



Validation of simulation models in context of railway vehicle acceptance

Oldrich Polach, Andreas Bottcher, Dario Vannuci, Jurgen Sima, Hennig Schelle, Hugues Chollet, Mayi Garcia Prada, G. Götz, Dirck Nicklisch, Laura Mazzola, et al.

► To cite this version:

Oldrich Polach, Andreas Bottcher, Dario Vannuci, Jurgen Sima, Hennig Schelle, et al.. Validation of simulation models in context of railway vehicle acceptance. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2015, 229 (6), 10.1177/0954409714554275 . hal-01512721

HAL Id: hal-01512721

<https://hal.science/hal-01512721>

Submitted on 24 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Validation of simulation models in context of railway vehicle acceptance

O. Polach¹, A. Böttcher², D. Vannucci³, J. Sima⁴, H. Schelle⁵, H. Chollet⁶, G. Götz⁷,
M. Garcia Prada⁸, D. Nicklisch⁹, L. Mazzola¹⁰, M. Berg¹¹, M. Osman¹²

¹)Bombardier Transportation (Switzerland) AG, Switzerland, ²)Alstom Transport Deutschland GmbH, Germany, ³)AnsaldoBreda, Italy, ⁴)Siemens AG, Germany, ⁵)TU Berlin, Germany (now at Bombardier Transportation GmbH, Germany), ⁶)IFSTTAR, France, ⁷)Bombardier Transportation GmbH, Germany (now at TU Berlin, Germany), ⁸)CAF, Spain, ⁹)DB Netz AG, Germany, ¹⁰) Politecnico di Milano, Italy, ¹¹)KTH, Sweden, ¹²)RSSB, UK

Abstract

Evaluation of a reliable validation method, criteria and limit values suitable for model validation in the context of vehicle acceptance was one of the objectives of the DynoTRAIN project. The presented investigations represent a unique amount of testing, simulations, comparisons with measurements and validation evaluations. The on-track measurements included several different vehicles tested in four European countries in a test train equipped with a simultaneous recording of track irregularities and rail profiles. The simulations were performed using vehicle models built with the use of different simulation tools by different partners. The comparisons between simulation and measurement were conducted in over 1,000 simulations using a set of the same test sections for all vehicle models. The results were assessed by three different validation approaches: Comparing values according to EN 14363, by subjective engineering judgement by project partners and using so called validation metrics, i.e. computable measures developed with the aim to increase the objectivity while still maintaining the agreement with engineering judgement. The proposed validation method uses the values computed by analogy with EN 14363 and provides validation limits to be applied on a set of deviations between simulation and measurement.

Keywords: Validation, railway vehicle, vehicle model, simulation, running dynamics, running characteristics, vehicle acceptance, authorisation, certification, homologation

1. Introduction

Railway vehicle acceptance is one of the significant cost and time drivers during the acquisition of railway rolling stock. Multi-body simulation tools, which are widely used in rolling stock design and development to conduct a wide range of investigations including the prediction of test results, can contribute to reduce the time and cost of the testing for the acceptance of running characteristics. Meanwhile, the reliability of simulations is becoming widely recognised and the opportunity to replace some physical tests by computer simulations has recently been considered in standards and product specifications. However, a reliable validation of the simulation model is the crucial condition when considering the application of simulations in the vehicle acceptance context.

The validation of the computer simulation model is a process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model [1]. In contrast to the verification which is primarily dedicated to the checking of the multi-body simulation code and conducted by the code developers, the model validation has to be carried out by the model developer and considers the particular model stage and the particular intended application of the model. The validation represents the comparisons with measurements assessing the quantitative accuracy of the simulation model in regard to the intended application, i.e. the simulations using the validated model. Simply said, the validation should check if the model is suitable for the intended simulations, if it is “fit for purpose”.

The comparison with measurements used for model validation should take into account all errors and scatter of conditions influencing both measurement as well as simulation: the errors of running dynamics measurement, the errors in the measurement of track layout and track irregularities, measurement of rail profiles and wheel profiles, as well as the scatter of the test conditions e.g. friction coefficient between wheel and rail. The validation assessment should also take into account the number of repeated tests used for validation and their reproducibility.

The surveys dedicated to validation by Cooperrider and Law [2] and by Gostling and Cooperrider [3], both from the advent of modern computer simulation techniques, represent the verification of the simulation tools / software, rather than the model validation. Meanwhile, computer simulations are widely used in the design of railway rolling stock and in research studies, but the progress of validation methodologies is rather small. A number of publications present particular comparisons between simulation and measurement documenting the validation of the particular simulation model, e.g. a validation of tramcar vehicle model in [4], validation of vehicle critical speed when negotiating a large radius curve in [5] or validation of tilting train in [6]. However, no systematic investigations are presented regarding the validation methodology considering simulations of railway vehicles. The state of the art papers by Evans and Berg [7] from 2009 as well as by Bruni et al. from 2011 [8] provide some hints regarding the validation of multi-body railway vehicle models.

Experience with the validation of railway vehicle models in the context of the vehicle acceptance process has been gained over many years in the UK and introduced in the Railway Group Standard Guidance Note GM/RC2641[9]. A vehicle model validated against the stationary tests according to GM/RC2641 can be used in the UK for the assessment of the resistance of railway vehicles to derailment according to the Railway Group Standard GM/RT2141 [10]. This model validation method has also been incorporated as recommended practice in to the European standard EN 15273-2:2013 [11] dealing with vehicle gauging.

The validation experience gained by dynamics specialists in the UK has been used during the preparation of the model validation process described in UIC 518:2009 [12].

Furthermore, two model validation trials were conducted by this committee. The experience with one of them dealing with the simulations of a locomotive acceptance tests is published in Ref. [14]. The results of the second validation trial related to a freight wagon with Y25 bogies were presented and discussed in the framework of the DynoTRAIN project.

