To cite this article:
Gernoth Götz & Oldrich Polach (2018) Verification and validation of simulations in a rail vehicle certification context, International Journal of Rail Transportation, 6:2, 83-100, DOI: 10.1080/23248378.2017.1412839

Verification and validation of simulations in a rail vehicle certification context

Gernoth Götz[a*], Oldrich Polach[b]
[a] *Institute of Rail Vehicles, Berlin Institute of Technology, 10623 Berlin, Germany*
[b] *Consultant, 8212 Neuhausen am Rheinfall, Switzerland*
**Corresponding author email: Gernoth.Goetz@TU-Berlin.de*

Abstract: This article presents recent progress in the application of validation methodologies for the assessment of rail vehicle dynamics by computer simulation. Peculiarities of vehicles' on-track tests, which represent the experiments, used for model validation, are discussed and the terms validation as well as application domain are introduced. This is followed by a comparison of recently suggested validation metrics for rail vehicle models. The validation approach developed in the DynoTRAIN project and introduced in EN 14363:2016 [1] is applied to investigate the impact of signal synchronisation and to evaluate new findings regarding the effect of section selection used for the validation. The DynoTRAIN validation methodology shows a small sensitivity to the inaccuracy of the signals synchronisation. The result of this model validation assessment remains either "validated" or "not validated", rather independently of the selected set of test sections. This confirms the robustness of the DynoTRAIN validation approach.

Keywords: Rail vehicle; vehicle dynamics, multi-body dynamics simulations, model verification, model validation

1. Introduction

The terms Verification and Validation (V&V) are used in various contexts and meanings. The ASME V&V Guide [2] provides following definitions: Verification is the process to determine the accuracy of the model implementation and to examine, if the model represents the conceptual description. The validation process determines the degree of quantitative accuracy of the simulation model with respect to the intended use. This article considers V&V in context of these definitions. V&V are subject of various papers from many different fields ranging from engineering and physics to operations topics [3]. However, only few publications are dedicated to V&V of rail vehicle multi-body simulations, although they represent an indispensable part of rolling stock development and engineering [4].
The development of computer technology enabled possibilities for engineers to simulate running dynamics of rail vehicles. Thus, computer codes were continuously further developed and results from various programs were verified against each other. Results from basic simulation tasks demonstrated a good forecast of the expected running behaviour. Later, more powerful and reliable multi-body simulation tools supported more and more rail vehicles' design and development process or were used to investigate complex vehicle-track-interaction issues. Nowadays, they are also largely recognized as a possible means to reduce costs and efforts for vehicle certification by reducing the amount of physical on-track testing under specific conditions. Consequently, the option of "virtual testing" by simulations has been introduced in standards dealing with testing of rail vehicles [1, 5, 6, 7]. But the essential condition for an application of simulation instead of physical testing is the validation of simulation model.
The standards and documents suggesting the use of simulations for the assessment of rail vehicle running dynamics contain also requirements on model validation; see UIC Code 518:2009 [5],

EN 14363:2016 [1], GM/RT2141 [6], EN 15273-2:2013 [8] and EN 14067-6:2010 [9]. However, those validation requirements relay on available tests rather than on systematic approaches based on theory of V&V. Moreover, they are often not mandatory, and do not contain the validation metric and the explicit quantitative validation limits (matching errors). Only few recently published studies discuss the possible validation metric and validation limit values with the aim to replace reviewer's subjective assessment by quantitative objective criteria [10, 11, 12, 13, 14].
This article provides a review of recent progress in V&V of rail vehicle simulation models and presents new investigations for applications in rail vehicle certification related to validation quantities, signal synchronisation and selection of measurement samples used for validation of simulation models. Chapter 2 discusses briefly the verification of multi-body simulation codes used for rail vehicle dynamics simulations. Chapter 3 gives a review of model validation methods for rail vehicle dynamics. The Sections of Chapter 3 discuss the peculiarities of experimental validation in rail vehicle dynamics, model validation methods in rail vehicle dynamic, validation and application domains, validation quantities and validation metrics. Chapter 4 includes investigations of effects on model validation results considering the influence of selected validation sections and the impact of inexact synchronisation of simulation and measurements signals. The conclusions are summarized in Chapter 5.

2. Verification of computer codes for rail vehicle dynamics

The development of information technology was accompanied by the development of program codes to investigate rail vehicle dynamics. Comparisons of simulation results with measurements demonstrated over the years a growing level of agreement (e.g. wheel-rail forces in curves, vehicle's critical speed) supporting the verification of simulation codes developed in different countries by different experts. As the measurements are always connected with uncertain parameters, a benchmark to compare results of different simulation program codes appeared to be initially as a useful verification approach [15].
The first well-known benchmark of simulation codes related to rail vehicle dynamics is the Benchmark ERRI B 176 (also ORE B 176) [16] introduced in the framework of call for bogies with improved behaviour in curves [17]. The aim of this benchmark was to assess the reliability of computer codes used for the development of tendered bogies with steered or steering wheelsets. The benchmark contained calculations using linear and nonlinear models of a passenger coach. Altogether, 5 railway companies, 11 rolling stock suppliers and 2 research institutes expressed their interest to participate and received tasks in April 1991. The participants used either commercial tools (VAMPIRE, MEDYNA) or computer codes developed by their companies. The benchmark results were provided by 16 participants, but only 7 of them submitted results for all tasks. The benchmark showed that the simulation results were often inconsistent. The tendencies of most results were plausible, but some of them were doubtful.
The Manchester Benchmark published in 1999 [18] specified simulations of two vehicles, namely, a passenger coach similar to the ERRI-Benchmark vehicle and a two-axle freight wagon with friction suspension. Four track cases were defined to allow comparisons of the capabilities of computer simulation packages. Suppliers of five major commercial simulation tools carried out the simulations: VAMPIRE, GENSYS, SIMPACK, NUCARS and ADAMS/Rail-MEDYNA (ADAMS/Rail and MEDYNA participated as one tool because of their close cooperation at that time). The progress in the development of commercial simulation packages has been demonstrated in the benchmark results [19]. The simulations carried out to compare different simulation tools gave good agreement with each other on predicting lateral to vertical wheel force ratios on twisted track as well as in other benchmark tasks.
Later, development of information technology and capabilities of multi-body dynamics tools extended application areas further. The mathematical models of crucial coupling elements like forces in the contact between wheel and rail are meanwhile proven by comparisons of simulations of various vehicles as well as by comparisons with laboratory roller rig experiments. The simulation codes are able to provide reasonably accurate results supposing that model parameters possess required accuracy and no modelling errors occur.
The simulation codes for rail vehicle dynamics can be considered as verified as they are commonly

used in rolling stock development, the users trust their capabilities and the outputs of multi-body simulations are widely accepted. However, each quantitative application of simulations, particularly in the certification process, requires a model validation with the aim to prove that there are no errors or mistakes in the model and the model parameters are sufficiently accurate for the intended application. The model quality can be improved by justification and calibration of uncertain model parameters based on comparisons with measurements. This is considered here as a part of the model validation.

