POLITECNICO DI TORINO
Repository ISTITUZIONALE

Kernel-based digital twins for evaluating the running safety of freight wagons during emergency braking

Original
Kernel-based digital twins for evaluating the running safety of freight wagons during emergency braking / Bosso, N., Cantone, L., Magelli, M., Trinchero, R., Zampieri, N.. - In: VEHICLE SYSTEM DYNAMICS. - ISSN 0042-3114. - (2026), pp. 1-29. [10.1080/00423114.2026.2692718]

Availability:
This version is available at: 11583/3016692 since: 2026-10-15T13:04:38Z

Publisher:
Taylor and Francis

Published
DOI:10.1080/00423114.2026.2692718

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026



Taylor & Francis
Taylor & Francis Group

Vehicle System Dynamics

International Journal of Vehicle Mechanics and Mobility

ISSN: 0042-3114 (Print) 1744-5159 (Online) Journal homepage: www.tandfonline.com/journals/nvsd20

Kernel-based digital twins for evaluating the running safety of freight wagons during emergency braking

N. Bosso , L. Cantone , M. Magelli , R. Trinchero & N. Zampieri

To cite this article: N. Bosso , L. Cantone , M. Magelli , R. Trinchero & N. Zampieri (29 Jun 2026): Kernel-based digital twins for evaluating the running safety of freight wagons during emergency braking, Vehicle System Dynamics, DOI: 10.1080/00423114.2026.2692718

To link to this article: https://doi.org/10.1080/00423114.2026.2692718

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 29 Jun 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=nvsd20

VEHICLE SYSTEM DYNAMICS
https://doi.org/10.1080/00423114.2026.2692718

OPEN ACCESS Check for updates

Kernel-based digital twins for evaluating the running safety of freight wagons during emergency braking

N. Bosso[a], L. Cantone[b], M. Magelli[a], R. Trinchero[c] and N. Zampieri[a]

[a]Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Torino, Italy; [b]Department of Enterprise Engineering, Tor Vergata University of Rome, Roma, Italy; [c]Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy

ABSTRACT

The paper defines a computational framework for the derivation of digital twins of multibody (MB) simulations. The framework integrates longitudinal train dynamics (LTD) and MB codes. The digital twins are based on kernel regressions and classifications, and they can predict the running safety of freight trains during emergency air brake operations. The digital twins include two types of models, namely scalar regressions for the evaluation of the running safety indexes and a binary classifier for the detection of unsafe running conditions. The inputs for the digital twins can be easily obtained from common LTD codes. With 400 training samples, the scalar regression models have an average R^2 score around 0.8, while a basic binary classifier can correctly label 90% of data. A weighted classifier can minimise misclassification of unsafe conditions, but this comes at the cost of worsened classification of safe behaviour. Overall, good accuracy can be achieved with the advantage of a small dataset. Once trained, the evaluation of the digital twins is 6 orders of magnitude faster than launching multibody simulations. The combination of good prediction capabilities and outstanding computational times can thrust the use of kernel regressions as surrogates of MB simulations.

ARTICLE HISTORY
Received 15 September 2025
Revised 29 March 2026
Accepted 16 June 2026

KEYWORDS
Multibody simulation; longitudinal train dynamics simulation; digital twins; machine learning; kernel regression

1. Introduction

The paradigm of digital twins (DTs) is currently receiving growing interest in the railway field, with possible application on several subsystems, including railway vehicles [1]. Among others, DTs can be a game changer in the fields of vehicle design [2] and vehicle dynamics [3]. With reference to the latter, nowadays, international rules worldwide like the European EN14363 standard [4] allow for virtual homologation of railway vehicles, thus assessing the vehicle dynamics behaviour with numerical simulations. The EN14363 defines different indexes with the corresponding thresholds to assess the running safety against derailment, including the derailment ratio Y/Q, the sum of lateral forces $\sum$Y and the unloading ratio DQ/Q. These parameters are derived from the calculation of the

CONTACT M. Magelli matteo.magelli@polito.it

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

wheel-rail contact forces, which require dedicated modules and routines available in most multibody (MB) simulation packages.

When assessing the vehicle dynamic behaviour according to the EN14363 standard, isolated vehicles or short fixed compositions of small groups of vehicles are considered. Whilst this approach can be suitable for most passenger vehicles, freight wagons can operate under different train compositions and a wide range of possible conditions. Furthermore, the forces acting on the coupling elements, especially longitudinal compressive forces (LCFs) arising during air brake operations on curved track sections, are known to increase the derailment risk.

Due to the big computational effort required by a detailed investigation of the full dynamics of all wagons inside the train consist, simplified models are developed to simulate the longitudinal train dynamics (LTD) and hence calculate the position and speed of each vehicle in the composition as well as the in-train forces [5,6]. As typical LTD codes, including those joining an international benchmarking activity launched in 2017 [7,8], only evaluate the longitudinal dynamics of the vehicles in the train consist, they do not output the wheel-rail contact forces, thus hindering the evaluation of the running safety indexes defined in international rules. To still evaluate the running safety of full train compositions, the IRS 40421 [9] describes a relative approach based on statistical methods that use the in-train forces obtained from LTD simulations. The LCFs values are compared with extrapolated permissible values obtained from experimental tests, as prescribed by the UIC 530 leaflet [10].

To further improve the running safety assessment of freight wagons considering the interactions inside the train composition and heterogeneity of operating conditions [11], it is necessary to combine LTD and MB simulations, with different strategies developed over the years to accomplish this goal [12,13]. However, while LTD simulations are usually faster than real-time, and can potentially be implemented in automatic train operation (ATO) and protection (ATP) systems, MB simulations usually demand high computational efforts, which make on-board implementation unfeasible.

A valid solution to tackle this issue is to develop DTs of MB simulations, with the aid of computationally efficient surrogate models developed with artificial intelligence (AI) and machine learning (ML) techniques, that can replace the computationally demanding MB simulations. In past works, the Authors built support vector machine (SVM) and least-square support vector machine (LS-SVM) kernel regressions, trained from MB simulations, to obtain surrogate models able to estimate the running safety indexes both when running at constant speed [14], solving the LTD equations with the LTDPoliTo code [15,16], and when performing emergency air brake operations [17], solving the pneumatic and dynamic problems with the *TrainDy* code [18]. Bernal et al. [19] conceived an augmented DT for railway system (ADTR) to be integrated within the framework of an ATO producing locomotive driving commands. Zhang et al. [20,21] trained and tested different neural networks (NNs) to predict in-train forces from ATO measurements that were emulated via LTD simulations run for a wide range of operating conditions.

The present paper relies on the computational framework built in previous works published by the Authors [14,17]. With respect to the original work [14], the simulations shown in this paper consider emergency air braking operations, simulated with the *TrainDy* code, which is the recommended code by UIC. The reference operation scenario, including emergency air braking after full traction, is the same considered in the conference paper

[17], which features higher concerns for the running behaviour. However, the present paper is an extended version of the conference paper, as it provides additional details on the modelling strategy and it considers a wider range of train combinations, as simulations are launched for different combinations of the main parameters of three wagons in the middle of the train (central trio). Moreover, as a major novelty, the computational framework is adapted to build binary classifiers for derailment detection, which can be combined with surrogate models for a more robust evaluation of the running conditions. Furthermore, with respect to the conference paper, this work more deeply investigates the effect of the number of training samples on the prediction capabilities of the derived surrogate models and classifiers.

The goal of the paper is to prove how DTs can provide excellent computational accuracy in the prediction of the running safety indexes and classification of derailment conditions at a minimal computational cost and with a small amount of training samples, even when considering a quite large number of input parameters. These are big advantages of DTs, which make them suitable for different possible applications both offline, to enhance the capabilities of common LTD codes, and on board the vehicles for different purposes. Nonetheless, the practical implementation of the trained DTs is outside the scope of the present paper, as each specific application is characterised by its own challenges, which require additional investigations for the identification of proper solutions.

The structure of the paper is as follows. Section 2 shows the computational framework developed to train and test the ML models, and provides the background of the adopted LTD code, of the MB model of freight wagon with Y25 bogie and of the kernel regression techniques used to build the surrogate models and classifier. Then, Section 3 presents the major results of the activity. Section 4 critically discusses the implications of the derived DTs, suggesting possible applications where DTs can be game changers. Finally, the Conclusions section summarises the outcomes of the work and suggests possible future developments.

2. Method

The ML models derived in this paper are trained using the results of dynamic simulations of emergency air brake operations of European freight wagons equipped with the Y25 bogie, run for different operating conditions. A reference train, with a head locomotive trailing 40 wagons, is considered, see Figure 1a. All wagons inside the train have a length of 18 m and axle-load of 22.5 tons, except for the wagon trio in the middle of the train consist. For these 3 wagons, the values of axle-load and length are model parameters that change to explore different possible wagon configurations. The train configuration is always built ensuring that the middle wagon, i.e. the 21st vehicle starting the count from the head of the train, is the lightest and hence the one where the highest derailment risk is expected. Therefore, the ML models are built to predict the running safety indexes on all wheels and wheelsets of this vehicle.

The reference freight train runs on a track with an initial 1000 m level straight section, followed by a clothoid transition and a full-curve section, see Figure 1b. Please note that Figure 1b only intends to qualitatively show the trend of rail vertical coordinate and curve radius, that are not constant in the simulations but depend on the LHS. Therefore, units and numerical values are not displayed in the plot for these quantities. The track superelevation

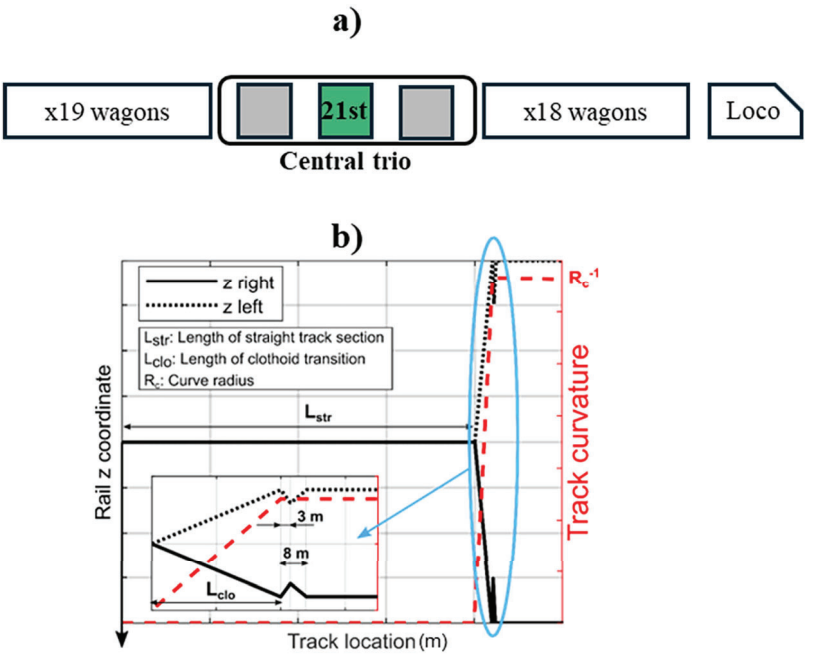


Figure 1. Reference train and track: (a) train composition and (b) track curvature and superelevation characteristics, with zoom on the twist defect at the end of the clothoid transition.

is calculated as a function of the curve radius only, assuming the unbalanced acceleration to be proportional to the curve radius. The length of the clothoid transition is calculated considering a nominal twist of 2‰. The single curve layout is selected over the S-shaped curved section since it is prevalent on common railway tracks. Therefore, this allows us to apply the trained surrogate models over typical operating conditions and track layouts. The reference train operation starts with an initial acceleration from a speed of 5 km/h until reaching the target initial speed at braking, and the braking command is activated.