The recent revision of prEN 14363:2012 [13] includes the possibility to use computer simulations under certain conditions as specified in this standard. The requirements for the model validation in this document have been evaluated using the experience in the UK as

well as during the preparation of UIC 518:2009. Unfortunately, neither document contains a specification of the allowable differences between simulation and on-track test results. Because of the lack of quantitative criteria, an assessment by an independent reviewer is required to ensure that the model provides a sufficient representation of reality for the intended application. To be able to remove this requirement was one of the main objectives for work package 5 (WP5) of the DynoTRAIN project.

Clear, quantitative and measurable criteria and limit values to assess the differences between simulation and measurement (called also matching error limits) in the model validation process represent a crucial requirement when applying simulations to reduce the amount of physical testing during the vehicle acceptance process. Such quantitative limits would enable the specialist carrying out simulations to understand if a particular model fulfils the validation requirements or if it needs an improvement, to visualise the model weaknesses and to motivate the specialists to improve their model if needed. Unambiguous quantitative validation criteria and limits will ensure that all vehicle models used in the vehicle acceptance context have achieved the sufficient quality level.

The objectives of DynoTRAIN WP5 were:

- To review the state of the art of building and validation of multi-body railway vehicle models
- To test vehicle models by comparisons between simulations and measurements
- To specify the requirements for validation of vehicle models in the context of vehicle acceptance.

The DynoTRAIN WP5 investigations were structured in to 5 Tasks. The investigations started with Task 1 dedicated to state of the art of vehicle modelling and validation. The review of the state of the art of suspension and vehicle modelling was summarised in a paper presented during the IAVSD Symposium in Manchester 2011 [8]. Questionnaires and presentations about model validation experience showed that the validation is typically carried out as a synthesis of stationary tests and on-track measurements, sometimes combined with validation of suspension components modelling. Measured track irregularities and rail profiles from along the test route during the on-track tests are very often not available. This missing data are usually mentioned as the reason for the presented deviations between simulations and measurements. Task 2 was dedicated to investigations about suspension modelling. It provided a variety of comparisons and allowed improved insight in to the modelling of suspension components (rubber components, suspension with friction, viscous dampers, and air springs); see the presentations of some of the results in [15] and [16]. The experience gained in Tasks 1 and 2 was used when modelling the vehicles used in the validation investigations. The preparation of vehicle models and the identification of uncertain or unknown parameters by comparisons with stationary tests was the topic of Task 3. Tasks 4 and 5 were dedicated to validation studies and analyses, which resulted in the proposed new approach for the validation of vehicle models with regard to the vehicle acceptance described in this article. The presented validation studies represent a unique work regarding the validation of railway vehicle models in the context of vehicle acceptance. The measurements with a test train with several different vehicle types conducted in four European countries and equipped with a simultaneous recording of track irregularities and rail profiles [17] were compared with a large set of simulations. The validation evaluations carried out in the framework of the presented investigations were performed using several vehicle models, built by 7 project partners using 3 different simulation tools. The proposed process, the criteria and the validation limits are based on a large investigation using the state of the art of modelling and simulation.

This article is structured as follows. The next chapter presents the tests used for evaluation, simulation models and model configurations with differing input parameters, selection of simulation input parameters and test sections selected for comparisons of simulations and measurements. Chapter 3 describes the comparisons investigated in regard to defining the model validation approach. Chapter 4 presents the evaluations related to the selection of a suitable validation method and validation limits (matching errors). Chapter 5 presents the proposed method, criteria and limit for validation of vehicle models used for simulations of on-track tests in the context of vehicle acceptance. Chapter 6 is dedicated to a discussion about the proposed validation method and about the influence of the model adjustments by comparisons with stationary tests. A summary and conclusion is provided in chapter 7.

2. Validation investigations in DynoTRAIN

2.1 On-track tests used for validation

The presented model validation investigations used on-track measurements conducted in the framework of DynoTRAIN WP1 as well as some measurement results provided by project partners.

The DynoTRAIN test campaign was conducted in October 2010 with several different vehicles which were equipped with 10 force measuring wheelsets and several other sensors [17]. The train travelled for a total of 20 days of test runs through Germany, France, Italy and Switzerland with speeds up to 120 km/h with freight wagons connected and up to 200 km/h without freight wagons. A measuring vehicle integrated in the test train continuously recorded the track irregularities and rail profile shapes along all test runs. The test train contained the following vehicles:

- Locomotive DB BR 120
- DB passenger coach Bim
- Empty freight wagons Sgns with Y25 bogies
- Loaded freight wagons Sgns with Y25 bogies
- Laas freight vehicle unit consisting of two 2-axle wagons with UIC link suspension; one empty and one fully loaded; whereby the empty wagon was equipped with measuring wheelsets.

Besides the vehicles tested in DynoTRAIN, another two vehicles were investigated:

- High speed EMU for TCDD (Turkey) manufactured by CAF
- DMU IC4 for DSB (Denmark) manufactured by AnsaldoBreda.

The measurements used for the validation investigations with these vehicles were done in the past during the running dynamic acceptance tests of these vehicles under contract with the vehicle supplier.

2.2 Vehicle models and model configurations

Multi-body vehicle models used for the evaluation of the validation method were prepared by project partners using different simulation tools, see examples of models built in simulation tool Simpack in Figure 1. Several versions of each vehicle model were prepared using different stages of model parameters, track irregularities, rail and wheel profiles as well as modelling depth. The differing model versions are called “model configurations” here. An overview about the vehicle models used in the presented investigations is shown in Table 1.