3. Review of model validation methods for rail vehicle dynamics

3.1 Peculiarities of experimental validation in rail vehicle dynamics

One example typically presented in publications about model validation is a tapered cantilever beam under a distributed load along a portion of the beam, see e.g. article by Schwer [20]. The investigated output is the beam deformation, while the primary factors contributing to variability of this quantity are beam material and torsional rigidity of the wall fixture.

Compared to this simple example, rail vehicle dynamics represents a more complex area with many more parameters. The quantities measured during on-track tests for vehicle acceptance of running characteristics according to EN 14363:2016 [1] using normal measuring method, are wheel/rail forces and accelerations. The running dynamics tests are carried out under specified target test conditions representative for operating conditions in European networks. The test runs and statistical analyses of their results are separated in to so called “test zones” with the focus to different sets of test conditions: Straight track and very large radius curves (test zone 1), large radius curves (zone 2), small radius curves (zone 3) and very small radius curves (zone 4). Measurements in each test zone have to contain specified number of track sections from each test zone with the length given in this standard (between 70 m and 500 m dependent on the test zone), which are then used for the assessment of running characteristics. The statistical analysis of the measurements in several track sections allows the determination of estimated median values and maximum values which are representative for the target test conditions in the particular test zone. Then, these estimated maximum and median values are compared with limit values.

The result of the on-track test for the acceptance of running characteristics thus depends besides vehicle parameters also on other parameters and conditions, namely: The measurement method and measured quantities, track layout, track irregularities, rail profiles of selected track section, weather conditions during testing and rail contamination (represented in the simulation model by the friction coefficient between wheel and rail) as well as the statistical evaluation. Although the test conditions are specified by the standard (target parameters for track layout, track quality, wheel/rail contact geometry and weather conditions), the real world may vary due to errors in the measurement of track layout and rail profiles as well as due to uncertainty regarding the wheel/rail friction coefficient.

3.2 Review of model validation methods

Fries et al. [13] have recently discussed the validation of rail vehicle dynamics simulation models. Their paper includes some broad suggestions on criteria for model validation, including single-value metrics and frequency domain comparisons. The authors of Ref. [13] recognize that validation criteria depend upon the intended use of the model and many other factors. They investigated the application of metric by Sprague and Geers (see [21, 22] for details about this metric) on a passenger rail vehicle with measurements of accelerations and suggested validation limits for the application of this metric.

Bogojevic and Lucanin [12] proposed new validation metric for validating rail vehicle models based on a comparison between the cumulative distribution functions (CDFs) of simulation and measurement. They developed a model of a freight wagon and evaluated the proposed validation metric by comparisons with measurements, which included wheel/rail forces. The extent of the matching between the CDFs obtained by measurement and simulation is estimated using the Fisher test, which determines the probability of their matching based on their variances and number of points of CDFs. As this metric is not sensitive to the difference of mean values, the result of the Fisher test is combined with a second measure, a relative difference of the mean values of the compared CDF functions, which is divided by the supremum [12] (i.e. the smallest upper bound on a set) of simulation and measurement mean values. The final metric consists of the product of these two measures and belongs to the range between 0 and 1, with the "1" denoting a perfect match of both CDFs.

The European research project DynoTRAIN (2009 – 2013) contained a work package dedicated to model validation. The validation method developed in this project is intended for vehicle acceptance approval and has recently been implemented in to EN 14363:2016 [1]. The exercises carried out in the framework of DynoTRAIN project used models of 11 vehicles, prepared by different modellers using 2 simulation tools Simpack and VOCO, see Polach et al. [11]. The investigated vehicles were tested in the framework of DynoTRAIN project [23] and partly also outside this project. Each running test provided besides acceleration measurements also measurements of wheel/rail forces. The proposed model validation process considers assessment of 12 quantities based on measured forces between wheel and rail and vehicle body accelerations [1, 11], see Table 1. The validation assessment is carried out by evaluating differences between the simulation values and the measurement values for each quantity on a minimum of three sections from each of the four test zones according to EN 14363, i.e. on at least 12 test sections. The test sections have to be selected from full curves and straight track. Their length has to be not shorter than the length according to EN 14363, i.e. between 70 and 500 m (dependent on the test zone and vehicle speed). Each quantity has to be evaluated using at least two signals (e.g. vertical car body acceleration above the leading and trailing bogie), thus, at least 24 pairs of simulated and measured values for each quantity. The validation result is assessed calculating the mean and the standard deviation of differences between simulation and measurement for each particular quantity as shown in Figure 1 on example of rms-values of vertical car body acceleration. The left diagram in Figure 1 displays the simulation values and the measurement values. The right diagram shows the differences, their mean value and standard deviation value, which are compared with the validation limits. Table 1 shows the quantities required for the validation comparisons, their filtering and processing and the validation limits for standard deviations of differences simulation-measurement. The validation limit for mean value of differences simulation-measurement is 2/3 of the corresponding limit for standard deviation. The fulfilment of the validation limits can be easily assessed displaying the results divided by the corresponding validation limits, i.e. as normalized values. A model is assessed as validated only if the magnitudes of all normalized mean values and standard deviations of differences simulation-measurement are not higher than one. Using our example in Figure 1, the mean value and the standard deviation of differences simulation - measurement of vertical car body accelerations are divided by corresponding limit values and displayed in the overall validation assessment in Figure 2. This figure represents an example of a successful validation result because no magnitudes of presented normalized values are higher than one.

Kraft et al. [14] proposed a new approach for validation of rail vehicle models based on on-track measurements analysing 10 vehicle models. The applied measurements of two vehicles included the wheel/rail forces, measurement of other vehicles contained accelerations only. The matching between simulation and measurement is assessed by least-square misfit functions (i.e. distances between simulation and measurement results) normalised with the lower response signal (either measurement or simulation). The misfit functions are calculated for a set of sections. The cumulative frequency distribution of the misfit functions is used to compare the quality of different vehicle models. The matching limit, called model quality indicator, is derived from the uncertainty of the on-track measurement evaluated from the on-track test repeatability analysis.

Table 1. Quantities and limit values for the DynoTRAIN validation methodology (from [10])

Quantity	Notation	Unit	Filtering	Processing	Validation limit for standard deviation
Quasi-static guiding force	Y_{qst}	kN	Low-pass filter 20 Hz	50%-value (median)	5
Quasi-static vertical wheel force	Q_{qst}	kN	Low-pass filter 20 Hz	50%-value (median)	4 (1+0.01 Q_0) Q_0 - static vertical wheel force (kN)
Quasi-static ratio *Y/Q*	$(Y/Q)_{qst}$	-	Low-pass filter 20 Hz	50%-value (median)	0.07
Quasi-static sum of guiding forces	ΣY_{qst}	kN	Low-pass filter 20 Hz	50%-value (median)	6
Guiding force, maximum	Y_{max}	kN	Low-pass filter 20 Hz	0.15%/99.85%-value *)	9
Vertical wheel force, maximum	Q_{max}	kN	Low-pass filter 20 Hz	99.85%-value *)	6 (1+0.01 Q_0) Q_0 - static vertical wheel force (kN)
Ratio *Y/Q*, maximum	$(Y/Q)_{max}$	-	Sliding mean (2 m window, step 0.5 m)	0.15%/99.85%-value *)	0.10
Sum of guiding forces, maximum	ΣY_{max}	kN	Sliding mean (2 m window, step 0.5 m)	0.15%/99.85%-value *)	9
Car body lateral acceleration, rms-value	$\ddot{y}^*_{rms}$	m/s^2	Band-pass filter 0.4 to 10 Hz	rms-value	0.15 **)
Car body vertical acceleration, rms-value	$\ddot{z}^*_{rms}$	m/s^2	Band-pass filter 0.4 to 10 Hz	rms-value	0.15 **)
Car body lateral acceleration, maximum	$\ddot{y}^*_{max}$	m/s^2	Band-pass filter 0.4 to 10 Hz	0.15%/99.85%-value *)	0.40 **)
Car body vertical acceleration, maximum	$\ddot{z}^*_{max}$	m/s^2	Band-pass filter 0.4 to 10 Hz	0.15%/99.85%-value *)	0.40 **)