The track is modelled without irregularities, to minimise the effect of stochastic parameters in the derivation of the surrogate models. Nonetheless, a twist defect is applied at the end of the transition zone. The applied twist error is equal to 6.5‰ over a base length of 3 m, which is the limit value of twist defect prescribed by the Italian railroad administration ('Rete Ferroviaria Italiana' – RFI), see the zoom in Figure 1b. The twist defect is specifically applied to increase the derailment risk and build conservative surrogate models. The track is modelled as a moving elastic foundation (see §2.3).

The computational framework built to train and test the ML models is sketched in Figure 2. First, a number of configurations of the input parameters is selected via Latin hypercube sampling (LHS) by means of a MATLAB routine. The simulation input parameters, described in detail in the following subsection, are related to the characteristics of each wagon inside the central trio as well as to the specific braking operation, in terms of target speed at braking start and curve radius. For each combination of the simulation input parameters, an LTD simulation is run with an enhanced version of the UIC approved *TrainDy* code [18] to calculate the in-train forces during the emergency braking operation. The in-train forces acting at the front and back of the central wagon trio are applied

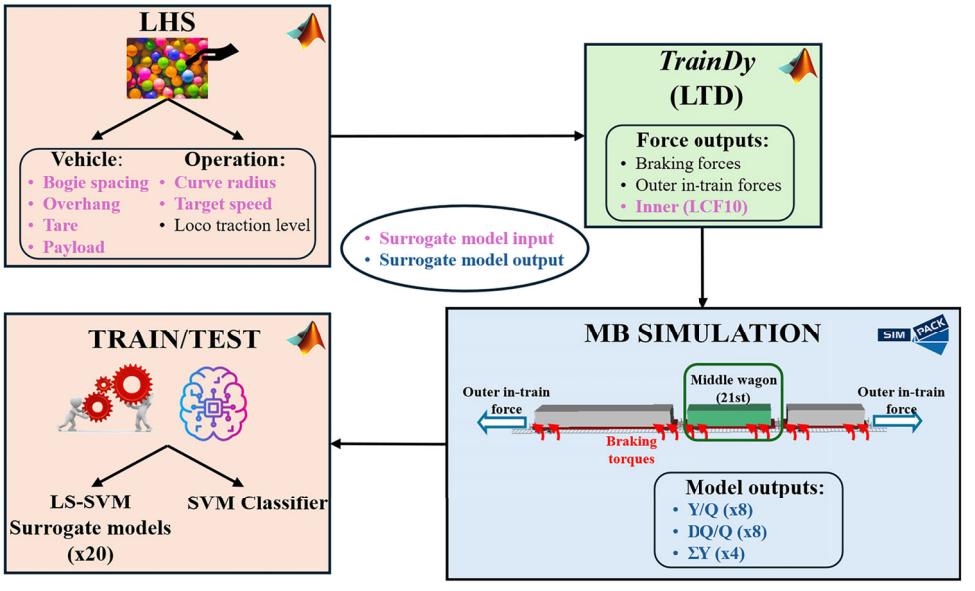


Figure 2. Framework for the derivation of DTs of MB simulations.

to a detailed MB model of the wagon trio implemented in SIMPACK. At the end of each MB simulation, 20 scalar parameters are recorded on the critical vehicle (middle wagon), namely the derailment ratio Y/Q (x8) and unloading ratio DQ/Q (x8) on each wheel, as well as the sum of guiding forces $\sum$Y on each wheelset (x4). The safety indexes represent the outputs of the dynamic simulations, and they are extracted using the filtering strategies described in the EN14363 standard [4]. These include the application of (i) low-pass filter with 20 Hz cut-off frequency, (ii) moving average with 2 m sliding window and step below 0.5 m and finally (iii) the selection of the maximum absolute value between the 0.15% and 99.85% percentile. Although the simulations are not run with the purpose of virtual homologation, the EN14363 processing strategy is applied to smooth over instantaneous peaks of the running safety indexes which are unlikely to bring to derailment. Furthermore, a Boolean parameter for derailment detection is calculated, which is set to 1 when any value of Y/Q or $\sum$Y, post-processed according to the EN14363 prescriptions, is above the corresponding threshold defined in international rules, and 0 otherwise. In the training stage, the surrogate models that relate the major inputs to the simulation outputs and the derailment classifier are obtained via ML kernel regressions using MATLAB functions. Conversely, in the testing stage, the accuracy of the previously derived ML regressions and classifier is assessed via comparison of the predicted outputs with the actual values calculated from the MB simulations.

Although this work relies on the *TrainDy* code for LTD simulations during emergency braking operations, it is important to stress that the proposed approach can be extended and adapted to any LTD code, as long as the code is capable of evaluating the air brake forces [22,23]. Similarly, the proposed framework can be adapted to any desired commercial or in-house MB software package.

2.1. Data generation

The simulation input parameters are drawn via LHS to generate train configurations with different wagons in the middle trio and to consider different operations. Precisely, 4 parameters are drawn for each wagon in the central trio to define the main vehicle data (tare, payload factor, bogie-spacing, overhang factor), and 3 parameters are used to define the operating conditions (locomotive traction level, curve radius and initial speed at braking start). The latter operation parameters are drawn for operating conditions that are known to increase the derailment risk, i.e. curve radius between 150 and 300 m and initial speed in the range 20-40 km/h, whereas the traction level is used to change the amount of LCF.

Concerning the vehicle data, for each wagon in the middle trio, a bogie spacing value is drawn from the typical range of two-bogie European freight wagons, that is assumed to be 5-20 m, based on the overhang diagrams provided by the UIC 530-2 leaflet [10]. The wagon length is determined based on the bogie spacing to comply with the overhang limits as prescribed by the same UIC 530-2 leaflet. This is achieved via the sampling of an additional overhang parameter per each of the wagons in the trio. The overhang parameter varies in the range 0-1, such that 0 corresponds to the minimum allowed overhang and 1 corresponds to the maximum possible overhang for the considered bogie spacing value. The weight-on-rails of the wagons in the central trio is determined by sampling two parameters for each vehicle. The first parameter is drawn to determine the vehicle tare from the corresponding variation range, which is defined based on typical values for the considered types of wagons (20–36 ton). The second parameter is a loading parameter that is equal to 0 when the wagon is unloaded while it is 1 when the maximum payload is considered. For the external vehicles in the trio, the latter is the payload that leads to a total weight-on-rails of 90 tons, while for the middle vehicle, the maximum payload is the payload that leads to a weight-on-rails equal to the lighter vehicle between the outer wagons. Anyways, the maximum allowed weight-on-rails for the middle wagon is limited to 60 tons, to increase the load unbalancing between adjacent vehicles and hence increase the derailment risk. It is worth noting that in this way, it is possible to reproduce virtually any operating conditions in terms of wagon length, tare and payload arrangement within the trio, assuming that the mass of the light wagon is limited to 60 tons.

It is important to stress that the inputs for the ML models slightly differ from the simulation inputs drawn via LHS. In fact, the ML kernel regression inputs explicitly include the values of 10 m LCF (LCF10) acting at the front and back of the central wagon and calculated from the *TrainDy* LTD simulations. This is done to avoid that the derived ML models show a direct dependency on the locomotive traction level and specific train composition, as well as to allow them to be used with the outputs available from typical LTD simulations, even when the locomotive is coasting, and a braking operation is commanded. Furthermore, the overhang ratio and bogie-spacing are not used directly as inputs in the model, and they are replaced by parameter combinations that are commonly easier to derive from real wagon data, namely the length over buffers and ratio between bogie-spacing and length over buffers. The main inputs for the derived trained regressions are summarised in Table 1. In total, 16 input parameters are required to evaluate each of the 20 surrogate models for the calculation of the running safety indexes, as well as for the binary classifier.

Table 1. Main input parameters for the kernel regressions.

Parameter	Range	Quantity	Notes
Tare	20–36 ton	x3	Sampled with LHS (one per vehicle in the trio)
Payload factor	0–1	x3	Sampled with LHS (one per vehicle in the trio): maximum weight on rails is 90 ton for outer vehicles and 60 ton for middle one
Length over buffers	–	x3	Calculated for each vehicle in the trio from parameters drawn with LHS (bogie spacing and overhang factor).
Bogie-spacing/length over buffers	-	x3	Calculated for each vehicle in the trio from parameters drawn with LHS (bogie spacing and overhang factor)
Target speed at braking start	20-40 km/h	x1	Sampled with LHS.
Curve radius	150-300 m	x1	Sampled with LHS.
LCF10 on middle wagon	-	x2	Calculated with LTD simulation (the values for the coupling systems in front and at the back of the middle wagon are extracted).

2.2. LTD simulations

The *TrainDy* software [18] was developed by the University of Rome Tor Vergata with the financial and technical support of Faiveley Transport (a Wabtec Group company). It was acquired by the UIC in 2007 and an international validation of the software was launched to check its conformity with more than 30 experimental tests provided by DB AG, SNCF and TRENITALIA. In 2009, UIC set up a 'special group' within UIC to further improve the tool. In 2021, the UIC published the International Railway Solution (IRS) 40421, where the *TrainDy* software, ratified by the Special Group, is recommended for LTD simulations. The software is written in MATLAB and in its basic version it is an executable program called by a GUI, therefore, it is not suitable for statistical analysis. The software used for the LTD simulations of this paper is an extended version of the UIC approved *TrainDy* code, capable of analysing train configurations with user modified inputs, such as those provided by the LHS sampling.

TrainDy solves the pneumatic and dynamic problems together, in various operational scenarios [18]. For this paper, the initial position is set as a possible input to *TrainDy*: in this way, for each configuration of input parameters generated by the LHS, it is possible to impose that the peak of the LCFs acting on the 21st vehicle occurs when the wagon is located between the clothoid and the full curve. By this way, despite the specific input parameters, all simulations have a common characteristic, which simplifies model training. To achieve this, for each set of input parameters, several preliminary LTD simulations are run before the final simulation is started, in order to adjust the initial position of the train so that the peak of the LCFs on the 21st vehicle occurs at the beginning of the full curve. For this purpose, a dedicated algorithm, similar to the MATLAB *fzero* function, was developed to find the zeros of a non-linear numerical equation. Although it is not guaranteed that derailment does occur at the exact location where the LCF peak is recorded, since derailment is triggered by a combination of longitudinal, lateral and yaw dynamics, this choice allows to set the initial position of the train with an optimisation of LTD simulations. Other choices, such as determining the initial position in order to have the maximum yaw moment at the twist defect location, would require an optimisation performed with a combination of LTD and MB simulations, which would make the computational times unfeasible.