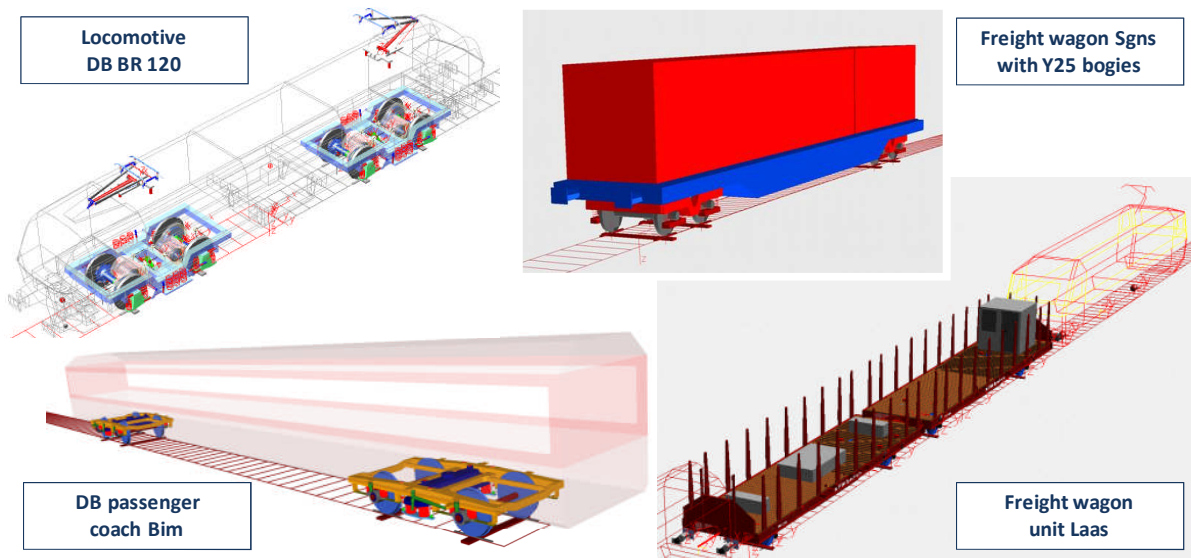


Figure 1. Examples of multi-body vehicle models of vehicles tested in DynoTRAIN.

Table 1 Overview about the multi-body simulation models used for the evaluation of the presented validation methodology.

Vehicle	Project partner	Simulation tool	Number of model configurations	On-track tests used for validation
Locomotive DB BR 120	Siemens	Simpack	24	DynoTRAIN
	IFSTTAR	VOCO	4	DynoTRAIN
DB passenger coach Bim	Bombardier Transportation	Simpack	13	DynoTRAIN
	IFSTTAR	VOCO	4	DynoTRAIN
Freight wagon Sgns, empty	Technical University Berlin	Simpack	8	DynoTRAIN
	IFSTTAR	VOCO	6	DynoTRAIN
Freight wagon Sgns, laden	Technical University Berlin	Simpack	7	DynoTRAIN
Laas freight vehicle, empty	Alstom	Simpack	5	DynoTRAIN
High speed EMU (TCDD)	CAF	SIDIVE	3	Provided by vehicle supplier CAF
DMU IC4, coach T3 (DSB)	Ansaldobreda	Simpack	2	Provided by vehicle supplier Ansaldobreda
DMU IC4, coach M4C (DSB)	Ansaldobreda	Simpack	2	Provided by vehicle supplier Ansaldobreda

The vehicle models used in the investigations represent fully non-linear three dimensional models as this is the state of the art in railway engineering and research. Rigid bodies representing vehicle body, bogie frame, wheelset, axle box etc. are connected by springs, dampers, friction elements and bump-stops modelling the suspension components. Damper models consist of a dashpot together with series stiffness. The non-linear wheel/rail contact

models use the respective contact evaluation method and the respective version of Kalker's computer code Fastsim implemented in the applied simulation tool.

The vehicle models were prepared under the partners' responsibility. The basic data regarding the vehicles tested during the DynoTRAIN test campaign was provided by DB; the remaining information was estimated or identified from archive material by partners modelling the vehicles. The identification of vehicle model parameters of vehicles tested outside the DynoTRAIN project was fully under the responsibility of the respective partner, which is also the supplier of the vehicle and provided the data from the running tests.

The initial vehicle models were prepared using the available vehicle data without considering the results of stationary tests. Project partners were however advised to adjust the model mass parameters before starting the comparisons to achieve a good agreement between the static model wheel loads and the static wheel loads measured during the on-track tests. Then, the initial models were adjusted with the aim to improve the agreement of the on-track tests with the simulation results, so that several differing configurations of the same model could be compared. The vehicle models adjusted based on the comparisons with the stationary tests represented other model configurations. In order to assess the effect of using the actual measured infrastructure parameters like track layout, track irregularities and rail profiles, the model configurations with estimated rail profiles and estimated track irregularities data were also prepared and compared with the on-track measurements.

A total of 78 model configurations were investigated, with differing knowledge of vehicle data, input parameters regarding the infrastructure, different usage of stationary tests (before and after model adjustments by comparisons with stationary tests) and applying a different depth of modelling detail. Moreover, some model configurations of the locomotive BR 120 by Siemens varied implementing the driving torque in test sections where this locomotive was used as a propelling vehicle. In order to assess the effect of using the actual measured infrastructure parameters like track layout, track irregularities and rail profiles, several model configurations were compared. Besides the model configuration applying measured input data, the configurations with estimated rail profiles and estimated track irregularities data were investigated. For the sake of brevity, readers interested in the effect of measured and estimated wheel and rail profiles as well as track irregularities data on the model validation results are referred to [18]. The effect of using the results of stationary tests for the model validation in regard to the simulation of the on-track tests was investigated by comparing the simulations of the on-track tests using vehicle models before and after the comparisons with the stationary tests, see discussion in chapter 6.2. Figure 2 shows the variety of investigated model configurations together with the assessed quantities which are described in more detail in chapter 3.