*) Absolute values of simulated and measured values
**) For freight vehicles and vehicles without bogies or without secondary suspension, these limits have to be doubled

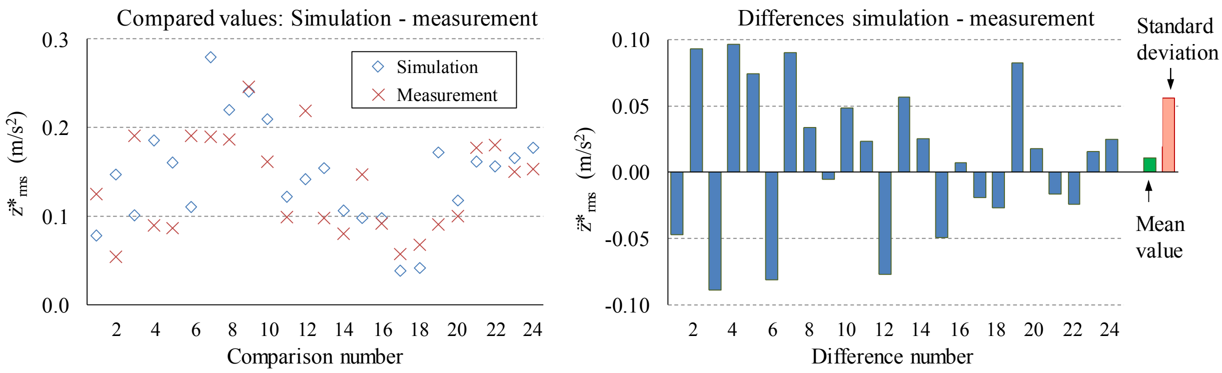

Figure 1. Example of the DynoTRAIN evaluation of mean and standard deviation of differences between simulation and measurement for one validation quantity

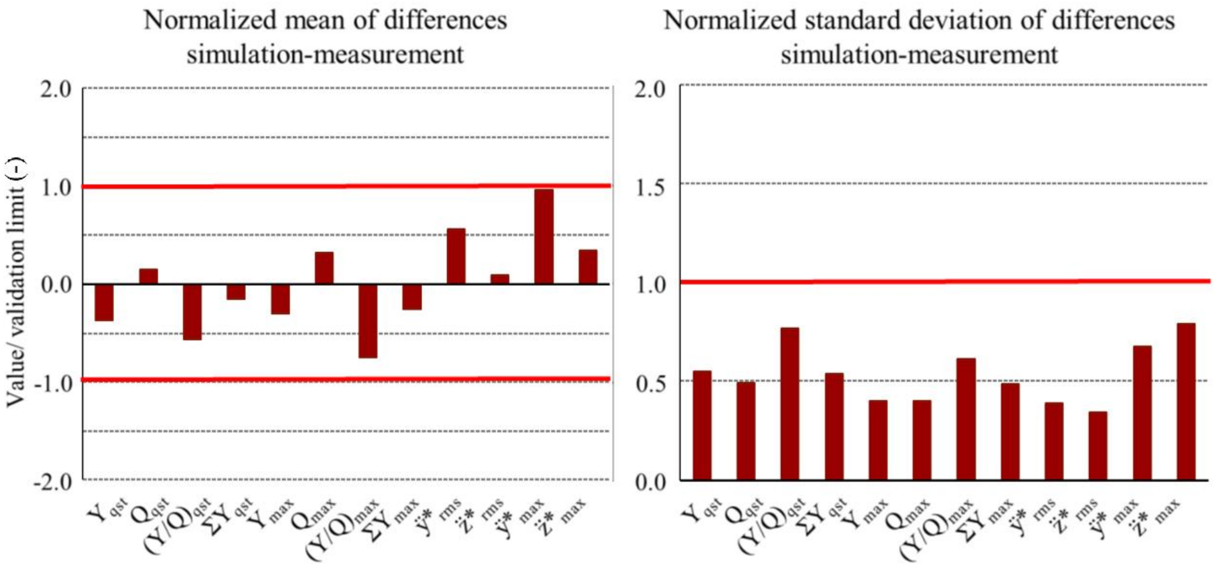

Figure 2. Example of a successful validation result using the DynoTRAIN validation methodology; passenger coach tested in the framework of DynoTRAIN project; a randomly selected set of test sections

3.3 Validation and application domains

The model validation has to consider the application domain, i.e. the intended area of the simulation model usage. The validation itself concerns the validation domain, i.e. the region of system parameters and physical complexity, in which the confidence in the predictive capability of the computation model has been demonstrated quantitatively by satisfactory agreement between simulations and experiments [3]. Outside of the validation domain, there is a degradation of confidence in the quantitative predictive capability of the model. This means that outside of the validation domain, the model may be credible but the quantitative capability has not been demonstrated. The engineering applications seek for the application domain inside of the validation domain or at least with a significant overlap of both domains.

The application domain has an important effect on the validation experiment, quantities as well as the validation metric and target matching between simulation and experiment. For example, the running safety of rail vehicles is related to risk of derailment and track shift, which is most exactly assessed by the measurement of forces between wheel and rail. Comfort assessment is based on the measurement of accelerations in the car body, thus requiring accurate simulations of vehicle body accelerations including the vibration of its elastic structure. The investigation of cross wind and aerodynamics, on other hand, should prove that the main vehicle characteristics affecting the side wind stability are represented with sufficient accuracy, which means accurate simulation of lateral body movement and its effect on the wheel unloading. Thus, the model validation is based on comparisons of vertical wheelset and wheel forces and on the flexibility coefficient (suspension coefficient) used for kinematic gauge calculation according to EN 14067-6:2010 [9]. The dynamic gauging requires accurate simulation of vehicle body displacements. The model validation process thus provides besides other quantities also requirements for the maximum matching error and the average error on the hysteresis loop of the primary and secondary roll angle evaluated in the stationary sway test, see EN 15273-2:2013 [8].