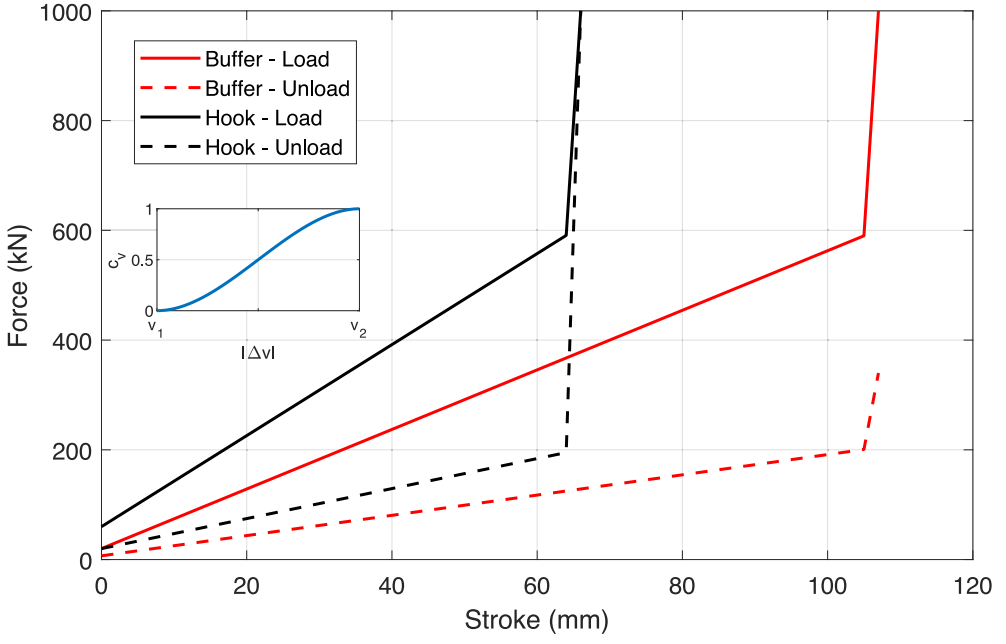


Figure 3. Force-stroke characteristics of buffer and hook modelled in *TrainDy* (the inner plot shows the trend of the hysteresis damping third-order polynomial function).

The in-train forces are calculated in *TrainDy* starting from the force-stroke characteristics of buffers and hook, see Figure 3. The hysteresis is modelled with a damping transition between loading and unloading curves depending on the value of deflection speed [18], as stated by:

$$F_{long}(\Delta x, \Delta v) = [1 - c_v(|\Delta v|)] \cdot F_{unload}(\Delta x) + c_v(|\Delta v|) \cdot F_{load}(\Delta x) \tag{1}$$

where F_{long} is the longitudinal force provided by the hook/buffers, Δx is the deflection, Δv is the deflection speed, F_{unload} and F_{load} are the unloading and loading forces, respectively, both calculated as a function of deflection, and finally c_v is the hysteresis damping function. The latter is defined as a third-order polynomial function that saturates to 0 when the absolute value of the deflection speed is below a lower threshold v_1 and equals 1 when the deflection speed in absolute value approaches an upper threshold v_2, see the inner plot in Figure 3. Of course, the characteristics implemented in the *TrainDy* code consider the series of two hooks as well as the parallel of the series of adjacent buffers. The buffers only act in compression, while the hook only provides tensile forces.

In addition, to streamline the simulations, *TrainDy* was transformed into a MATLAB function capable of being run in parallel and modifying inputs written to disk, such as payload distribution and wagon tare/length, without the need to have a different input on disk for each case. In this way, it has not been necessary to create a different train configuration/test/train operation for each set of input parameters.

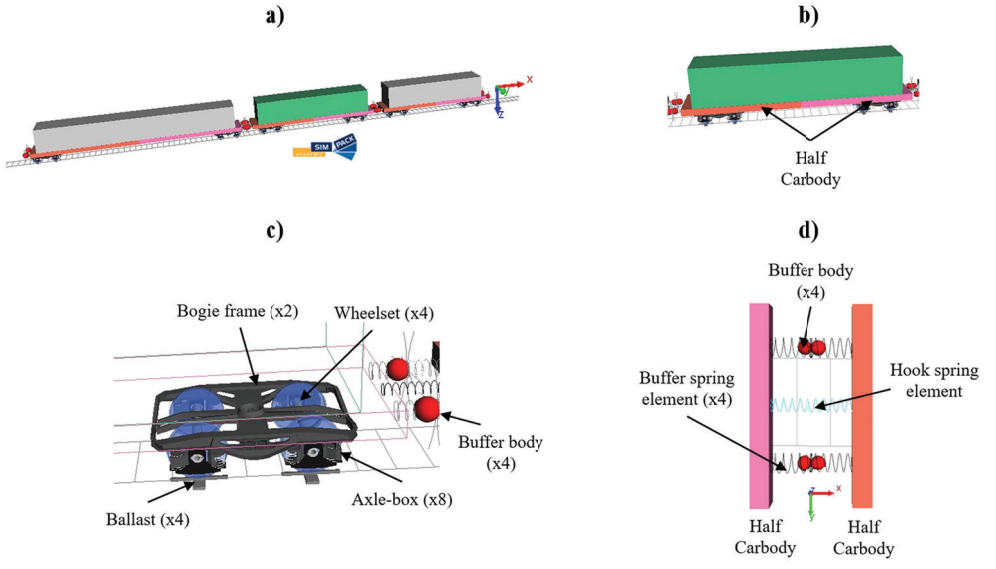


Figure 4. SIMPACK MB model used in this work: (a) wagon trio; (b) detail of single vehicle with two half carbodies; (c) bogie model; (d) buffer-hook connection system.

2.3. MB model

For each configuration of the simulation input parameters drawn via LHS, a MB simulation is run to calculate the full dynamics of the wagons in the central trio, with the 21st wagon in the middle. The MB model, built in SIMPACK, includes 3 railway vehicles, see Figure 4a, and it is derived from a well-established MB model of the Y25 bogie, developed by the Politecnico di Torino (PoliTo) railway research group in past works [24]. Each wagon in the model is made up of 20 bodies, namely 2 half-carbodies, 2 bogie frames, 4 wheelsets, 8 axle-boxes and 4 buffers, see Figures 4b and c. All bodies, except for axle-boxes and buffers, feature 6 degrees of freedom (DOFs) along the track, defined through the SIMPACK general rail track joint. The axle-boxes are only allowed to rotate with respect to the wheelset they belong to, while for buffers, a different joint type is applied on the inner bodies (buffers between the middle wagon and the outer ones) and outer bodies (buffers that connect the outer wagons of the trio to the rest of the composition, which is not modelled in SIMPACK). The inner buffers have a translational DOF in longitudinal direction with respect to the half-carbody, while the outer buffers are rigidly fixed to the half carbodies, thus featuring zero DOFs. The main inertial properties of the bodies in the MB model and the specified joints are given in Appendix 1, see Table A1.

The track is modelled as a moving elastic track foundation with the definition of one ballast body per wheelset, see again Figure 4c. The ballast body features 3 DOFs defined through the Wheel/Rail Track Sleeper joint available in SIMPACK, namely lateral and vertical translation as well as a roll rotation. Each ballast body is connected to the reference frame with a dedicated force element with lumped stiffnesses and damping along the lateral, vertical and roll directions. The values of stiffness and damping along each direction

are set to the default values available in SIMPACK, that are derived from a ERRI (ORE) benchmarking activity [25]. In total, the MB model features 72 bodies and 212 DOFs.

The Y25 MB model considers the nonlinearities of both primary (Lenoir link and bilinear vertical stiffness characteristic) and secondary suspension (centreplate and side bearers) stages, and the interested reader can refer to [24]. The two half-carbodies are connected to each other with a force element applying high stiffness values in all directions except for roll, for which the torsional stiffness of the carbody is prescribed. The wheel-rail contact forces are calculated using the SIMPACK equivalent elastic method for the solution of the normal problem (calculation of the contact patch area and normal force) and FASTIM [26] for the tangential problem (calculation of the tangential forces).

The SIMPACK model of the wagon trio includes a detailed model of the buffer-hook systems connecting the middle wagon to the outer vehicles, see Figure 4d. Each inner buffer is connected to the corresponding half-carbody with a point-to-point (Ptp) nonlinear element with hysteresis, for which the mechanical impedance characteristic of a single buffer is specified by means of input functions for the loading and unloading curves, see Figure 3. Additionally, the contact between adjacent buffers is modelled with a Hertzian contact force element that solves the contact between spherical bodies with a radius of 2.75 m. Finally, the hook connection between the central wagon and the outer wagons of the trio is modelled through the definition of two dedicated Ptp nonlinear elements with hysteresis, which implement the hook mechanical impedance characteristic, specified through input functions. The loading and unloading curves specified for the hook force element correspond to the stiffness series of the mechanical impedance characteristics given in Figure 3.

The main force elements defined in the MB model for the connection of the main bodies (primary suspension, secondary suspension, connection between the half-carbodies, buffer-hook system) are summarised in Appendix 1, see Table A2.

Extra force elements are defined on the MB model to consider the effects of the train composition. The forces on the coupling systems connecting the outer wagons of the central trio to the rest of the train are applied as equivalent tractive forces acting along the track longitudinal direction, with values coming from the LTD simulation run with the *TrainDy* code. Similarly, the air brake forces on each vehicle, obtained from the pneumatic module of the *TrainDy* code, are applied in the model as equivalent braking torques. Finally, additional resistant forces are applied on each vehicle to account for the ordinary and curving resistances. These forces are provided as pure longitudinal forces and calculated from the same expression used in the LTD simulations run with *TrainDy*. Precisely, the ordinary and curving resistances are calculated with the following equations suggested in the international benchmark of LTD simulators [7,8]:

$$F_{ord} = M_{tot} \cdot \left[2.943 + \frac{89.2}{Q_{ax}} + 0.0306V + \frac{0.122V^2}{M_{tot}}\right] \tag{2}$$

$$F_{crv} = M_{tot} \cdot \frac{6116}{|R_c|} \tag{3}$$

where F_{ord} is the contribution due to ordinary resistances, F_{crv} is the contribution due to curving, V is speed (to be expressed in km/h), R_c is the curve radius, Q_{ax} is the vehicle axle-load (to be expressed in ton) and M_{tot} is the total mass of the vehicle, which includes the rolling inertias and is expressed in ton.

As several simulations are run with the MB model of the central trio, for different configurations of the input quantities drawn via LHS, different parameters in the MB model must be changed from one simulation to another. The mass of the half-carbodies of all three wagons is properly adjusted for each simulation to obtain the target value of axle-loads, corresponding to the tare and payload factors drawn via LHS. Similarly, the values of bogie-spacing and length over buffers are modified based on the parameters (bogie-spacing and overhang factor) drawn with LHS. Coherently, the wheelset pitch inertia is adjusted in each simulation to obtain a constant ratio between the rolling inertias and the wagon tare mass equal to 7.4%. This is the value that corresponds to an inertia of the wheelset $I_{yy,w} = 110\,\text{kg}\cdot\text{m}^2$ for a tare load of 28 ton and nominal wheel radius of 0.46 m. With a similar approach, the torsional stiffness of the carbody, defined as a lumped stiffness applied between the two half-carbodies, is assumed to be inversely proportional to the bogie-spacing, such that a torsional stiffness of 281,000 Nm/rad is applied for a reference bogie-spacing value equal to 12 m.

2.4. Introduction to LS-SVM regression and SVM classifier for metamodeling

Metamodeling, also known as surrogate modelling, aims to build efficient and closed-form approximations of complex and computationally expensive models. Among various machine learning approaches, this work investigates the effectiveness of Least Squares Support Vector Machine (LS-SVM) regression [27] and Support Vector Machine (SVM) classification [28], as they provide efficient solutions for regression and classification tasks, respectively, even with small datasets. Precisely, an LS-SVM is adopted to build surrogate models for the prediction of the relevant safety parameters, while a SVM classifier is used to obtain a binary classifier for derailment.