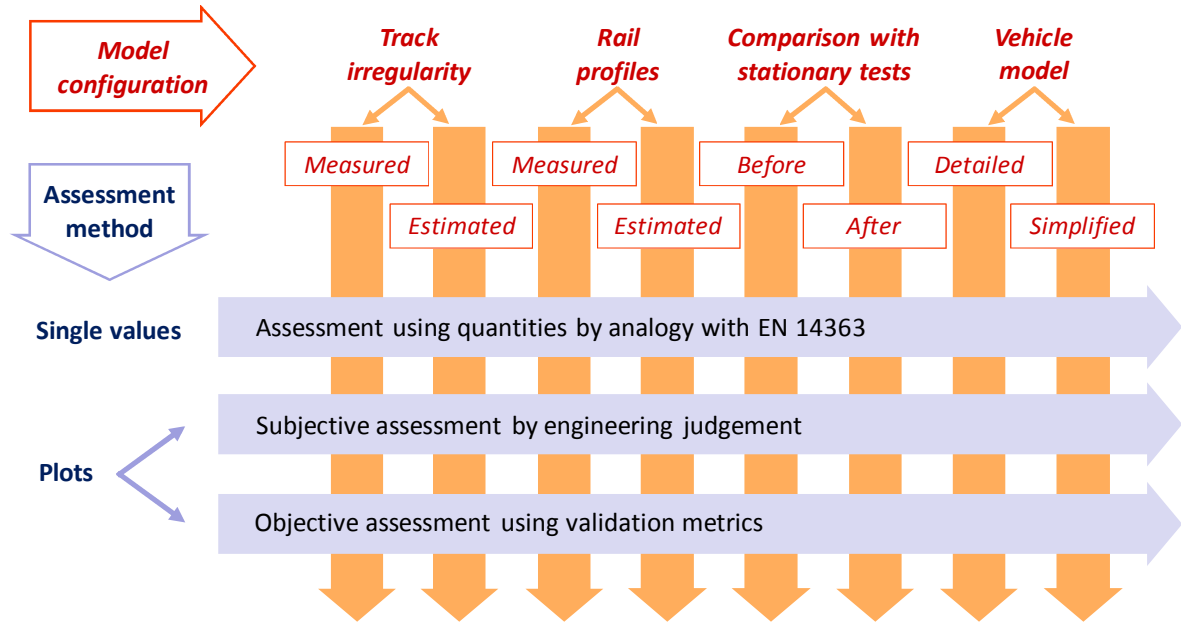


Figure 2. Overview of model configurations and assessment methods evaluated in the framework of the presented investigations.

2.3 Simulation input parameters

2.3.1 Track layout and track irregularities

The simulations of vehicles tested in DynoTRAIN used measured track layout as well as measured track irregularities. Because the track layout (curvature) as well as track irregularities are usually measured simultaneously, a correct separation and introduction of these data in the simulations represents one of the difficulties considering model validation.

The track geometry data were measured during the DynoTRAIN test campaign by the DB track recording car “RAILab I” [17]. The data were stored in binary files with a sampling distance of 0.16 m.

The preparation of measured track irregularities in a format suitable for simulations has been carried out by DB Netz AG. As the inertial platform based RAILab I system uses a special filter algorithm to separate long wave lengths caused by the track layout from the track irregularities to be assessed, the recorded data were de-coloured using corrective filters before using them in the vehicle dynamics simulations. For each of the selected track sections the relevant RAILab I data were transformed in to the format used in multi-body simulation package Simpack. Two input data files were created for each track section; one of them containing the track layout (curvature and cant using high-pass filters 70 m) and the second describing irregularities (lateral and vertical position of the left and right rail with band-pass filters 1-70 m).

There were no measurements of track irregularities available for the on-track measurements conducted outside of DynoTRAIN. Thus, the simulations with vehicle models DMU IC4 and High Speed EMU Turkey were carried out using estimated track irregularities. This estimated track irregularity data used either in case of missing measured data or for comparisons regarding the importance of the track data knowledge were either generated based on the power spectral density according to ORE B176 [19] or measured

track irregularities from other measurements. The selection of track irregularities to be used instead of the actual measured data was the responsibility of the respective partner.

2.3.2 Rail profiles

The rail profiles were measured during the DynoTRAIN test runs by means of an optical measuring device [17] and recorded with a spacing interval of 0.25 m. For the synchronisation of the measured rail profiles with all the other measuring channels the time stamp and counter signal provided by the track recording car RAILab I was combined with the odometer signal of the rail profile measuring system and both stored together in an additional synchronisation file.

The implementation of measured rail profiles in multi-body simulations generates several questions. A typical recommendation is to use a “representative profile”. However, how do you identify this representative profile? As the rail profiles in curves are wearing differently on the outer and inner rail as well as differently from straight and curve transitions, the use of one profile for each rail along the whole investigated section will obviously provide incorrect results either outside the full curve or in the full curve, unless there is no wear of rails and the rail profiles are thus identical over the whole track section. Continuously varying rail profiles along the track section has been implemented in some of the simulation packages but it is still not a state of the art and thus not applied here. After several investigations and discussions regarding this topic, it was finally agreed to calculate averaged rail profiles from the measured rail profiles of the part of the actual track section with constant track curvature (i.e. one profile for left and one for right rail) and to use these averaged profiles for simulations of this particular track section. Thus, the profile used may be incorrect in curve transitions and accompanied straight track parts. Moreover, if the actual rail profile changes along the distance, e.g. in some longer sections, the applied averaged rail profiles may not be fully representative.

The preparation of the averaged rail profiles was conducted by DB Netz AG. At first the profiles were smoothed and their running surfaces (down to an appropriate profile gradient) were approximated by high-order polynomials. Then all profiles of the same rail within the respective track section were aligned to each other vertically at the rail top and laterally at the gauge measuring point (14 mm below the top of rail). In order to allow for superposition of measured track irregularities, finally the resulting rail profiles were shifted in the lateral direction to meet the 1435 mm nominal track gauge.

For each simulation exercise a mean profile for left rail and a mean profile for right rail were provided by taking into account all rail profiles of the track section with constant radius, i.e. section C-D in Figure 3. These single mean profiles for left and right rail were used in simulations of the complete particular section.

The model configurations with “estimated rail profiles” used design rail profile and rail inclination of the particular country. There were no measurements of rail profiles available for the on-track measurements conducted outside of DynoTRAIN. Thus, the simulations of vehicle models DMU IC4 and High Speed EMU Turkey used both the respective design rail profiles and rail inclinations of the particular country.