The present article concentrates on simulations of on-track tests for approval of running characteristics according to EN 14363:2016 [1] as used in context of vehicle certification in European countries. The validation experiment and validation domain should therefore consist of running tests (on-track tests). Traditionally, the experiment used for validation of rail vehicle models consists of tests specified and used for the certification anyway rather than from tests specified to cover the targeted validation domain. Stationary tests are used to approve and possibly calibrate the model parameters also when the simulation model is intended for the simulation of running tests only [1, 5, 6, 24]. The application domain can contain parameters outside the validation domain, if the confidence in the quantitative predictive capability of the model is not reduced. In the validation process according to EN 14363:2016 [1] this is believed to be fulfilled if the resulting running dynamics behaviour of the modified simulation model with a parameter outside the validation domain is close to the vehicle's behaviour in the validation domain.

3.4 Validation quantities

As stated previously, quantities to be assessed in validation depend on the application domain. Nevertheless, we have to highlight that the aim of validation is to provide confidence about the overall model accuracy and not only about the matching of a particular quantity.
Considering the on-track tests for approval of running characteristics according to EN 14363:2016 [1], the assessment quantities are:

Running safety:

- Dynamic sum of guiding forces ΣY_{max},
- Dynamic ratio of lateral to vertical wheel force $(Y/Q)_{max}$,
- Instability criterion (assessed by root mean square of the sum of guiding forces ΣY calculated over 100 m distance).

The ratio $(Y/Q)_{max}$ is evaluated in curves for the outer wheels only.

Track loading:

- Quasi-static lateral wheel/rail force Y_{qst},
- Quasi-static vertical wheel/rail force Q_{qst},
- Maximum vertical wheel/rail force Q_{max}.

These wheel/rail forces are evaluated in curves for the outer wheels only.

Oscillation behaviour:

- Car body acceleration in vertical and lateral directions, measured on the vehicle's floor above the running gears, and the lateral bogie frame acceleration.

One validation approach is to consider quantities used for the approval of running characteristics as the quantities relevant for the validation. This means that the validation comparison concentrates on the quantities supposed to achieve high values and assessed in the framework of vehicle acceptance. An agreement between simulation and measurement for a single quantity, however, can be easily achieved by manipulation of uncertain measurement parameters like e.g. friction coefficient between wheel and rail. This is illustrated using examples of two test sections in Figure 3 (an Italian test section, curve radius R = 292 m) and Figure 4 (a French test section, curve radius R = 294 m). Both figures show comparisons of simulated and measured guiding forces using the model of passenger coach that was tested within DynoTRAIN [23]. The simulation results are based on two different wheel/rail friction coefficients: 0.3 and 0.6. The dashed lines limited with plus signs define the full curve part and the dashed lines limited with circle signs define the test section with the length according to EN 14363. If there is a disagreement of quasi-static guiding force on the outer leading wheel, the usual suggestion is to adjust the estimated wheel/rail friction coefficient used in the simulation. As shown in Figure 3, the wheel/rail friction coefficient was increased from 0.3 to 0.6, which improved significantly the agreement on the left (curve outer) and similarly on the right (curve inner) leading wheels. The already high agreement level on the trailing wheels is rather insensitive to the higher friction coefficient and differences remain small. In this example, the justification of wheel/rail friction coefficient considering only the curve outer leading wheel (quantity assessed in context of testing according to EN 14363) provided an overall improvement of agreement.

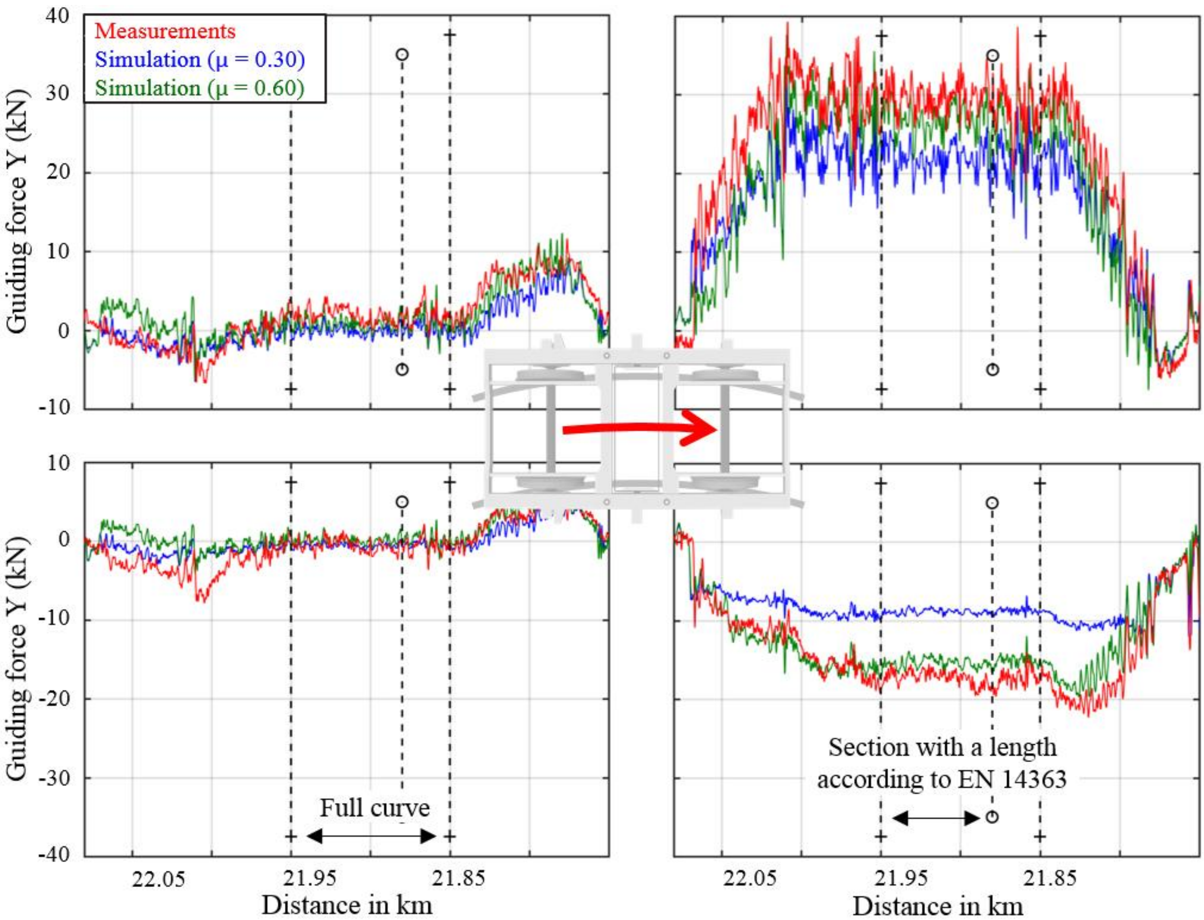

Figure 3. Influence of increasing wheel/rail friction coefficient on Y-forces (an Italian test section, right curve; the plots relate to the nearest wheel in the central picture of the bogie)

In the second example in Figure 4, the justification of wheel/rail friction coefficient from 0.3 to 0.6 also improves the agreement on the outer leading wheel (quantity assessed in tests according to EN 14363). However, at the same time, a significant disagreement of the guiding force occurs on the inner leading wheel. Moreover, further disagreements are observed on the trailing wheels; the forces on the trailing wheels are not reduced by the increase of the wheel/rail friction coefficient from 0.3 to 0.6. The adjustment of the wheel/rail friction coefficient based on the measured guiding force on the curve outer leading wheel is in this case clearly not a correct adjustment considering all contacts between wheel and rail; the observed disagreements are caused by other parameters and require further investigations.