2.4.1. Least squares SVM (LS-SVM) regression

LS-SVM regression is a supervised learning algorithm that efficiently trains nonlinear regression models by solving a least-squares optimisation problem [27,29]. Given a set of training pairs $\{(\boldsymbol{x}_l, y_l)\}_{l=1}^{L}$, where $\boldsymbol{x}_l \in \mathbb{R}^d$ are the input features and $y_i \in \mathbb{R}$ are the corresponding output values, the LS-SVM regression in the primal space is expressed as:

$$\mathcal{M}(\boldsymbol{x}; \boldsymbol{w}, b) = \boldsymbol{w}^T \boldsymbol{\Phi}(\boldsymbol{x}) + b \tag{4}$$

where $\boldsymbol{\Phi} : \mathbb{R}^d \rightarrow \mathbb{R}^D$ is a nonlinear mapping that transforms the d-dimensional input space to a higher D-dimensional feature space, and $\boldsymbol{w}$ and b are the regression coefficients and the bias term, respectively.

These parameters are computed during the training phase as the solution of the following optimisation problem [27,29]:

$$\min_{\boldsymbol{w},b} \frac{1}{2} ||\boldsymbol{w}^2|| + \frac{\gamma}{2} \sum_{l=1}^{L} (y_l - \mathcal{M}(\boldsymbol{x}_l;\ \boldsymbol{w}, b))^2 \tag{5}$$

where the first term regularises the model by controlling its flatness to prevent overfitting, and γ is a model hyperparameter tuned via cross-validation [30].

The primal space formulation of the LS-SVM regression admits the following dual formulation [27,29,31]:

$$\mathcal{M}(\boldsymbol{x}) = \sum_{l=1}^{L} \alpha_l k(\boldsymbol{x}_l, \boldsymbol{x}) + b \tag{6}$$

where $k(\cdot,\cdot)$ is the kernel function, α_l are the coefficients of the dual formulation of the model (also known as the dual coefficients or Lagrange multipliers), and b is the bias term. Thanks to the kernel trick, the above dual formulation makes the complexity of the model independent of input dimensionality. Indeed, the kernel function implicitly maps data into a higher-dimensional space, enabling it to handle nonlinear regressions without explicitly computing the feature transformations $\boldsymbol{\Phi}$ [27]. This allows the LS-SVM to model complex and nonlinear input-output relationships while maintaining computational efficiency [27,29].

Several kernel functions are available. In this work, a Gaussian Radial Basis Function (RBF) kernel is used:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\frac{1}{2\sigma^2}||\boldsymbol{x} - \boldsymbol{x}'||^2\right) \tag{7}$$

where σ denotes the kernel width and represents a kernel hyperparameter. The regression model involves two hyperparameters, namely the kernel width σ and the regularisation parameter γ . These hyperparameters are optimised during training via cross-validation or a validation set [30].

LS-SVM regression is particularly useful for nonlinear function approximation and offers computational efficiency compared to standard SVM regression, since its training requires solving a linear system [14,27,29].

2.4.2. Support vector machine (SVM) for classification

The SVM classifier is a supervised learning algorithm that finds the optimal hyperplane to separate data while maximising the margin between different classes. Given a set of training pairs $\{(\boldsymbol{x}_l, y_l)\}_{l=1}^{L}$, where again $\boldsymbol{x}_l \in \mathbb{R}^d$ are the input features and $y_l \in [-1, 1]$ are the corresponding class labels (i.e. discrete outputs), the goal of the SVM classifier is to find the following decision function [28]:

$$\mathcal{F}(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{\Phi}(\boldsymbol{x}) + b \tag{8}$$

where again $\boldsymbol{\Phi}(\boldsymbol{x})$ maps the input space to a higher-dimensional feature space, allowing to handle complex and nonlinear decision boundaries, and $\boldsymbol{w}$ and b are the model coefficients and bias term, respectively.

The decision function $\mathcal{F}$ divides the space into two parts, predicting class $+1$ when $\mathcal{F}(\boldsymbol{x}) > 0$ and class -1 when $\mathcal{F}(\boldsymbol{x}) < 0$, such that:

$$\hat{y} = \text{sign}(\mathcal{F}(\boldsymbol{x})) \tag{9}$$

The parameters $\boldsymbol{w}$ and b of the SVM classifier are estimated during model training by solving the following optimisation problem [28]:

$$\min_{\boldsymbol{w},b} \frac{1}{2}||\boldsymbol{w}^2|| + C\sum_{l=1}^{L} w_{y_l}\xi_l \tag{10}$$

subject to

$$\begin{gathered} y_l(\boldsymbol{w}^T \boldsymbol{\Phi}(\boldsymbol{x}_l) + b) \geq 1 - \xi_l \\ \xi_l \geq 0 \end{gathered} \tag{11}$$

for $l = 1, \ldots, L$.

In the above equations, C is the regularisation term, ξ_l are the slack variables that allow margin violations, and w_{y_l} are relative weights associated with each training sample to favour one class over the other. The objective is to minimise the norm of the weight vector $\boldsymbol{w}$, while controlling the slack variables ξ_l to allow limited misclassifications, to achieve better class separation and improved model generalisation.

The function $\boldsymbol{\Phi}$ maps the input data into a higher-dimensional feature space, allowing SVM to learn complex and nonlinear decision boundaries. In the dual form, the decision function is rewritten in terms of the Lagrange multipliers α_l acting as model parameters:

$$\mathcal{F}(\boldsymbol{x}) = \sum_{l=1}^{L} \alpha_l y_l k(\boldsymbol{x}_l, \boldsymbol{x}) + b \tag{12}$$

where like LS-SVM regression, $k(\cdot, \cdot)$ is a kernel function that enables nonlinear classification without explicitly computing the feature transformations and b denotes again the bias term. Also in this case, the two model hyperparameters (i.e. the kernel width σ and the regularisation parameter C) are optimised during training via cross-validation or a validation set [30].

This kernel trick allows SVM to model complex relationships while maintaining computational efficiency, benefiting from the sparsity of the resulting model [27,28]. SVM is widely used due to its ability to handle both linearly and nonlinearly separable data, making it robust for high-dimensional problems and small datasets [27,28].

2.4.3. Application of LS-SVM and SVM kernel regressions to railway vehicle dynamics

The results of the LTD and MB simulations were used to train 20 independent scalar LS-SVM regression surrogates (one for each simulation output) to predict the running safety indices of each wheel and wheelset of the middle wagon. Input features were min–max normalised to the interval [−1,1], while output variables were standardised (z-score normalisation). An RBF kernel was adopted. The models were trained using the MATLAB LS-SVMlab Toolbox [32], and the hyperparameters were tuned via leave-one-out cross-validation combined with simplex optimisation. It is important to remark that each surrogate model was trained independently and requires the tuning of only two hyperparameters (the model parameter γ and the RBF kernel width σ).

Moreover, a binary SVM classifier, as described in Section 2.4.2, was developed to predict train derailment conditions. Input features were min–max normalised to the interval [−1,1]. The dataset was divided into two classes: 'derailed' and 'not derailed'. In the simulations considered in this study, no actual derailment occurred. Therefore, the label 'derailed' is assigned to operating conditions for which at least one safety index obtained from the MB simulation exceeds the limits prescribed by the EN14363 standard.

The classifiers were implemented using the SVM function available in the MATLAB Statistics and Machine Learning Toolbox, adopting an RBF kernel. Two hyperparameters, namely the regularisation parameter C and the RBF kernel width σ, were optimised via

Table 2. Summary of the main characteristics of the LS-SVM regression and SVM classification models, including kernel choice, hyperparameters, tuning strategy, cross-validation scheme, and data scaling procedures.

Model	Kernel	Hyperparameters	Tuning Method	CV Strategy	Data Scaling
LS-SVM Regression	RBF	σ, γ	Simplex optimisation	Leave-one-out	Inputs: min–max [−1,1] Outputs: z-score
SVM Classification	RBF	σ, C	Bayesian optimisation	5-fold CV	min–max [−1,1]

Bayesian optimisation with 5-fold cross-validation. Two different classifiers were trained, corresponding to two weighting strategies applied to the training samples, see Equation (10). The first classifier is unweighted, with i.e. $w_{yl} = 1 \; \forall l$. The second classifier is weighted, where the samples labelled as 'not derailed' are assigned half the weight of those labelled as 'derailed', i.e. $w_{yl} = 1$ if $y_l = 1$; $w_{yl} = 0.5$ if $y_l = 0$. This is intentionally done to favour the capability of predicting unsafe conditions and hence limit the number of false negatives predicted by the classifier, i.e. limit the number of predictions of safe behaviour when instead the combination of LTD and MB identify that the limits set by the EN14363 standard are overcome. Please note that the selected weighting values represent a preliminary attempt to limit the number of false negatives. The investigation of alternative weighting configurations to assess their impact on classifier sensitivity and overall performance is going to be faced in future upgrades of the activity.

For clarity, Table 2 summarises the main characteristics of the regression and classification models, including kernel choice, hyperparameters, tuning strategy, cross-validation scheme, and data scaling.

3. Results

3.1. Dynamic simulations

To ensure the validity of the approach shown in the previous section, it is essential that the speed profile of the wagon trio in the MB simulation matches the one calculated from the LTD simulation. Figure 5 shows the speed profile of the central wagon of the middle trio for one configuration of the input parameters drawn with LHS. An excellent agreement exists between the *TrainDy* and SIMPACK simulations, thus confirming the validity of the implemented strategy to combine LTD and MB simulations, and the correct calculation of the ordinary and curving resistances in the two models. From Figure 5, it is possible to observe that at the beginning of the simulation, speed increases due to the traction effort provided by the head locomotive, until reaching a target speed. Then, the emergency air brake operation is commanded from the locomotive, and speed decreases until standstill. This kind of operation is repeated for all configurations of the input parameters, with proper adjustment of target speed at braking start, track layout and data of the wagons in the middle trio, see Table 1.

As already mentioned, for each simulation, corresponding to a specific configuration of the input parameters, the main running safety indexes are extracted for each wheel (derailment ratio and unloading ratio) and wheelset (sum of the guiding forces). For illustration purposes, Figure 6 provides the time histories of the main outputs calculated through numerical simulations on a single simulation run. Precisely, Figure 6a displays the values of

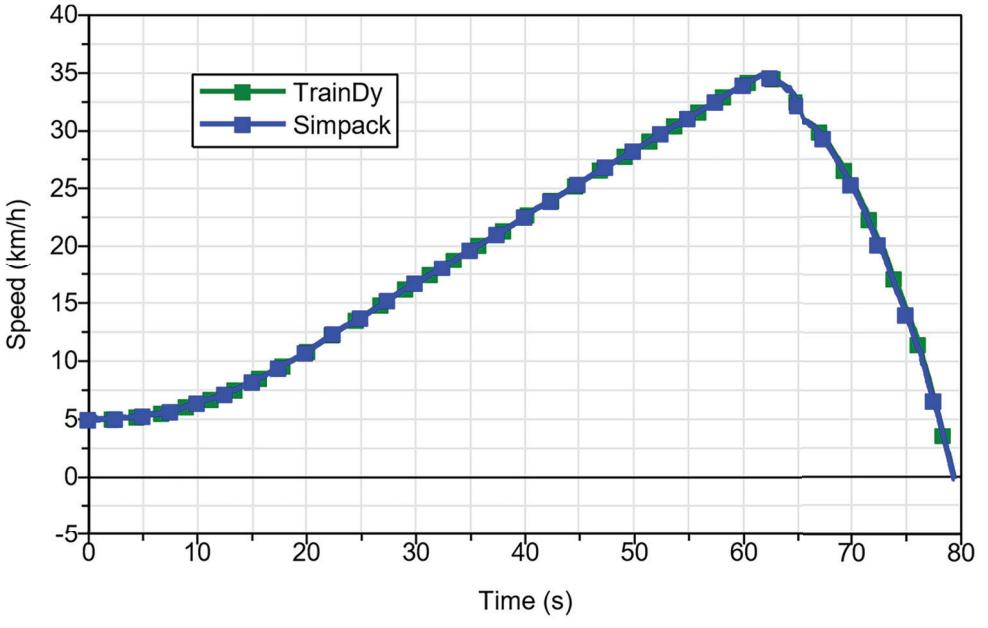


Figure 5. Comparison of speed profile obtained with *TrainDy* and Simpack simulations for one of the simulations run with different configurations of the input space parameters (curve radius = 237.75 m; target speed = 34.85 km/h).