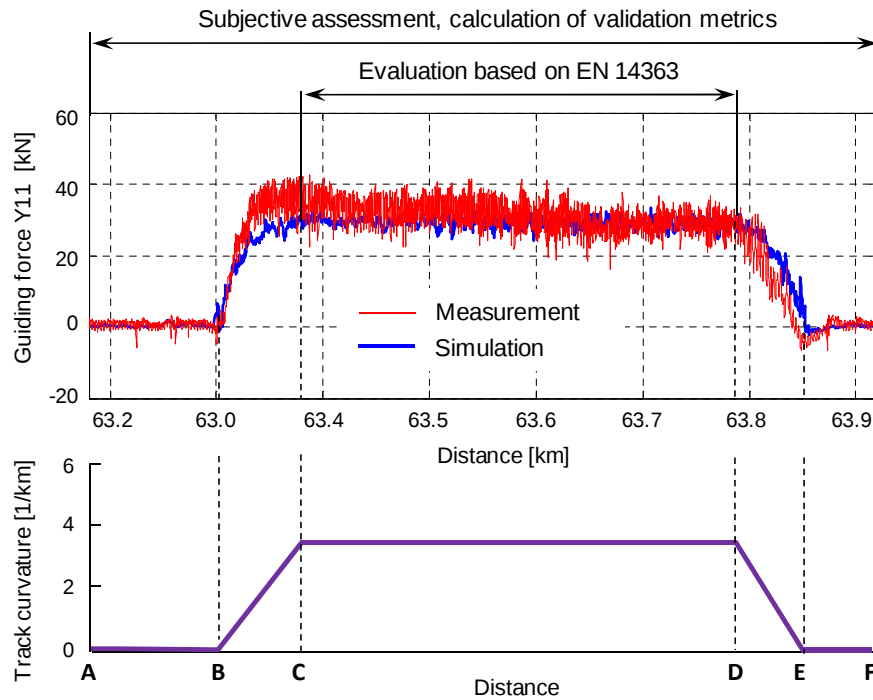


Figure 3. Example of validation exercise with the specification of track sections used for different kind of assessments.

2.3.3 Wheel profiles

The wheel profiles of vehicles tested in DynoTRAIN were measured before and after the test campaign and the measured data were used for simulations. The details regarding the wheel profile implementation were the partner's responsibility. The model configurations with "estimated wheel profiles" were carried out using the design wheel profile S1002. There were no measurements of the wheel profiles available for the on-track measurements conducted outside of DynoTRAIN. Hence, the vehicle models DMU IC4 and High Speed EMU Turkey used respective design wheel profiles: Profile S1002 (DMU IC4) or profile GV 1/40 (High Speed EMU Turkey), respectively.

2.3.4 Friction coefficient between wheel and rail

The value of friction coefficient between wheel and rail represents an uncertain simulation input parameter. The selection of this parameter was the responsibility of the partner carrying out the simulation. All test runs selected for validation from the DynoTRAIN measurements were carried out on dry rail. The partners used wheel/rail friction coefficients in the range 0.45 – 0.60. The majority of simulations used a constant value of friction coefficient identical in the tread and on the flange; only a few simulations used a lower friction coefficient on the flange. The majority of simulations were carried out using friction coefficients of 0.45 or 0.5; a few model configurations used a lower friction coefficient than 0.45 with the aim to test for an improvement of the agreement with the measurements.

The simulations of test results which were provided by vehicle suppliers used a value of wheel/rail friction coefficient of 0.45 (simulations of DMU IC4 by AnsaldoBreda) or a friction coefficient of 0.35 (simulations of High speed EMU Turkey by CAF), respectively.

2.4 Validation exercises

The comparisons between simulation and measurement were carried out for all vehicle models and model configurations under the same conditions and in the same manner as for selected representative sections of test runs, called validation exercises. One validation exercise consists of one curve passing scenario including both transitions and parts of straight track as shown in Figure 3. In this context the word “section” means a part of track; it does not mean section according to the definition in EN 14363:2005 [20]. A total of 17 validation exercises were selected representing all four track zones according to EN 14363:2005: Straight track and very large radius curves were represented by 5 sections, large radius curves ($R > 600$ m) by 2 sections; four sections were from small radius curves ($400 \text{ m} \leq R \leq 600 \text{ m}$) and 6 from very small radius curves ($250 \text{ m} \leq R < 400 \text{ m}$). Table 2 shows the parameters of the test sections selected for vehicles tested in DynoTRAIN regarding the location, track layout, section length as well as the speed of the test train in the respective section. It should be noted that the number of test sections in each test zone according to EN 14363 reported in this article do not fully comply with the final recommended validation procedure, because the procedure and the conditions to be used were not known at the start of the investigations. Moreover, the test conditions during the DynoTRAIN running tests did not fully comply with EN 14363, see [17].

The selection of test sections considered geometrical track quality (irregularities) and wheel/rail contact geometry with the aim to include varying conditions. The track sections for exercises 2, 3 and 5 have been included because of a high vertical disturbance of track irregularities. The properties of the wheel/rail contact geometry were assessed by calculation of equivalent conicity and radial steering index over the constant curvature sections using the measured rail profiles averaged over 100 m distance together with nominal design wheel profile S1002 and mean track gauge over the respective track section. The definition of radial steering index has been introduced in UIC 518:2009 [12] to assess the available rolling radius difference between left and right wheel. The index values lower than 1 represent contact geometry which provides sufficient rolling radius difference for self steering wheelsets while values higher than 1 represent an insufficient rolling radius difference for the respective curve radius. The curve test sections 4, 5, 7, 9, 14 and 15 provide radial steering index below 1 and thus good contact geometry regarding curving, while sections 1, 2, 3, 6, 12, 13, 16 and 17 give radial steering index higher than 1, i.e. disadvantageous contact geometry conditions regarding self steering of wheelsets. The equivalent conicities (calculated for a wheelset lateral displacement of 3 mm) in section 8 were medium values of 0.20-0.25 and in section 9 were values around 0.1. The sections 10 and 11 were selected because of the occurrence of very high conicities; the conicity calculated per 100 m distance varies from medium values up to a few very high values of around 1.