The presented examples show that the comparisons considering all measured quantities provide more reliable assessment of the overall agreement between simulation model and measurement than comparisons considering only quantities assessed during the vehicle acceptance test according to EN 14363. This is why the DynoTRAIN validation method [10, 11] specifies comparisons of more measured quantities, and the assessment of matching between simulation and measurement is carried out comparing maximum as well as quasi-static value of all wheel/rail quantities, e.g. also values not evaluated in context of the vehicle acceptance test according to EN 14363 like the quasi-static sum of guiding forces ΣY_{qst} or maximum value of guiding force Y_{max}.

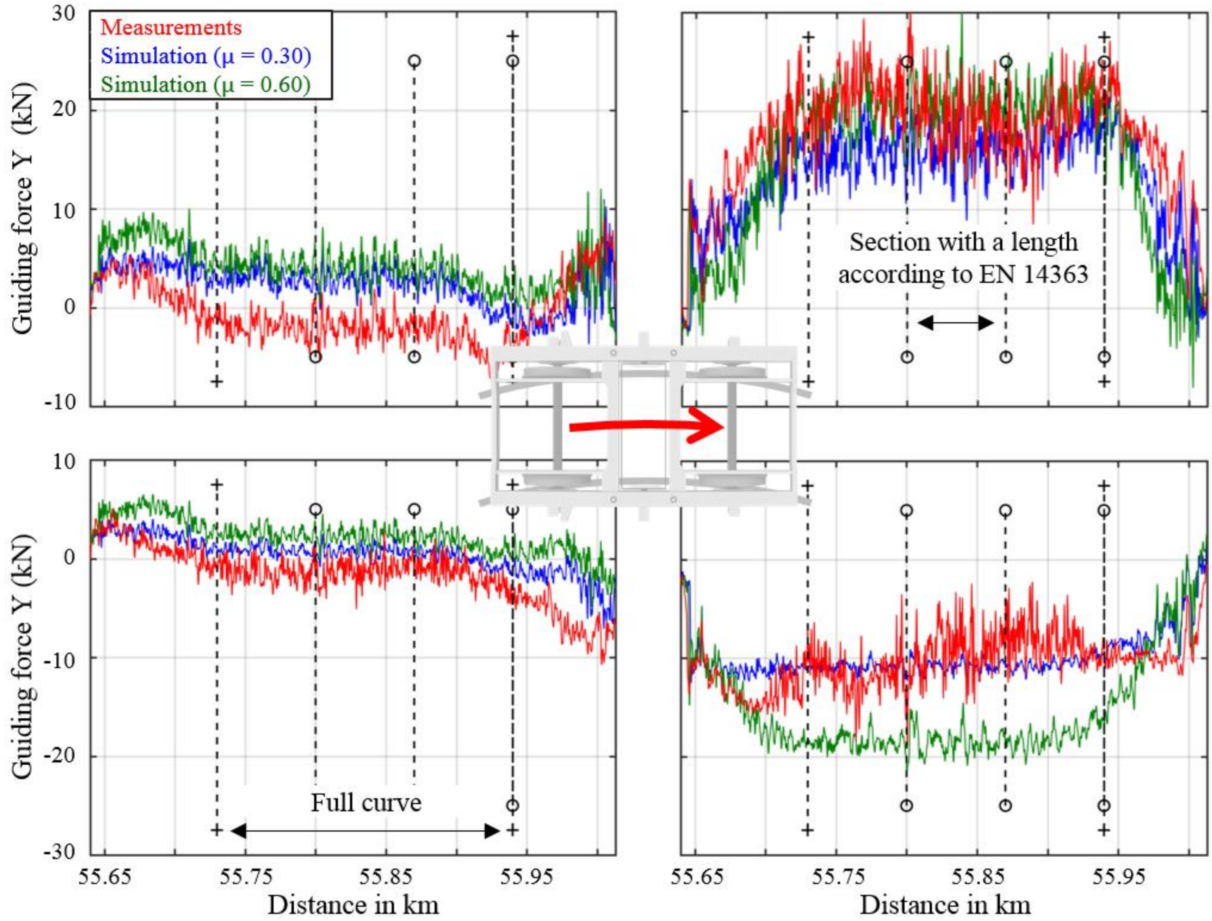

Figure 4. Influence of increasing wheel/rail friction coefficient on Y-forces (a French test section, right curve; the plots relate to the nearest wheel in the central picture of the bogie)

3.5 Validation metrics

The validation metric contains the specification of quantities to be evaluated, the metrics to be applied for comparison of matching between simulation and measurement and the criteria (admissible matching errors) for the validation approval. Moreover, the overall validation result depends not only on quantities and metrics used but also on the set of sections selected for the comparison of matching. The specifications and recommendations stated in standards and papers considering the validation of rail vehicle models often do not provide all those details. Table 2 shows an overview of the validation metrics proposed recently [11, 12, 13, 14] as mentioned in section 3.2. The application domains are very similar; however, the details of these metrics are quite different. The quantities to be evaluated are clearly specified only in the DynoTRAIN method [11].

The metrics are mostly defined as a relative value, whereby the reference (norm) varies: It is either the measured value [13], the higher of simulation and measurement values [12] or the lower of them [14]. One could wonder, why this difference? The relation of matching difference to the measured value considers the measurement as reference, i.e. the value mostly close to the real world. The use of the higher of simulation and measurement values provides an advantageous scaling of the results in the range between 0 and 1, while the use of the lower of simulation and measurement values provides the highest relative value as matching error. The common disadvantages of all relative metrics are high matching errors for small reference values; in asymptote, the matching error will be infinite for reference values approaching zero (expect of the proposal by Bogojevic and Lucanin [12]). The relative metrics thus require high accuracy of modelling unimportant results with values close to zero, which can be hardly considered as necessary. This is why the project DynoTRAIN chose the absolute differences instead of relative differences as the validation metric. The consequence is, that the validation limits need to be defined individually for each quantity and eventually in relation to other parameters as e.g. the static vertical wheel force.

Table 2. Overview of recent metric proposals for the validation of rail vehicle models

Validation criterion (metric)	Application domain	Are the quantities to be evaluated specified?	Type of validation metric	Does the metric apply for a single quantity?	Does the metric apply for a single section?	References
Comprehensive metric by Sprague & Geers	Simulations of rail vehicle on-track testing	No	Relative to measured value	Yes	Yes	Fries et al. [13]
Matching probability of cumulative density functions	Simulation of rail vehicle on-track testing	No	Relative, related to the higher of simulation and measurement values	Yes	Yes	Bogojevic and Lucanin [12]
Cumulative frequency distribution of least-square misfit functions for a set of sections	Various simulations, e.g. assessment of track geometry, investigation of derailments, vehicle gauging	No	Relative, related to the lower of simulation and measurement	Yes	No	Kraft et al. [14]
Mean and standard deviation of differences between simulation and measurement for a set of quantities in a set of test sections	Simulations for the acceptance of running characteristics according to EN 14363	Yes: A set of 12 quantities	Absolute difference between simulation and measurement	No	No	Polach et al. [11]

Two of the compared metrics [12, 13] are applicable to a single quantity from a single section. The metric proposed by Kraft et al. [14] evaluates a set of sections for each quantity, while the metric proposed in DynoTRAIN [10, 11] is only applicable using all specified quantities at least in the specified minimum number of sections. To comply with the validation requirements, all of the investigated quantities have to be below the validation limits.