in-train forces calculated with *TrainDy* at the front and rear of the middle wagon, with horizontal lines corresponding to the 10 m LCFs extracted as summarising parameters to build the surrogate models. On the other hand, Figure 6b shows the values of derailment ratio Y/Q on all eight wheels of the middle wagon. It can be observed from Figure 6a that positive in-train forces are obtained for the first part of the simulation, where the locomotive generates the tractive effort to accelerate the train from 5 km/h. Then, due to the activation of the air brakes, negative values, corresponding to compressive forces, arise. Concerning instead the derailment ratio, obtained from the MB model, low values are observed until the vehicle enters the clothoid transition, and peaks arise on the outer wheels of the leading wheelsets of the front and rear bogies close to the twist defect location. From the time history, for each wheel, a single scalar value is extracted, which is calculated using the filtering and percentile selection strategy defined in the EN14363 standard [4]. The same approach is adopted for the DQ/Q and ΣY parameters, which are not displayed here for the sake of brevity. The EN14363 filtering strategy is applied to build static surrogate models, that can calculate a single scalar value from the major simulation inputs. The 20 Hz low-pass and moving average filtering strategy with subsequent percentile selection as defined by the EN 14363 standard is implemented to smooth instantaneous peaks that would be unlikely to cause derailment, as some time is required for the applied wheel-rail contact forces to actually bring to derailment.

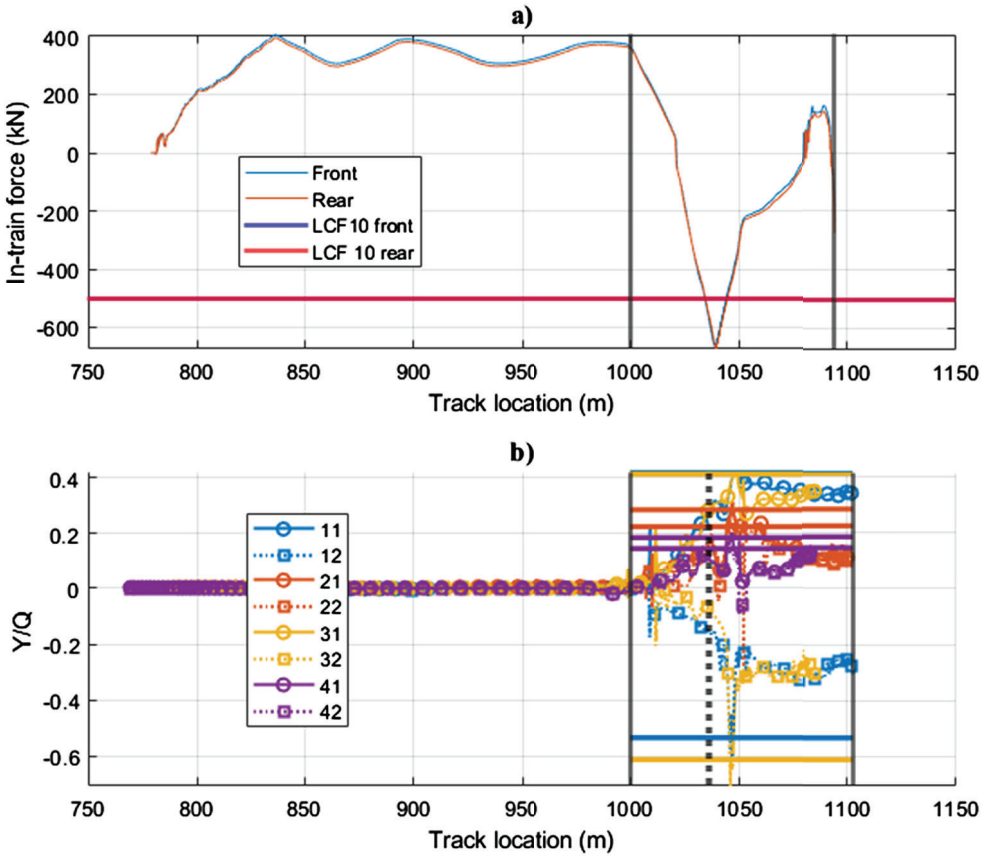


Figure 6. Results of numerical simulations for a reference configuration of the input parameters (curve radius = 237.75 m; target speed = 34.85 km/h): (a) in-train forces calculated with *TrainDy* and (b) derailment ratio (the dashed vertical line corresponds to the twist defect location, while horizontal solid lines correspond to the extracted values based on the application of the filtering strategy of the EN 14363).

To classify the running conditions for each simulation run, risk ratios are calculated for both flange climbing and track lateral shift, as stated by:

$$DF_N = \max \frac{\left|\frac{Y}{Q_{EN}}\right|_{ij}}{0.8} \quad i = 1-4; j = 1-2 \tag{13}$$

$$DF_{Ph} = \max \frac{|\Sigma Y_{EN}|_i}{0.85 \cdot (10 + \frac{P_{F0}}{3})} \quad i = 1-4 \tag{14}$$

where DF_N and DF_{Ph} refer to the risk ratio in terms of flange climbing and track shift, respectively, Y/Q_{EN} is the derailment ratio extracted with the EN filtering strategy, $\sum Y_{EN}$ is the sum of guiding forces extracted with the EN post-processing strategy, P_{F0} is the axle-load (in kN), and finally i is the wheelset counter, while j is the wheel side counter, which is 1 for right-hand side wheels and 2 for left-hand side wheels. Based on the definitions in Equations (13) and (14), when the risk ratio is above 1, the limits set by the EN14363 are

exceeded. Please note that in Equation (13), the limit value of the derailment ratio is taken as 0.8 with the explicit purpose of building conservative classifiers. Furthermore, the value of 0.8 is chosen over the Nadal's limit since the paper deals with the simulation of dynamic conditions, due to the emergency braking, while the Nadal's limit is believed to be more appropriate for quasi-static conditions. Additionally, a global risk ratio DF_{EN} is calculated as the maximum between the flange climbing and track shift risk ratios:

$$DF_{EN} = \max[DF_N; DF_{Ph}] \tag{15}$$

Therefore, labelling of running conditions for training and testing of the SVM binary classifier is performed according to the global risk ratio, so that simulations with global risk ratio above 1 are labelled as 'derailed', otherwise they are classified as 'not derailed'.

In total, 695 simulations were run with the approach detailed in Section 2. Out of the 695 samples, 295 random samples were selected as the testing dataset. From the remaining dataset, increasing random numbers of training samples (100, 200, 300 and 400) were selected to build the kernel regressions and classifiers, assessing the improvement of their prediction capabilities on the testing dataset. The total number of simulations, as well as the numbers of training samples and test samples, were preliminarily selected to ensure a trade-off between computational effort and accuracy of the derived DTs. Future work is demanded to explore different numbers of samples and the partitioning between training and testing dataset. Figure 7a shows the probability distribution of the individual risk ratios for flange climbing and track lateral shift, while Figure 7b displays the cumulative probability distribution of the overall risk ratio. From Figure 7b, it can be observed that around 37% of the simulations is classified as derailed and that the overall risk ratio is always above 0.5. This confirms that the reference simulation scenario is indeed characterised by a high derailment risk. On the other hand, Figure 7a highlights that flange climbing is more dangerous with respect to track lateral shift, as expected.

More in detail, for all 695 simulations, Figure 8 shows the probability distribution of derailment ratio and unloading index recorded on all wheels of the central wagon. On the other hand, Figure 9 provides the probability distribution for the track lateral shift risk ratio. For a proper understanding of Figures 8 and 9, please note that wheelsets are identified with one number, while two digits are used to identify wheels. The first number, in the range 1-4, identifies the wheelset, with 1 corresponding to the head wheelset. For wheels, the second number, instead, can be either 1 or 2, with 1 corresponding to right-hand side (inner) wheels and 2 corresponding to left-hand side (outer) wheels. From Figure 8, it can be inferred that, as expected, the left-hand side wheels (outer wheels) of the leading wheelsets belonging to the front and rear bogies feature the highest values of derailment index, and the critical wheel tends to be wheel 32, namely the outer wheel of the leading wheelset of the rear bogie. This means that the middle wagon tends to suffer more from long twist.

3.2. Prediction capabilities of the surrogate models

The predictive performance of the LS-SVM regressions is assessed in terms of the coefficient of determination (R^2 score) as well as with error-based metrics, namely RMSE and MAE. Additionally, since the Y/Q and DQ/Q indexes have values falling in a very

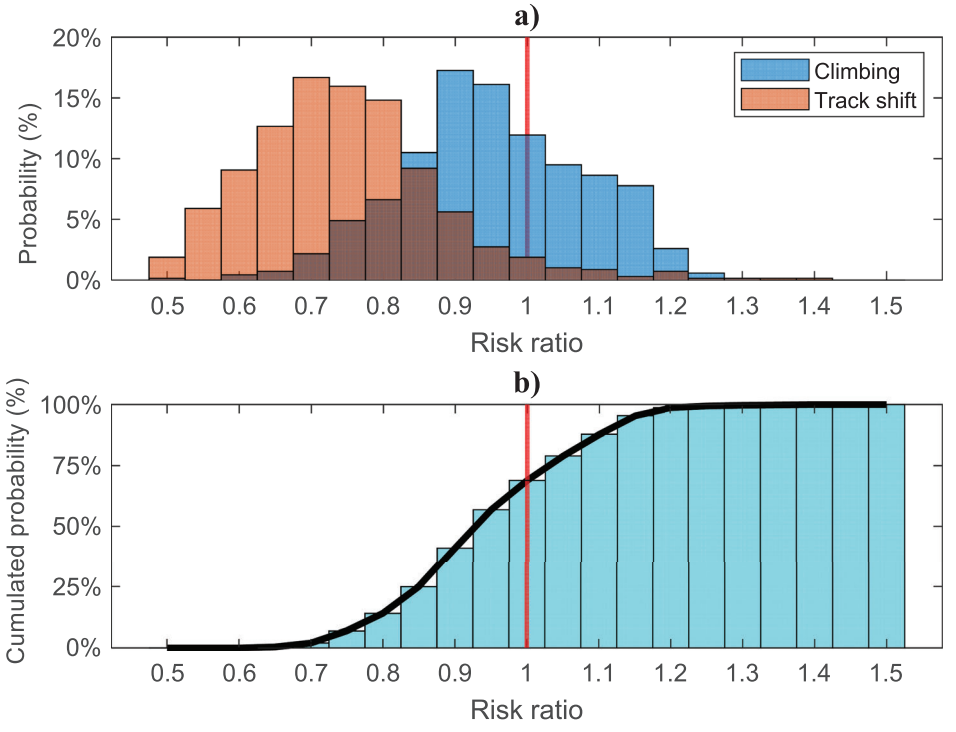


Figure 7. Risk ratios: (a) probability distribution of risk ratios for flange climbing and track shift, and (b) cumulative probability distribution of the global risk ratio.

different range with respect to the ΣY, normalised error-based metrics are considered. Precisely, the relative RMSE (rRMSE) and mean absolute percentage error (MAPE) metrics are introduced, which are calculated as:

$$rRMSE = \frac{RMSE}{\sqrt{\frac{1}{N}\sum y_i^2}} \tag{16}$$

$$MAPE = \frac{100}{N}\sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{17}$$

where y is any of the 20 scalar outputs extracted from MB simulations, $\hat{y}$ is the corresponding predicted scalar output, calculated with kernel regression, i is a summation counter and finally N is the number of observations.