Table 2 Test runs and parameters of track sections used in the validation exercises performed with vehicles tested in DynoTRAIN.

Exercise number	Line	Country	Test zone acc. to EN 14363	Curve radius [m]	Cant [mm]	Speed [km/h]	Section length: whole section A-F / constant curvature section C-D [m]
1	Geislingen-Westerstetten	Germany	4	282	120	68	740 / 400
2	Geislingen-Westerstetten	Germany	4	312	100	68	280 / 140
3	Geislingen-Westerstetten	Germany	3	572	155	110	1080 / 320
4	Uffenheim-Ansbach	Germany	3	580	150	110	870 / 490
5	Uffenheim-Ansbach	Germany	3	581	110	110	1130 / 680
6	Uffenheim-Ansbach	Germany	2	864	115	120	750 / 360
7	Uffenheim-Ansbach	Germany	2	694	160	121	690 / 190
8	Uffenheim-Ansbach	Germany	1	infinity	0	120	1760 / 1760
9	Würzburg-Fulda	Germany	1	5600 / 6000	75	200	3300 / 2644
10	Lichtenfels-Bamberg	Germany	1	infinity	0	160	3200 / 3200
11	Lichtenfels-Bamberg	Germany	1	infinity	0	160	3200 / 3200
12	Pisa-Firenze	Italy	4	295	140	76	504 / 110
13	Pisa-Firenze	Italy	4	292	140	76	968 / 771
14	Biasca-Göschenen	Switzerland	4	278	150	74	424 / 280
15	Biasca-Göschenen	Switzerland	4	294	142	74	384 / 192
16	St. Giovanni-Firenze	Italy	3	442	140	90	510 / 250
17	St. Giovanni-Firenze	Italy	3	406	150	90	651 / 426

As the freight vehicles were included in the test train only at speeds up to 120 km/h, the Laas wagon and the Sgns freight wagons were missing in the runs of the exercises 9, 10 and 11. Each simulation was performed for a part of the test run, called “part of interest” (A-F in Figure 3) and some outputs were evaluated over this part, while other outputs were evaluated over the part of the track with constant curve radius (C-D in Figure 3) only.

3. Simulation output and comparisons with measurements

3.1 Introduction

The simulations of selected on-track tests were evaluated in the same manner by all partners conducting simulations. This required an agreement and specification of the output data and its format.

As the aim of the validation is the application of simulation for the vehicle acceptance, a comparison of quantities as they are measured and evaluated according to EN 14363:2005 [20] was logically considered as one possible assessment method. Another typical validation assessment is a judgement of the comparison between the time domain signals from simulation and measurement. In contradiction to the quantities according to EN 14363 which are assessed primarily in the track sections with constant curvature, the judgement of time diagrams allows to assess also the behaviour in transitions as well as the signals' frequency content. A subjective judgement of time or distance diagrams thus represented another kind of assessment. An engineering judgement is however not measurable; a replacement of such assessment by calculable quantitative criteria would be preferred. The evaluation of so called validation metrics conducted recently by the Transportation Technology Center, Pueblo, USA [21] motivated the DynoTRAIN project partners to include the evaluation of the validation metrics as the third kind of assessment. These three kinds of validation assessment were applied on the investigated vehicle models and model configurations as shown schematically in Figure 2. The definition of these assessments and agreed simulation outputs are presented in the following chapters.

3.2 Assessment using values based on EN 14363

The comparisons between simulation and measurements were performed using an agreed set of output quantities which are used in the testing according to EN 14363. The simulation and measurement results were filtered and processed by analogy with EN 14363 requirements and compared against each other, whereby this evaluation considers the part of track with constant curvature only, i.e. section C-D in Figure 3. Table 3 shows the list of output quantities, their filtering, processing as well as the nomenclature and unit. A total of 2 wheelsets were used for the validation assessment of each vehicle, which resulted in a total of 28 quantities related to wheel/rail forces. The bogie accelerations were measured on the bogie frame above all wheelsets in lateral direction and above the wheelsets of one bogie in vertical direction, resulting in a total of 12 bogie frame acceleration quantities (not applicable for the 2-axle wagon). The vehicle body accelerations were measured on the floor level above both bogie centre pins in lateral and vertical direction resulting in a total of 8 car body acceleration quantities. Thus, a total of 48 quantities per model configuration and test section (36 for the 2-axle wagon) were applied consisting of quasi-static as well as dynamic wheel-rail forces and vehicle body and bogie frame accelerations.