Although the DynoTRAIN validation method provides specification of quantities to be evaluated, the metrics to be applied for comparison of matching between simulation and measurement as well as the validation criteria, there are still some parameters, which can be selected by the user of this method. One of them is the length of the compared sections, which must be not shorter than the length according to EN 14363, but can be longer. Another open point is the number of compared sections, which must be at least 12, but can be higher. These effects were investigated by the authors in [25] using selected examples.

The overall validation result using any kind of metrics finally depends on the compilation of selected test sections used for comparison with simulation results. This effect, not investigated so far in any previous studies, as well as the impact of signal synchronisation level on the validation result are investigated in the next chapter applying the DynoTRAIN validation method.

4 Investigation of effects on model validation results

4.1 Effect of selected validation sections

The DynoTRAIN project suggested selection of a minimum of three test sections per test zone for the validation of rail vehicle models. However, the set of sections being available because of on-track tests can be much larger and questions about the selection of sections and their influence on model validation results may occur. Compared to previous investigations presented in [25], the authors created for this investigation new compilations of test sections with a maximum number of nine sections per test zone with a length according to EN 14363 (maximum 36 sections in the sum) from the DynoTRAIN project database [23]. This database includes of about 5000 km of measured track data (track geometry and track irregularities). The process to create new sections can be described as follows. Initially, all available sections fulfilling several requirements (e.g. almost constant speed of test train) are created. The numbers of newly created sections varied per test zone and per country. Since the European on-track tests are carried out usually in one country, the authors decided to create country-specific compilations. According to the experiences of the authors, the standard deviation values of the validation quantities are rather crucial whether a model is assessed to be validated. Thus, the authors selected the maximum nine sections per test zone and country out of the maximum available number (per country and test zone) using the requirement that the highest standard deviation value of the validation quantity, which was calculated with all sections, should remain almost the same in the reduced selection of sections. These representative sets of sections are created for two rail vehicles tested in the framework of DynoTRAIN project and modelled using multi-body simulation tool Simpack: Passenger coach Bim 263.5 with MD 36 bogies and loaded freight wagon Sgns 691 with Y25 bogies. All possible validation results are then calculated by applying the DynoTRAIN methodology to all possible combinations of test section compilations using the requirement to select three sections per test zone. This leads to a large variation of validation results as shown in Figures 5 and 6. A boxplot format is selected to visualize the scatter in the calculated mean and standard deviation values. In the box, 50% of the validation results per validation quantity are included. The horizontal red line describes the median value while the upper and lower horizontal blue line describes the 0.25- and the 0.75-quartil.

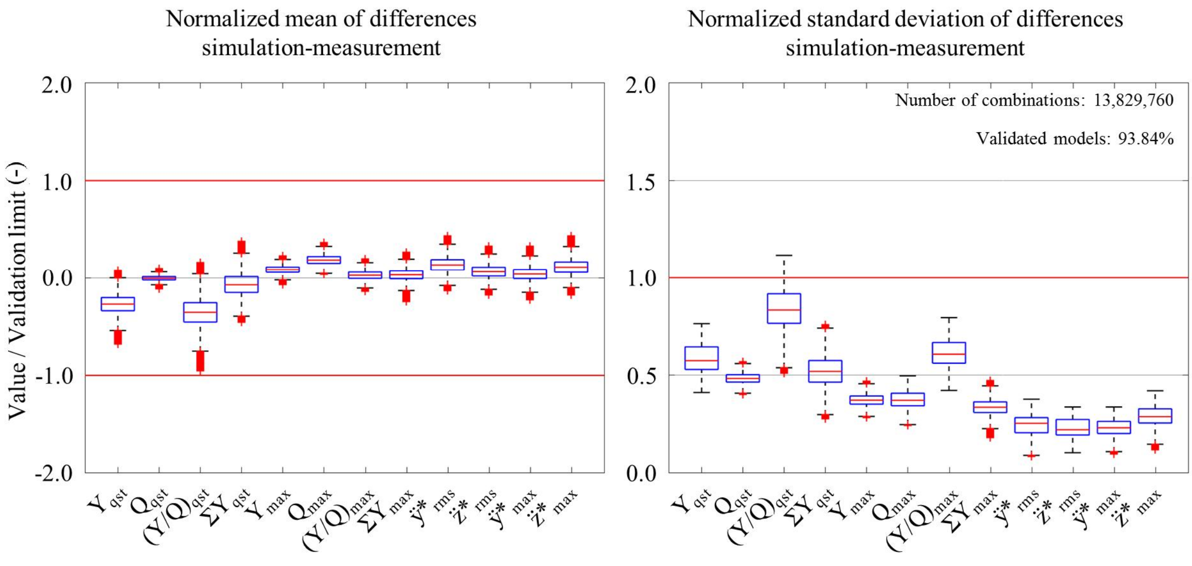

Figure 5. Influence of section selection on model validation results (passenger coach)

The red crosses represent the outliers. Figure 5 shows the scatter of validation results of the passenger coach (see Figure 2 for a single validation result). The section compilation influences rather negligible the overall model validation result. In 93.84% of all section compilations, the model is assessed to be validated. In comparison, the section selection for the loaded freight wagon (Figure 6) has no influence on the overall model validation result. In every section compilation, the model is assessed to be not validated. Both examples demonstrate that in spite of the highest effort in varying selected sections for model validation, the overall assessment of the model is not affected. Further investigations of the presented effects are described in [26].

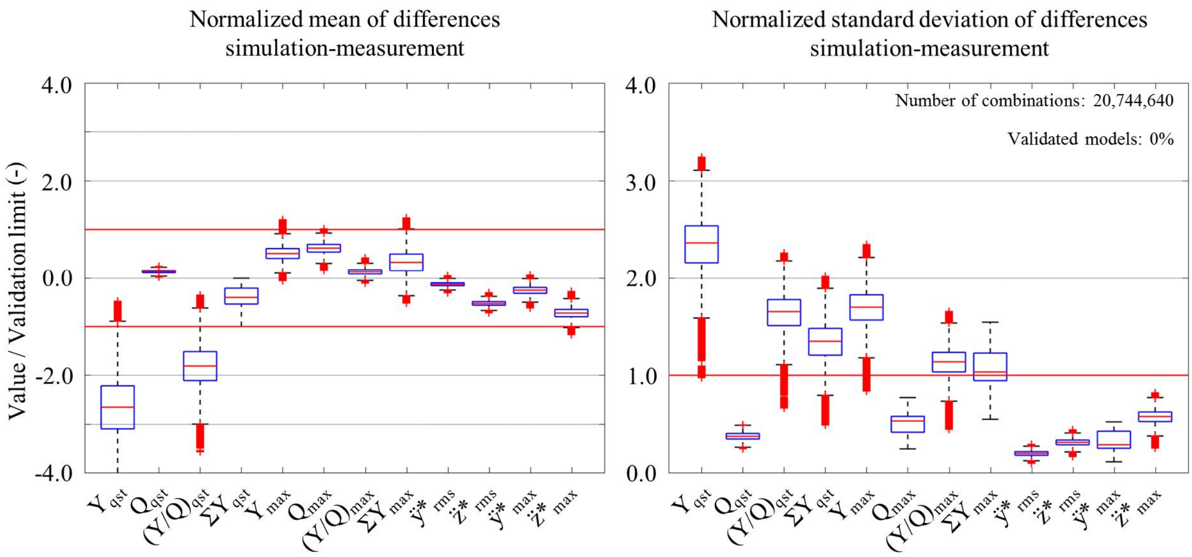

Figure 6. Influence of section selection on model validation results (loaded freight wagon)