Figures 10a and b show the frequency observation of the R^2 score and rRMSE, respectively, obtained with increasing numbers of training samples for the 20 scalar surrogate models. Please note that since the MAPE and rRMSE normalised error-based metric feature the same trends, only the rRMSE is displayed, for the sake of brevity. For each bar plot, the x-axis is either the R^2 or rRMSE value, while the y-axis corresponds to the relative number (i.e. probability) of surrogate models with the considered metrics value falling within a certain range. With increasing number of training samples, the frequency distribution for the R^2 score tends to shift towards right and becomes narrower, meaning that the average value of R^2 increases and its variance decreases for all models. On the other hand, the

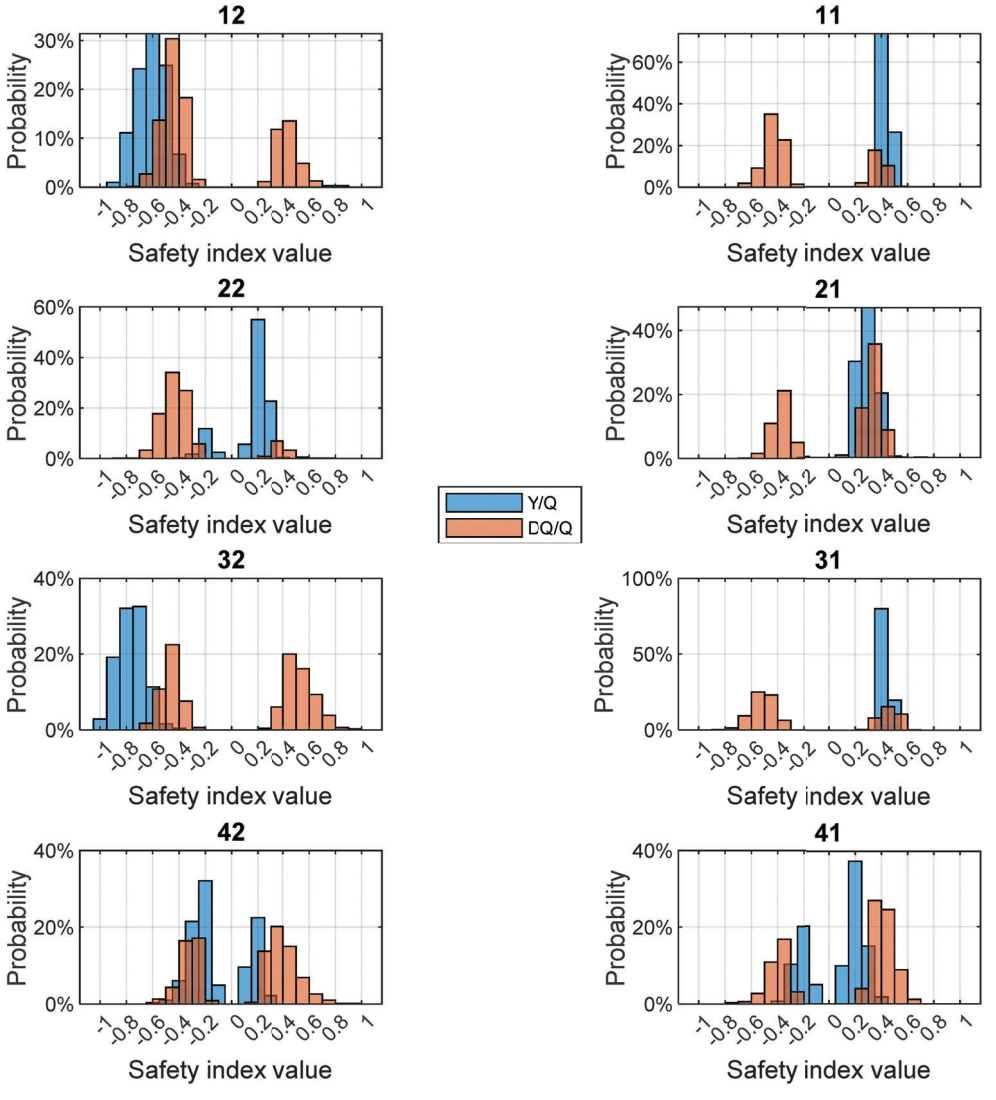


Figure 8. Frequency observation bar plots for the derailment and unloading ratios recorded on each wheel.

rRMSE metrics in Figure 10b tend to shift towards left (lower values) and become narrower with increasing number of training samples. This suggests a progressive improvement of the predictive performance of the LS-SVM surrogate models in the considered range of training samples.

To further prove the prediction capabilities of the trained LS-SVM surrogate models, the bar plots in Figures 11a and b display the values of R^2 and rRMSE for each of the 20 derived surrogate models, for an increasing number of training samples. Again, Figure 11 highlights that with an increasing number of training samples, the R^2 score tends to increase, as well, for all derived surrogate models used to compute the running safety indexes. It

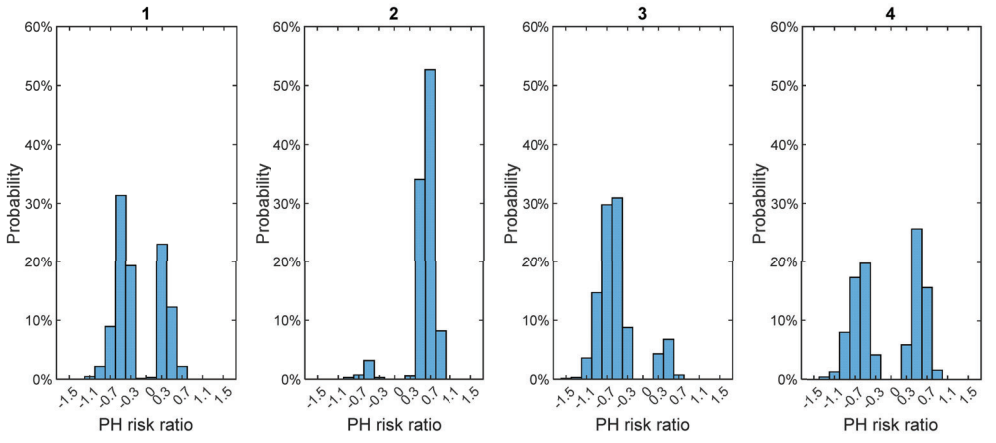


Figure 9. Frequency observation bar plots for the track shift risk ratio recorded on each wheelset.

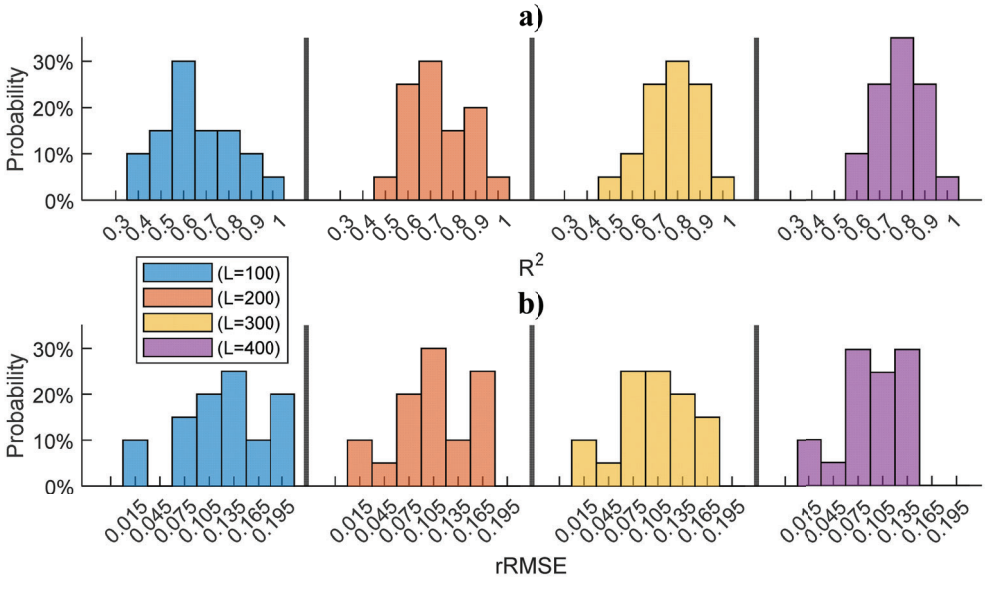


Figure 10. Frequency observation bar plots for the metrics used to assess the capabilities of the LS-SVM models: (a) R^2 score and (b) rRMSE.

can be observed from Figure 11a that the surrogate models derived to predict the derailment ratio Y/Q on the outer wheels of the leading wheelsets belonging to the front and rear bogies (wheelsets 1 and 3) feature a very high accuracy, with an R^2 score above 0.8 when the number of training samples is equal to at least 200. This aspect is particularly relevant, as these indicators are those which exhibit the strongest correlation with potential derailment events, see Figure 8, and therefore they are critical for the evaluation of the wagon running safety. Similarly, from Figure 8b, it can be observed that for all safety indexes, the rRMSE metrics decreases with increasing number of training samples. With 400 training samples, the Y/Q and DQ/Q on the outer (left) wheel of the third wheelset,

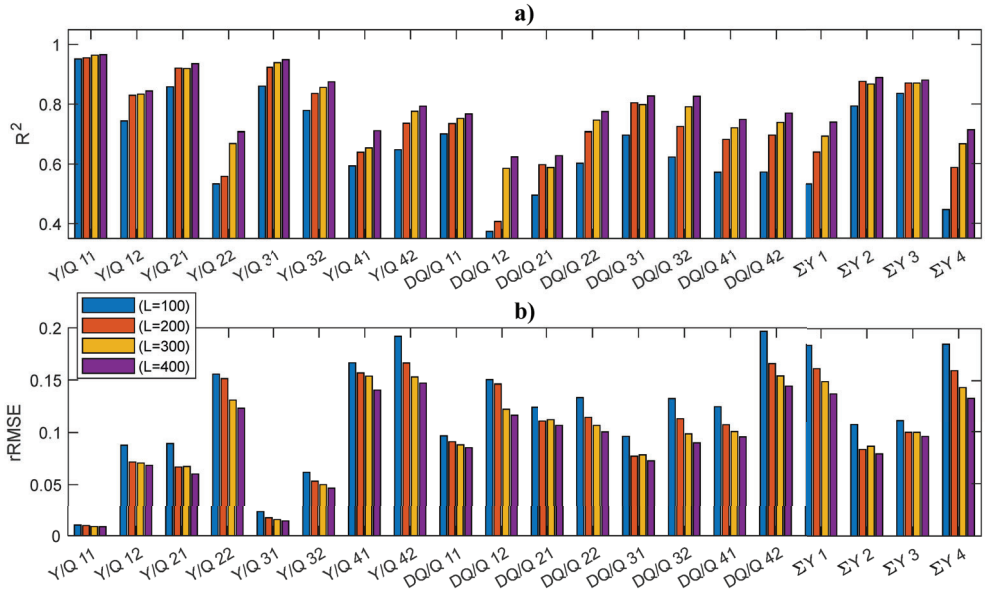


Figure 11. Bar plots with values of R^2 score of the LS-SVM models for increasing number of training samples.

