:422–38.
[59] Urquhart M, Ljungskog E, Sebben S. Surrogate-based optimisation using adaptively scaled radial basis functions. Appl Soft Comput 2020;88:106050.
[60] Han X, Xiang H, Li Y, Wang Y. Predictions of vertical train-bridge response using artificial neural network-based surrogate model. Adv Struct Eng 2019;22(12):2712–23.
[61] Li H, Wang T, Wu G. Probabilistic safety analysis of coupled train-bridge system using deep learning based surrogate model. Struct Infrastruct Eng 2021;1–20.
[62] Müller J, Shoemaker CA. Influence of ensemble surrogate models and sampling strategy on the solution quality of algorithms for computationally expensive black-box global optimization problems. J Global Optim 2014;60(2):123–44.

[63] Bacigalupo A, Gnecco G, Lepidi M, Gambarotta L. Computational design of innovative mechanical metafilters via adaptive surrogate-based optimization. Comput Methods Appl Mech Engrg 2021;375:113623.
[64] Buhmann MD. Radial basis functions: Theory and implementations, vol. 12. Cambridge University Press; 2003.
[65] Gutmann H-M. A radial basis function method for global optimization. J Global Optim 2001;19(3):201–27.
[66] Duan Y. A note on the meshless method using radial basis functions. Comput Math Appl 2008;55(1):66–75.
[67] Shi L, Yang R-J, Zhu P. An adaptive response surface method for crashworthiness optimization. Eng Optim 2013;45(11):1365–77.
[68] Roy S, Crossley WA, Stanford B, Moore KT, Gray JS. A mixed integer efficient global optimization algorithm with multiple infill strategy-applied to a wing topology optimization problem. In: AIAA Scitech 2019 forum. 2019, p. 2356.
[69] Liu Y, Li L, Zhao S, Song S. A global surrogate model technique based on principal component analysis and Kriging for uncertainty propagation of dynamic systems. Reliab Eng Syst Saf 2021;207:107365.
[70] Wang H, Ye F, Li E, Li G. A comparative study of expected improvement-assisted global optimization with different surrogates. Eng Optim 2016;48(8):1432–58.
[71] Garrido-Merchán EC, Hernández-Lobato D. Dealing with categorical and integer-valued variables in Bayesian optimization with Gaussian processes. Neurocomputing 2020;380:20–35.
[72] Regis RG, Shoemaker CA. A stochastic radial basis function method for the global optimization of expensive functions. INFORMS J Comput 2007;19(4):497–509.
[73] Davis E, Ierapetritou M. A kriging based method for the solution of mixed-integer nonlinear programs containing black-box functions. J Global Optim 2009;43(2):191–205.
[74] Picheny V, Gramacy RB, Wild S, Le Digabel S. Bayesian optimization under mixed constraints with a slack-variable augmented Lagrangian. Adv Neural Inf Process Syst 2016;29.
[75] Wang Y, Shoemaker CA. A general stochastic algorithmic framework for minimizing expensive black box objective functions based on surrogate models and sensitivity analysis. 2014, arXiv preprint arXiv:1410.6271.
[76] Shang Y, Nogal M, Wolfert ARM. A co-simulation solution for vehicle-track interaction dynamics problems. In: The fifth international conference on railway technology: Research, development and maintenance. Montpellier, France: Elsevier; 2022.
[77] Li Z, Wu T. On vehicle/track impact at connection between a floating slab and ballasted track and floating slab track's effectiveness of force isolation. Veh Syst Dyn 2009;47(5):513–31.
[78] Kuo C-M, Huang C-H, Chen Y-Y. Vibration characteristics of floating slab track. J Sound Vib 2008;317(3–5):1017–34.
[79] Ortega RS, Pombo J, Ricci S, Miranda M. The importance of sleepers spacing in railways. Constr Build Mater 2021;300:124326.
[80] De Man A. Dynatrack: A survey of dynamic railway track properties and their quality (Ph.D. thesis), Delft University of Technology; 2004.
[81] Esveld C. Modern railway track, vol. 385. MRT-productions Zaltbommel; 2001.
[82] Oregui M, Núnez A, Dollevoet R, Li Z. Sensitivity analysis of railpad parameters on vertical railway track dynamics. J Eng Mech 2017;143(5):04017011.
[83] Li T, Su Q, Kaewunruen S. Influences of dynamic material properties of slab track components on the train-track vibration interactions. Eng Fail Anal 2020;115:104633.
[84] Kaewunruen S, Remennikov A. Dynamic properties of railway track and its components: a state-of-the-art review. 2008.
[85] Ripke B, Knothe K. Simulation of high frequency vehicle-track interactions. Veh Syst Dyn 1995;24(sup1):72–85.
[86] Wang H, Markine V, Liu X. Experimental analysis of railway track settlement in transition zones. Proc Inst Mech Eng F 2018;232(6):1774–89.
[87] Hasan N. Rail pad stiffness and classification system. J Transp Eng A Syst 2019;145(5):04019012.