Table 3 Output quantities used for the assessment by analogy with EN 14363

Quantity		Filtering	Processing	Notation	Unit
Wheel/rail forces, quasi-static values	Guiding force	Low-pass filter 20 Hz	50th percentile (median)	Y_{qst}	kN
	Wheel load	Low-pass filter 20 Hz	50th percentile (median)	Q_{qst}	kN
	Ratio Y/Q	Low-pass filter 20 Hz	50th percentile (median)	Y/Q_{qst}	-
	Sum of guiding forces	Low-pass filter 20 Hz	50th percentile (median)	ΣY_{qst}	kN
Wheel/rail forces, dynamic values	Guiding force	Low-pass filter 20 Hz	0.15 percentile, 99.85 percentile	Y_{max}	kN
	Wheel load	Low-pass filter 20 Hz	99.85 percentile	Q_{max}	kN
	Ratio Y/Q	Low-pass filter 20 Hz	Sliding mean (window 2 m, step 0.5 m) 0.15 and 99.85 percentile	Y/Q_{max}	-
	Sum of guiding forces	Low-pass filter 20 Hz	Sliding mean (window 2 m, step 0.5 m) 0.15 and 99.85 percentile	ΣY_{max}	kN
Bogie frame acceleration, rms values	Lateral acceleration	Band-pass filter 0.4 to 10 Hz	rms value	$\ddot{y}_{rms}^+$	m/s ²
	Vertical acceleration	Band-pass filter 0.4 to 10 Hz	rms value	$\ddot{z}_{rms}^+$	m/s ²
Bogie frame acceleration, dynamic values	Lateral acceleration	Low-pass filter 10 Hz	0.15 percentile, 99.85 percentile	$\ddot{y}_{max}^+$	m/s ²
	Vertical acceleration	Low-pass filter 10 Hz	0.15 percentile, 99.85 percentile	$\ddot{z}_{max}^+$	m/s ²
Car body acceleration, rms values	Lateral acceleration	Band-pass filter 0.4 to 10 Hz	rms-value	$\ddot{y}_{rms}^*$	m/s ²
	Vertical acceleration	Band-pass filter 0.4 to 10 Hz	rms-value	$\ddot{z}_{rms}^*$	m/s ²
Car body acceleration, dynamic values	Lateral acceleration	Band-pass filter 0.4 to 10 Hz	0.15 percentile, 99.85 percentile	$\ddot{y}_{max}^*$	m/s ²
	Vertical acceleration	Band-pass filter 0.4 to 10 Hz	0.15 percentile, 99.85 percentile	$\ddot{z}_{max}^*$	m/s ²

3.3 Subjective assessments

A subjective engineering judgement is based on a visual impression from time history plots and power spectral density (PSD) diagrams. A selected set of quantities consisting of 20 plots per vehicle model configuration and test section (for all vehicles besides the Laas freight vehicle with lower number of plots) was issued and provided to project partners for the assessment.

The following quantities were displayed and issued in the form of distance or time plots:

- Lateral wheel/rail forces (Y-forces): 4 diagrams per vehicle model configuration and test section
- Vertical wheel/rail forces Q (wheel loads): 4 diagrams
- Ratios Y/Q : 4 diagrams
- Bogie frame lateral accelerations above wheelsets 1 and 2: 2 diagrams
- Car body vertical acceleration above bogie 1: 1 diagram
- Car body lateral acceleration above bogie 1: 1 diagram

The simulation as well as measurement signals were filtered by a low-pass filter of 20 Hz, without any other processing, and displayed for the whole investigated test section (section A-F in Figure 3).

Moreover, power spectral densities of 4 acceleration signals were also provided as diagrams for subjective assessments:

- Bogie frame lateral accelerations above wheelsets 1
- Bogie frame vertical accelerations above wheelsets 1
- Car body vertical acceleration above bogie 1
- Car body lateral acceleration above bogie 1

These signals were filtered by low-pass 20 Hz and PSDs displayed for frequency range 0 – 10 Hz.

The project partners were asked to assess the diagrams presenting the comparisons of the measurement and simulation signal quantities by a simple binary assessment “Yes/No”.

Assessing a diagram with “Yes” states, that for the displayed signal quantities of the particular diagram the assessor considers the model as validated and vice versa.

As the form of the diagram (size, number of compared curves, scaling of axis, colours, position of curve in front or background, respectively) can influence the result of this judgement, it was first necessary to select and agree on a suitable form of diagrams. It was decided to present only two curves in each diagram, comparing measurement and simulation of the quantity’s distance or time history.

The selection of the scaling of vertical axis turned out to be a more difficult question.

Figures 4-7 show examples of comparisons between simulation and measurement for the following 4 investigated vehicle models:

- Locomotive DB BR 120 by Siemens
- DB passenger coach Bim by Bombardier Transportation
- Loaded freight wagon Sgns by Technical University Berlin
- Laas freight vehicle by Alstom.

Figure 4 presents the guiding force on the outer wheel of the leading wheelset from the test section 1 (curve radius 282 m) using the same scale for all vehicles to illustrate the differences in the level of the investigated values. As can be seen, the position of the signal is not exactly the same in regard to the distance. This may lead to slight differences when calculating the values in the specified interval with constant curvature. We can also observe other effects such as e.g. a signal offset (locomotive model by Siemens).

For illustration, the same results are displayed in Figure 5 in the original form as submitted for the subjective assessment, together with the percentage of positive assessments by project partners. The diagrams were assessed by 10 partners, i.e. 40% means that 4 from 10 partners considered the presented results as documenting a validated model. Besides a differing scale of vertical axis, the Laas freight vehicle results are displayed as a time diagram over a longer interval compared to other vehicles which are presented as distance diagrams. We can see that the assessment of the very light two axle wagon Laas is rather strict compared to the results of the locomotive or loaded freight wagon.

Figure 6 shows the ratio Y/Q at the outer wheel of the leading wheelset from the test section 2 (curve radius 312 m), Figure 7 the vertical car body acceleration from the test section 8 (straight). While the ratio Y/Q is of similar level at all vehicles, the accelerations vary significantly between the vehicles. This opens the question of the selection of scaling for the presentation of results. When using an equal scaling, the comparison for light vehicles can hardly be assessed as they have low vertical as well as lateral wheel/rail forces. Also the assessment of accelerations of soft suspended vehicles would be difficult. Alternatively, the use of automatically adjusted scaling leads to the impression of large differences even if the values are very small. To allow better comparable assessment, it was proposed to the project partners to use a fixed scaling with one of three specified scale groups; however, the final decision was up to the partner conducting the simulations.

Consequently, the values presented in the evaluated diagrams are sometimes rather small, while in other cases the peaks are outside of the diagram; both effects make the subjective assessment more difficult.