Table 3. Values of the statistical metrics used to assess the performances of the LS-SVM corresponding to 400 training samples (all quantities in the table are non-dimensional, except for the RMSE and MAE for the sum of guiding forces, which have units of N).

Parameter	R^2	RMSE	MAE	rRMSE	MAPE
Y/Q 11	0.966	0.0036	0.0029	0.0082	0.66
Y/Q 12	0.844	0.0423	0.0332	0.0684	5.53
Y/Q 21	0.936	0.0175	0.0133	0.0601	4.97
Y/Q 22	0.711	0.0269	0.0184	0.1225	8.98
Y/Q 31	0.949	0.0058	0.0044	0.0137	1.05
Y/Q 32	0.874	0.0357	0.0257	0.0465	3.53
Y/Q 41	0.710	0.0316	0.0249	0.1407	12.02
Y/Q 42	0.793	0.0354	0.0269	0.1471	12.57
DQ/Q 11	0.765	0.0362	0.0281	0.0857	6.97
DQ/Q 12	0.622	0.0520	0.0381	0.1166	8.85
DQ/Q 21	0.635	0.0378	0.0287	0.1056	8.20
DQ/Q 22	0.776	0.0434	0.0278	0.1003	6.85
DQ/Q 31	0.827	0.0366	0.0284	0.0727	5.86
DQ/Q 32	0.821	0.0472	0.0355	0.0914	7.05
DQ/Q 41	0.747	0.0398	0.0296	0.0958	7.35
DQ/Q 42	0.767	0.0564	0.0432	0.1454	12.01
ΣY 1	0.739	2163.5	1690.5	0.137	11.82
ΣY 2	0.889	1762.6	1128.2	0.078	5.43
ΣY 3	0.880	2102.4	1481.9	0.095	7.68
ΣY 4	0.717	2828.2	2229.5	0.132	11.22

as well as the ΣY on the third wheelset feature rRMSE values below 10%. For the sake of completeness, Table 3 presents the values of all the statistical metrics (R⊃2, RMSE, MAE, rRMSE, MAPE), corresponding to the surrogate models built from 400 training samples.

Shifting focus to the binary SVM classifier, its capabilities can be inferred from Figure 12, which provides the confusion matrices for an increasing number of training samples.

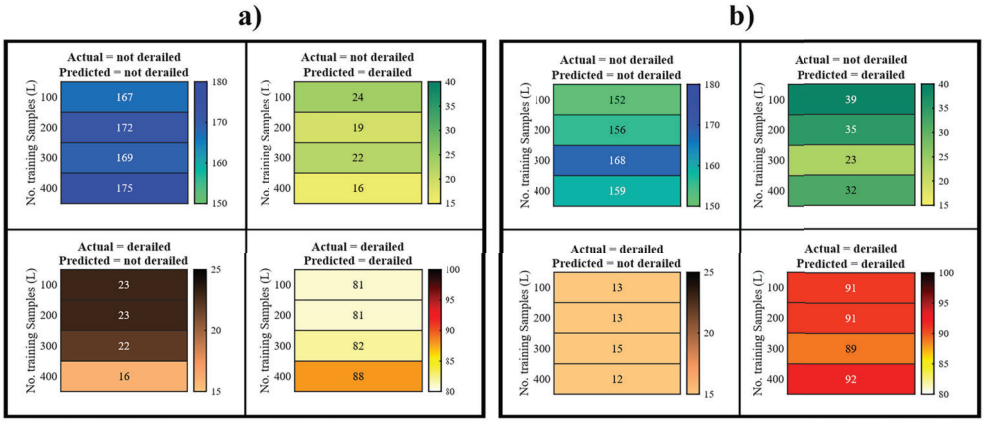


Figure 12. Confusion matrices of SVM classifiers for increasing number of training samples: (a) the basic SVM classifier and (b) the weighted SVM classifier. Please note that different colormaps are used for the 4 zones of the confusion matrices.

Please note that each matrix is a 2 × 2 matrix, where moving along rows, the label of actual data changes from 'not derailed' to 'derailed', while moving along columns, it is the predicted data label that changes from 'not derailed' to 'derailed'. Each cell provides the number of occurrences of a certain combination of actual and predicted labels, with classifier trained with the four numbers of training samples. This means that, for a fixed number of training samples, the sum of the numbers appearing in each cell is equal to 295, i.e. the total number of test samples. Please note that the cells along the matrix diagonal refer to conditions where the predicted label coincides with the actual one, with true negatives in cell (1,1) and true positives in cell (2,2). On the other hand, the out-of-diagonal cells are related to conditions where the SVM classifier misclassifies data, with false positives in cell (1,2) and false negatives in cell (2,1), which is the most undesirable condition.

From Figure 12, it can be observed that, on average, the number of false negatives in cell (2,1) obtained with the weighted classifier is almost halved with respect to the basic classifier. However, halving the weights associated to the training samples belonging to the 'not derailed' class also leads to an increased number of false positives in cell (1,2), as well as to a decrease of the observations of true negatives in cell (1,1). This is not intentional, but the application of weights during the training stage forces the SVM classifier to prefer the 'derailed' class over the 'not derailed' one. The weighted classifier tends to offer a more conservative evaluation, with positive and negative consequences in terms of safety and operativity, respectively.

With 400 training samples, the highest number of true positives and the lowest number of false negatives, see cell (2,1), are obtained, with an intermediate performance in terms of true negatives and false positives. When comparing the two classifiers, it should be observed that the basic model trained with 400 samples is characterised by a slightly higher number of false negatives when compared to the corresponding weighted model, but it has superior performances in overall data labelling.

Concerning the computational times, on a reference computer desktop, the average wall clock time required for one MB run, including time integration and output measurements,

is around 210 s, for an average simulated time of 71 s. With 400 samples, training the 20 LS-SVM surrogate models takes around 58 s, while 7 s are needed for each binary classifier. After training, the evaluation of the 20 surrogate models on 295 test samples requires 31.2 ms, while it takes 7 ms for data labelling with a classifier. This means that when the LTD simulations outputs are available, the evaluation of the running safety indexes with the trained surrogate models is faster than running a MB simulation by more than 6 orders of magnitude. Overall, it can be stated that, after training, the evaluation time of the surrogate models and binary classifiers is negligible, and it is compatible with on-board implementation needs.

4. Discussion

The surrogate models built in the present paper feature good numerical accuracy and excellent computational efficiency with the advantage of a small training sample dataset. This is enabled by the selected kernel regression techniques, which considerably simplify the training procedure compared to more flexible structures such as artificial neural networks, which typically require tuning several architectural and optimisation hyperparameters (e.g. number of layers, number of neurons per layer, learning rate, etc.). Additionally, with respect to traditional linear regressions, for static modelling, the number of unknowns to be determined with kernel regressions is independent of the complexity of the computational model providing the input-output relationship.

The surrogate models are built considering random wagons, that do not correspond to existing wagon types, despite featuring values of length over buffers, bogie spacing, overhang and axle load complying with European rules. This increases their applicability to a wide range of real freight train compositions, built with combinations of different existing wagon types. At the same time, the developed framework can be easily adapted to consider a specific fleet of freight wagon types, hence better targeting the dynamic behaviour of a smaller set of wagons. This would bring possible advantages in terms of numerical accuracy for the considered wagon fleet; however, it would strongly affect the possible range of applicability of the surrogate models.

Furthermore, it is important to remark that for the sake of simplicity, the surrogate models shown in this paper were built for emergency air brake operations on a level track, with no gradients. Gradients on the line affect the load transfers on the vehicle but also change the values of the in-train forces during the train operation, hence modifying the values of the calculated running safety indexes. However, introducing the track gradient among the simulation input parameters drawn via LHS is not straightforward, as the vehicle running behaviour is not only affected by the track gradient, but also by its location along the track, that in general does not coincide with the clothoid transition from straight to curved section. Therefore, the effect of line gradients is demanded to future upgrades of the activity, as it will require a dedicated strategy to deal with the definition of the location of the track elevation variation, with a slight increase in the complexity of the training stage.

Similarly, the surrogate models were obtained considering an ideal track without irregularities apart from the local twist error at the end of the clothoid transition. This was intentionally done to simplify the training stage minimising the effect of stochastic quantities. Nonetheless, future developments of the activity will deal with irregularities, training different models for different levels of stochastic irregularities. This will be beneficial to

define a statistical range of variation for the calculated running safety indexes, depending on the irregularity level.

One of the biggest limitations of the derived surrogate models based on kernel regressions is that they only provide static output values, whilst not calculating the time history of the running safety indexes. Future activities will deal with an analysis and comparison of advanced methods for the derivation of dynamic surrogate models, able to predict time histories for the quantities of interest.

Regardless of the previously mentioned limitations, the combination of numerical accuracy and computational efficiency of the developed surrogate models can pave the way for several possible applications, whose practical implementation, however, is outside the scope of the present paper. The straightforward application is the enhancement of the modelling capabilities of common LTD simulators that could be provided with DT modules of MB simulations for a fast estimation of the running safety indexes under specific operating conditions. This could simplify the offline optimisation of freight train compositions with the aim of maximising running safety, based on a direct estimation of the running safety indexes, rather than solely on the in-train forces. Additionally, DTs could be implemented on-board freight trains within ATO/ATP or monitoring systems. Nonetheless, the on-board implementation is not immediate, as several challenges need to be faced. In fact, the developed surrogate models require knowledge of the main characteristics of the wagons in the train composition and track layout. Concerning train composition data, they could be extracted from train composition sheets with automatic methods, e.g. OCR techniques. The payload could be additionally estimated from the weighing valves in the air brake system. Nonetheless, this type of measurement is not easy to perform, due to the lack of electrical power on-board freight wagons. Concerning track layout data, two possible solutions could be deployed. The first solution is the determination of curvature and slopes from a pre-constructed database based on the GPS location, as GPS is usually available on locomotives. The second solution could be an online estimation by means of sensors installed in the locomotive [33]. Similar challenges should be faced also for the application of the derived DTs within online monitoring systems.

5. Conclusions

The present paper proves that kernel-based techniques can be effectively adopted to derive surrogate models for the prediction of the running safety indexes as well as to build binary classifiers that can discriminate between safe and unsafe conditions. The main outcomes of the present work are given in the following bulleted list.

- With respect to previous works by the authors, the kernel regressions obtained in this paper include a larger number of input parameters, including vehicle data of a group of three wagons. This makes the derived models independent from the specific train configuration considered during training.
- When the number of training samples increases from 100 to 400, the R^2 score of the 20 surrogate models increases as well, with average values around 0.8. At the same time, the considered error-based metrics (RMSE, MAE, rRMSE and MAPE) tend to decrease. This suggests that good prediction capabilities can be achieved with the benefit of a limited number of training samples.

- From the LTD and MB simulations, it is found that the more critical wheels are wheels 12 and 32, namely the outer wheels of the leading wheelset of the front and rear bogies. The R^2 score of the LS-SVM surrogate models for the calculation of the Y/Q on these wheels is above 0.85, while the rRMSE is below 5%, thus ensuring that good prediction capabilities can be achieved on the critical wheels.
- With a sufficient, but still limited, amount of training samples, the basic SVM classifier features excellent prediction capabilities. The weighted SVM classifier is a valuable solution to further reduce the number of false negatives without requiring additional training samples. However, this comes at the cost of worsened performances in the detection of true negatives and false positives, highlighting that a trade-off exists between safety and overall accuracy of the classifier.
- Once the kernel regressions are trained, the evaluation of the safety indexes with the surrogate models is about 6 order of magnitude faster compared to running MB simulations, with clear advantages in terms of computational times.