4.2 Effect of signal synchronisation

The model validation based on comparisons of simulation and measurement requires a good synchronisation of the compared signals. The sensitivity to the exactness of synchronisation, however, can differ dependent on the definition of the validation metric. The DynoTRAIN validation metric was developed with the aim of allowing reliable validation also without a perfect signals' synchronisation. To achieve this aim, this validation is based on quasi-static values, rms and maximum values of the compared measurement and simulation sections instead of direct comparison of time plots.

The effect of the synchronisation of compared signals is evaluated using the example of a randomly created compilation of 12 test sections from the DynoTRAIN test campaign (Italian test sections with a length according to EN 14363; three sections per test zone). The DynoTRAIN methodology [1, 11] is applied to calculate the validation result. The differences between simulation and measurement are calculated for each section, value and quantity. The mean value and the standard deviation of differences between simulation and measurement are then calculated for each quantity as described in Section 3.2 and compared with the validation limits defined in Table 1. Figure 7 shows for each quantity the mean and the standard deviation of differences between simulation and measurement normalised by the correspondent validation limits. All signals from simulation and measurement in each section are initially synchronized with a very high accuracy. Successively, the synchronisation level is worsened with an offset between both signals of 1 m, 10 m and 20 m. The presented example shows very small effect of synchronisation error on the validation result. However, is this a typical result?

By applying the method to investigate the influence of the selection of sections on model validation result as introduced in Section 4.1, the signal synchronisation is further investigated using the example of the passenger coach. Figure 5 shows the scatter of the validation results due to selection of test sections when the measured and simulated signals are perfectly synchronized. The model is assessed to be validated in 93.84% of the combinations of test sections. Successively, the accuracy of the signal synchronisation is worsened. Figure 8 shows the validation results when the simulated and measured signals are not perfectly synchronized but possess an offset of 40 m. The model is assessed to be validated in 91.17% of the combinations of test sections, which is nearby the same as when the signals are perfectly synchronized. Table 3 shows the percentage values of validated models for the offsets between simulated and measured signals varying from zero to 40 m. These results confirm the achievement of the intended approach of the DynoTRAIN project to develop a reliable validation methodology which is rather insensitive to the synchronisation level of the measurement and simulation signals.

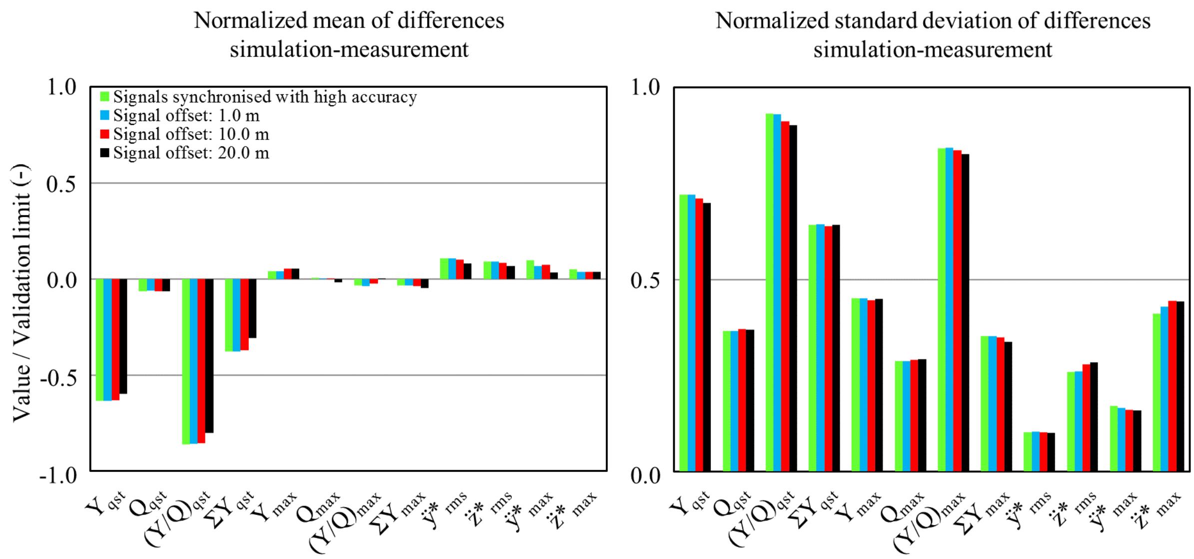

Figure 7. Influence of signal synchronisation on model validation on example of one random compilation of test sections (passenger coach)

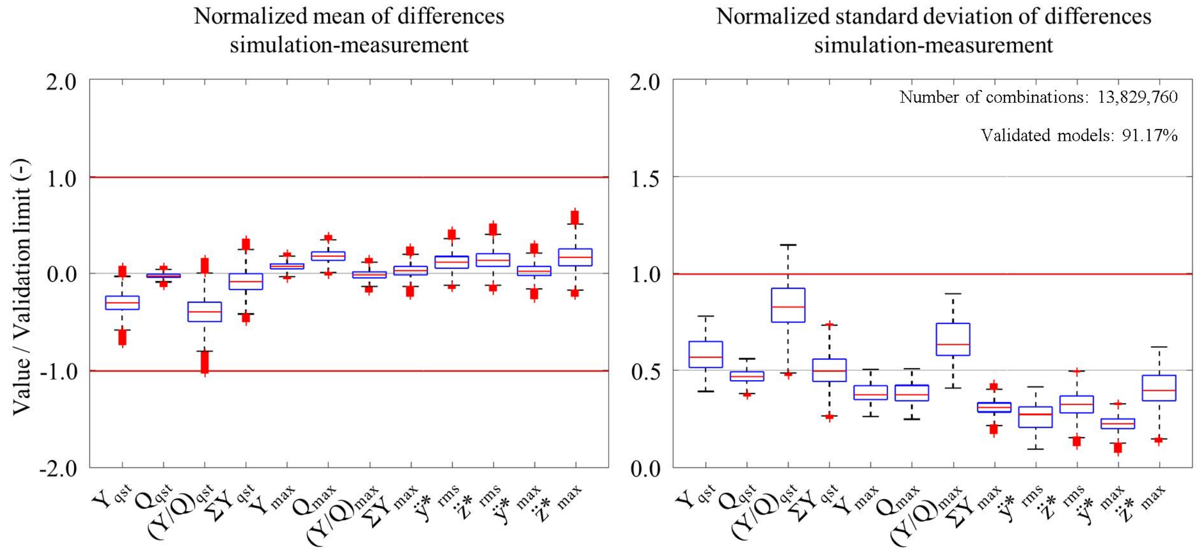

Figure 8. Influence of section selection on model validation results with an offset of 40 m between simulation and measurement signals (passenger coach)