In view of their excellent computational efficiency, the surrogate models can be used as DTs of MB simulations, hence leading the path towards several possible upgrades and applications. The straightforward application is adding the trained regressions within common LTD codes, thus making it possible to upgrade their capabilities and estimate the running safety indexes. Therefore, the trained regressions could be applied within numerical optimisation routines, e.g. to identify optimal train compositions that maximise safety. Similarly, as LTD simulations are known to have real-time computational performance capabilities, the use of DTs replacing MB simulations could thrust the on-board live assessment of the train running safety. However, for a proper detection of unsafe conditions, and to minimise the risk of false negatives, it is strongly recommended that the outputs of the surrogate models are combined with measured data to improve the robustness of on-board systems and algorithms. Overall, it is believed that the adoption of fast-to-evaluate surrogate models acting as DTs of MB simulations can enhance the safety and competitiveness of railway freight transportation. Nonetheless, the implementation of digital twins within ATO/ATP systems or monitoring units requires solving specific challenges, such as the online estimation of the track layout and main train composition data.

Author contributions

CRediT: **N. Bosso:** Conceptualization, Methodology, Supervision; **L. Cantone:** Conceptualization, Data curation, Software, Writing – review & editing; **M. Magelli:** Data curation, Software, Visualization, Writing – original draft; **R. Trinchero:** Data curation, Software, Writing – original draft, Writing – review & editing; **N. Zampieri:** Conceptualization, Funding acquisition, Methodology, Software, Writing – review & editing

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Ministero dell'Università e della Ricerca [grant number FISA-2023-00314].

References

[1] Thompson EA, Lu P, Alimo PK, et al. Revolutionizing railway systems: a systematic review of digital twin technologies. High-speed Railw. 2025;3(3):238–250. doi:10.1016/j.hspr.2025.05.005

[2] Bosso N, Magelli M, Zampieri N. Digital twins for rail vehicle designs. In: Spiryagin M, Bruni S, Cole C, editors. Handbook on digital twin and artificial intelligence techniques for rail applications. Boca Raton (FL): CRC Press; 2025. p. 119–132.

[3] Spiryagin M, Edelmann J, Klinger F, et al. Vehicle system dynamics in digital twin studies in rail and road domains. Veh Syst Dyn. 2023;61(7):1737–1786. doi:10.1080/00423114.2023.2188228

[4] EN 14363:2016. *Railway applications - testing and simulation for the acceptance of running characteristics of railway vehicles - running behaviour and stationary tests*.

[5] Wu Q, Spiryagin M, Cole C. Longitudinal train dynamics: an overview. Veh Syst Dyn. 2016;54(12):1688–1714. doi:10.1080/00423114.2016.1228988

[6] Cole C, Spiryagin M, Wu Q, et al. Modelling, simulation and applications of longitudinal train dynamics. Veh Syst Dyn. 2017;55(10):1498–1571. doi:10.1080/00423114.2017.1330484

[7] Spiryagin M, Wu Q, Cole C. International benchmarking of longitudinal train dynamics simulators: benchmarking questions. Veh Syst Dyn. 2017;55(4):450–463. doi:10.1080/00423114.2016.1270457

[8] Wu Q, Spiryagin M, Cole C, et al. International benchmarking of longitudinal train dynamics simulators: results. Veh Syst Dyn. 2018;56(3):343–365. doi:10.1080/00423114.2017.1377840

[9] IRS 40421/E/2:2021. *Rules for the consist and braking of international freight trains*.

[10] UIC 530-2:2008. *Wagons - running safety*.

[11] Krishna VV, Berg M, Stichel S. Tolerable longitudinal forces for freight trains in tight S-curves using three-dimensional multi-body simulations. Proc Inst Mech Eng F: J Rail Rapid Transit. 2019;234(5):454–467. doi:10.1177/0954409719841794

[12] Magelli M, Zampieri N. A novel approach for longitudinal train dynamics simulations with multibody codes. Veh Syst Dyn. 2025;63(5):978–996. doi:10.1080/00423114.2024.2362949

[13] Magelli M, Corrêa PHA, Antunes dos Santos A. Assessing the effect of modelling approaches for longitudinal train dynamics simulations with a multibody code. Proc Inst Mech Eng K: J Multi-body Dyn. 2025;239(4):326–338. doi:10.1177/14644193251368540

[14] Bosso N, Magelli M, Trinchero R, et al. Application of machine learning techniques to build digital twins for long train dynamics simulations. Veh Syst Dyn. 2024;62(1):21–40. doi:10.1080/00423114.2023.2174885

[15] Bosso N, Magelli M, Zampieri N. Long train dynamic simulation by means of a new in-house code. In: G Passerini, JM Mera, R Takagi, editors. In: *WIT transactions on the built environment* 199. Southampton: WIT Press; 2020. p. 249–259. doi:10.2495/CR200231

[16] Bosso N, Magelli M, Zampieri N. Development and validation of a new code for longitudinal train dynamics simulation. Proc Inst Mech Eng F: J Rail Rapid Transit. 2021;235(3):286–299. doi:10.1177/0954409720923497

[17] Bosso N, Cantone L, Gugliotta A, et al. Introduction of Digital twins in the longitudinal train dynamics simulation of freight train air brake operations. In: Huang W, Ahmadian M, editors. Advances in dynamics of vehicles on roads and tracks III. IAVSD 2023. Lecture Notes in Mechanical Engineering. Ottawa: Springer Nature Switzerland; 2025. p. 297–306. doi:10.1007/978-3-031-66971-2_32

[18] Cantone L. TrainDy: The New Union Internationale Des Chemins de Fer Software for Freight Train Interoperability. Proc Inst Mech Eng F: J Rail Rapid Transit. 2010;225(1):57–70. doi:10.1243/09544097JRRT347

[19] Bernal E, Wu Q, Spiryagin M, et al. Augmented digital twin for railway systems. Veh Syst Dyn. 2024;62(1):67–83. doi:10.1080/00423114.2023.2194543

[20] Zhang S, Huang P, Constable T, et al. A machine learning approach for predicting railway in-train forces from ATO measurements. In: Huang W, Ahmadian M, editors. Advances in dynamics of vehicles on roads and tracks III. IAVSD 2023. Lecture Notes in Mechanical Engineering. Ottawa: Springer Nature Switzerland; 2025. p. 16–22. doi:10.1007/978-3-031-66971-2_3

[21] Zhang S, Huang P, Yan W. A data-driven approach for railway in-train forces monitoring. Adv Eng Inform. 2024;59:102258. doi:10.1016/j.aei.2023.102258
[22] Wu Q, Cole C, Spiryagin M, et al. Freight train air brake models. Int J Rail Transp. 2023;11(1):1–49. doi:10.1080/23248378.2021.2006808
[23] Bosso N, Magelli M, Zampieri N. A numerical method for the simulation of freight train emergency braking operations based on the UIC braked weight percentage. Rail Eng Sci. 2023;31(2):162–171. doi:10.1007/s40534-022-00296-9
[24] Bosso N, Magelli M, Zampieri N. Dynamical effects of the increase of the axle load on European freight railway vehicles. Appl Sci. 2023;13(3):1318. doi:10.3390/app13031318
[25] Benchmark Problem Results and Assessment Bogies with Steered or Steering Wheelsets. ORE (ERRI). 1993.
[26] Kalker JJ. A fast algorithm for the simplified theory of rolling contact. Veh Syst Dyn. 1982;11(1):1–13. doi:10.1080/00423118208968684
[27] Suykens J, Van Gestel T, De Brabanter J, et al. Least squares support vector machines. Leuven: World Scientific; 2002.
[28] Vapnik VN. Statistical learning theory. Wiley; 1998.
[29] Trinchero R, Larbi M, Torun HM, et al. Machine learning and uncertainty quantification for surrogate models of integrated devices with a large number of parameters. IEEE Access. 2019;7:4056–4066. doi:10.1109/ACCESS.2018.2888903
[30] Ghojogh B, Crowley M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. *arXiv (preprint)*, 2019.
[31] Soleimani N, Trinchero R, Canavero FG. Bridging the gap between artificial neural networks and kernel regressions for vector-valued problems in microwave applications. IEEE Trans Microwave Theory Tech. 2023;71(6):2319–2332. doi:10.1109/TMTT.2022.3232895
[32] De Brabanter K, Karsmakers P, Ojeda F, et al. LS-SVMlab Toolbox User's Guide version 1.8. Leuven-Heverlee: ESAT-SISTA, Katholieke Universiteit Leuven; 2011.
[33] Rodríguez M, Urda P, Escalona JL. Calculation of the design horizontal geometry of a railway track using inertial sensors and GPS measurements recorded in line vehicles. Veh Syst Dyn. 2025: 1–22. doi:10.1080/00423114.2025.2516102

Appendix 1: MB model

The MB model is built defining a reference system with x-axis along the running direction, z-axis along the vertical direction (pointing downwards) and y-axis along the lateral direction determined with the right-hand rule. The following Table A1 provides the values of inertial properties and joint type defined for each body in the MB model. On the other hand, Table A2 summarises the main force elements defined in the model.

Table A1. Joint and inertial properties of the bodies in the MB model.

Body	Joint type	Mass (kg)	Ixx (kg·m^2)	Iyy (kg·m^2)	Izz (kg·m^2)
Half-carbody (outer wagons)	General rail track joint (6 DOFs)	5410–40410 kg depending on LHS (total weight on rails limited to 90 ton)	Scaled with axle-load (59188 kg·m^2 @ 22.5 ton)	Scaled with axle-load and (721517 kg·m^2 @ 22.5 ton)	Scaled with axle-load (726122 kg·m^2 @ 22.5 ton)
Half-carbody (middle wagon)	General rail track joint (6 DOFs)	5410–25410 kg depending on LHS (total weight on rails limited to 90 ton)	Scaled with axle-load (59188 kg·m^2 @ 22.5 ton)	Scaled with axle-load (721517 kg·m^2 @ 22.5 ton)	Scaled with axle-load (726122 kg·m^2 @ 22.5 ton)
Bogie frame	General rail track joint (6 DOFs)	1450	1103	715	1783
Axle-box	1 DOF (rotation along y)	170	1.48	6.58	5.23
Wheelset	General rail track joint (6 DOFs)	1080	750	Scaled with tare load (110 kg·m^2 @ 28 ton)	750
Ballast	Rail track sleeper (3 DOFs: z, y, roll)	330	10	1	10
Inner buffers	1 DOF (x translation)	150	-	-	-
Outer buffers	0 DOF	150	-	-	-

Table A2. Description of the main connection force elements in the MB model.

Connection	Description
Primary suspension	16x Shear spring elements (2 elements per axle-box): empty-loaded springs 8x Stick-slip force elements (1 element per axle-box): Lenoir link 8x Stick-slip force elements (1 element per axle-box): Lenoir link piston
Secondary suspension	2x Bushing elements (1 element per half-carbody): centreplate 2x Expression elements providing friction torques (1 element per half-carbody): centreplate friction. 4x Nonlinear spring damper elements (2 elements per half-carbody): side bearers 4x Stick-slip elements (2 elements per half-carbody): side bearer friction 4x Spring damper elements (2 elements per half-carbody): side bearer bump stops
Half-carbody connection	1x Bushing with torsional stiffness value scaled with bogie-spacing (281000 Nm/rad @ 12 m).
Buffer-hook connection (inner systems)	2x Hysteresis elements: Hook mechanical impedance characteristics 4x Hysteresis elements: Buffer mechanical impedance characteristics 4x Hertzian contact elements: contact between buffer heads.