Table 3. Influence of signal synchronisation on the percentage of validated models (passenger coach; 13,829,760 combinations of test section compilation)

Synchronization error	Synchronized signals	± 5 m	±10 m	±20 m	±40 m
Percentage of validated models	93.84%	93.69%	92.42%	93.62%	91.17%

5. Summary and conclusions

This article summarises the latest research on methodologies for validation of simulation models in the context of rail vehicle certification. First, the topics considering peculiarities of on-track testing of rail vehicles, validation domain, validation quantities and validation metrics are introduced. This is followed by a comparison of four recently proposed approaches for validation of rail vehicle simulation models and complemented by new research results regarding one of them.
The validation method developed in project DynoTRAIN provides an entire specification of quantities to be evaluated, validation metrics, required number of comparisons as well criteria to be applied for the validation approval. Another three reviewed validation approaches provide metrics which are applicable to single quantity from a single section, or a single quantity from a set of sections. As these methods do not specify the numbers of quantities and sections to evaluate and the matching errors to fulfil for a successful validation, the assessment is rather a subjective decision of the expert conducting this validation.
The presented new research evaluates the robustness of the DynoTRAIN validation methodology. This method is rather insensitive against the accuracy of the signal synchronisation because it is based on quasi-static, rms and maximum values inside of the compared measurement and simulation sections. The investigations considering the influence of section selection on model validation result confirm the robustness of this method. As the overall DynoTRAIN assessment depends on a large set of comparisons between simulation and measurement values of different quantities, the overall validation result remains either “validated” or “not validated”, rather independently of the set of test sections selected for validation.

Acknowledgement

This investigation used data from the research project DynoTRAIN (2009-2013) supported by the European 7th Framework Programme, contract number: 234079. The research presented in this paper received no specific grant from any funding agency.

References

[1] EN 14363:2016 Railway applications — Testing and simulation for the acceptance of running characteristics of railway vehicles — Running Behaviour and stationary tests. CEN, Brussels, March 2016.
[2] ASME V&V 10-2006. Guide for verification and validation in computational solid mechanics, New York, 2006.
[3] Oberkampf WL, Trucano TG, Hirsch C Verification, Validation, and Predictive Capability in Computational Engineering and Physics. Appl. Mech. Rev 2004, 57(5): 345-384.
[4] Polach O, Berg, M, Iwnicki SD Chapter 12: Simulation. In: Iwnicki S., Editor: Handbook of railway vehicle dynamics. CRC Taylor&Francis, Boca Raton London New York 2006.
[5] UIC Code 518:2009. Testing and approval of railway vehicles from the point of view of their dynamic behaviour – Safety – Track fatigue – Running behaviour. International Union of Railways, September 2009.
[6] GM/RT2141, Issue 3, Resistance of Railway Vehicles to Derailment and Roll-Over. RSSB, London, June 2009.
[7] 49 CFR 213, FRA Regulations, Electronic code of federal regulations, Title 49 Transportation, Part 213—Track safety standards. https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=1&SID=1ad0d8d4e99f4915f5c2a81cd4f82c6e&ty=HTML&h=L&r=PART&n=49y4.1.1.1.8
[8] EN 15273-2:2013 Railway applications - Gauges - Part 2: Rolling stock gauge. CEN, Brussels, May 2013.
[9] EN 14067-6:2010. Railway applications - Aerodynamics - Part 6: Requirements and test proce-

dures for cross wind assessment. CEN, Brussels, January 2010.
[10] Polach O, Böttcher A A new approach to define criteria for rail vehicle model validation. Vehicle System Dynamics 2014; 52:sup1: 125-141.
[11] Polach O, Böttcher A, Vannucci D, Sima J, Schelle H, Chollet H, Götz G, Garcia Prada M, Nicklisch D, Mazzola L, Berg M, Osman M Validation of simulation models in context of railway vehicle acceptance. Proc IMechE Part F: J Rail and Rapid Transit 2015; 229(6): 729–754.
[12] Bogojevic N, Lucanin V The proposal of validation metrics for the assessment of the quality of simulations of the dynamic behaviour of railway vehicles. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, February 2016 vol. 230 no. 2 pp. 585-597.
[13] Fries R, Walker R, Wilson N Validation of dynamic rail vehicle models. In: 23rd International Symposium on Dynamics of Vehicles on Roads and Tracks, August 19-23, 2013, Qingdao, China. Paper No. 10.1, ID416.
[14] Kraft S, Causse J, Coudert F An approach for the validation of railway vehicle models based on on-track measurements. Vehicle System Dynamics, 2015, Vol. 53, No. 10, pp. 1480–1499.
[15] Polach O, Evans J Simulations of running dynamics for vehicle acceptance: Application and validation. International Journal of Railway Technology 2013; 2: 59-84.
[16] ERRI B 176/3 Benchmark problem – Results and assessment, B176/DT290, Utrecht 1993.
[17] ORE B 176 Bogies with steered or steering wheelsets. Report No. 1: Specifications and preliminary studies. Vol. 2, Specification for a bogie with improved curving characteristics. ORE, Utrecht 1989.
[18] Iwnicki SD Manchester Benchmarks for Rail Vehicle Simulation. Vehicle System Dynamics 30(1998): no. 3-4, S. 295-313.
[19] Iwnicki, SD The results of the Manchester Benchmarks, Vehicle System Dynamics Supplement, 31(1999): 2-12.
[20] Schwer LE An overview of the end to end example: An illustration of the concepts of verification and validation. In: Proceedings of PACAM XI, 11th Pan-American Congress of Applied Mechanics, January 04-08, 2010, Foz do Iguaçu, PR, Brazil.
[21] Sprague MA, Geers TL A spectral-element method for modelling cavitation in transient fluid-structure interaction. International Journal for Numerical Methods in Engineering 2004; 60: 2467-2499.
[22] Schwer LE Validation metrics for response histories: perspectives and case studies. Engineering with Computers 2007; 23: 295–309.
[23] Zacher M, Kratochwille R Stationary and on-track tests using various vehicles. Proc IMechE Part F: J Rail and Rapid Transit 2015; 229(6): 668-690.
[24] Marquis BP, Tajaddini A, Kotz H-P, Breuer W, Trosino M Application of new FRA vehicle–track system qualification requirement. The Stephenson Conference - Research for Railways, London, 21-23 April, 2015, pp. 581-592.
[25] Götz G, Polach O Influence of varying input parameters on model validation results. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2017, 231(5) "Stephenson Conference Special Issue", pp. 598–609.
[26] Götz G, Polach O Model validation – A random or a reliable result? Paper ID 154, 25th International Symposium on Dynamics of Vehicles on Roads and Tracks, Rockhampton, Australia, August 14-18, 2017.