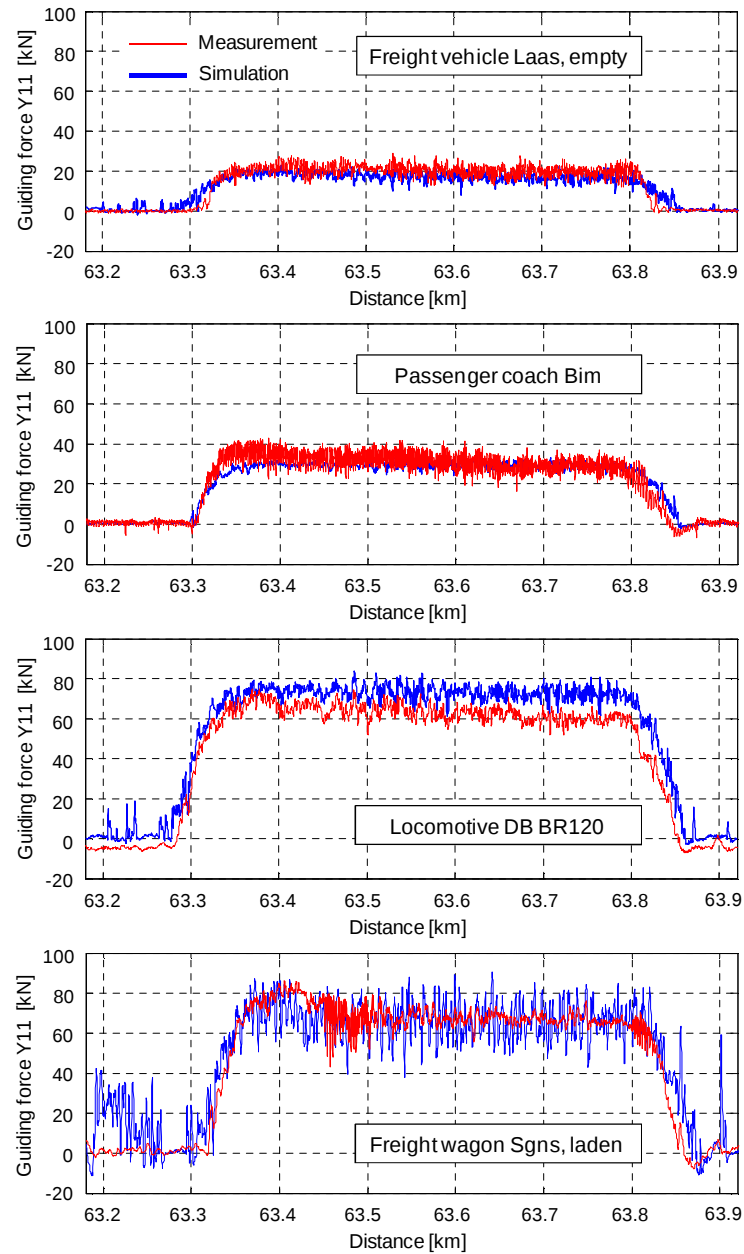


Figure 4. Validation examples: Guiding force on the outer wheel of leading wheelset, exercise 1, Germany, line Geislingen-Westerstetten, curve radius 282 m, cant 120 mm, speed 68 km/h.

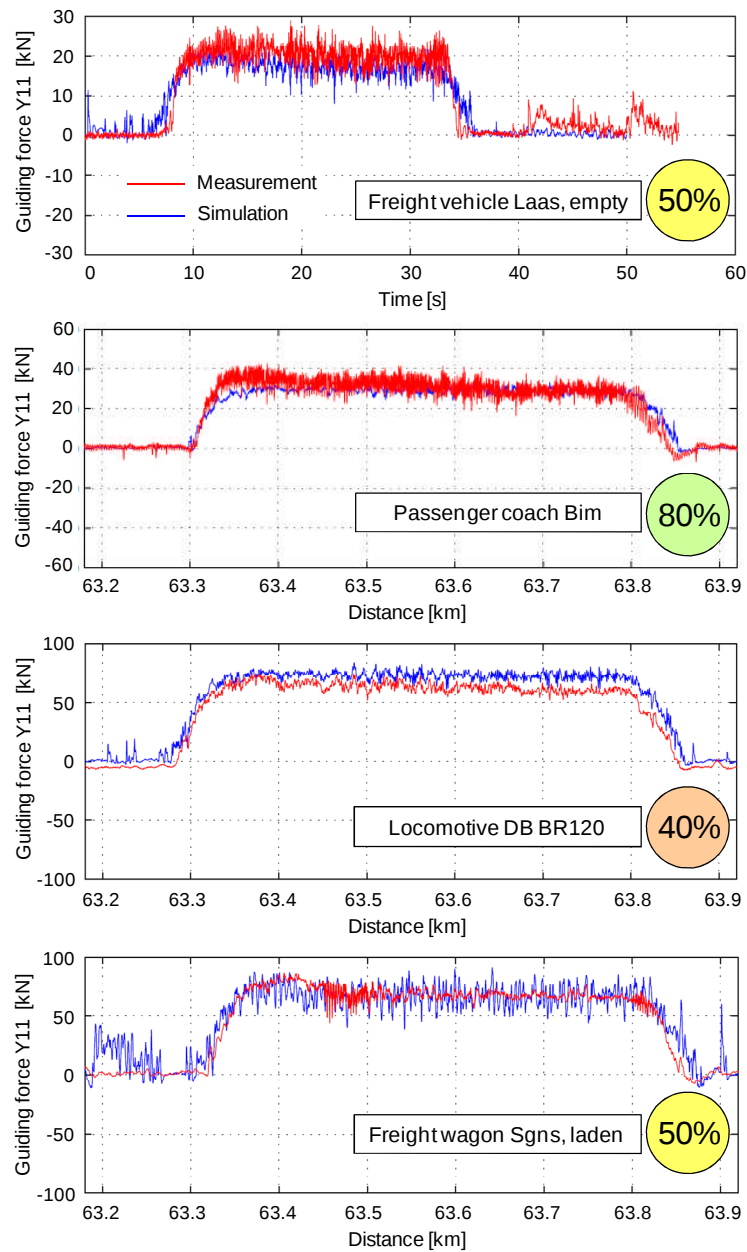


Figure 5. Validation examples and subjective assessments. Diagrams from the exercise 1 (as in Figure 4) in the form used for the subjective assessments by project partners. The values in the circles of each diagram display the percentage of positive assessments.

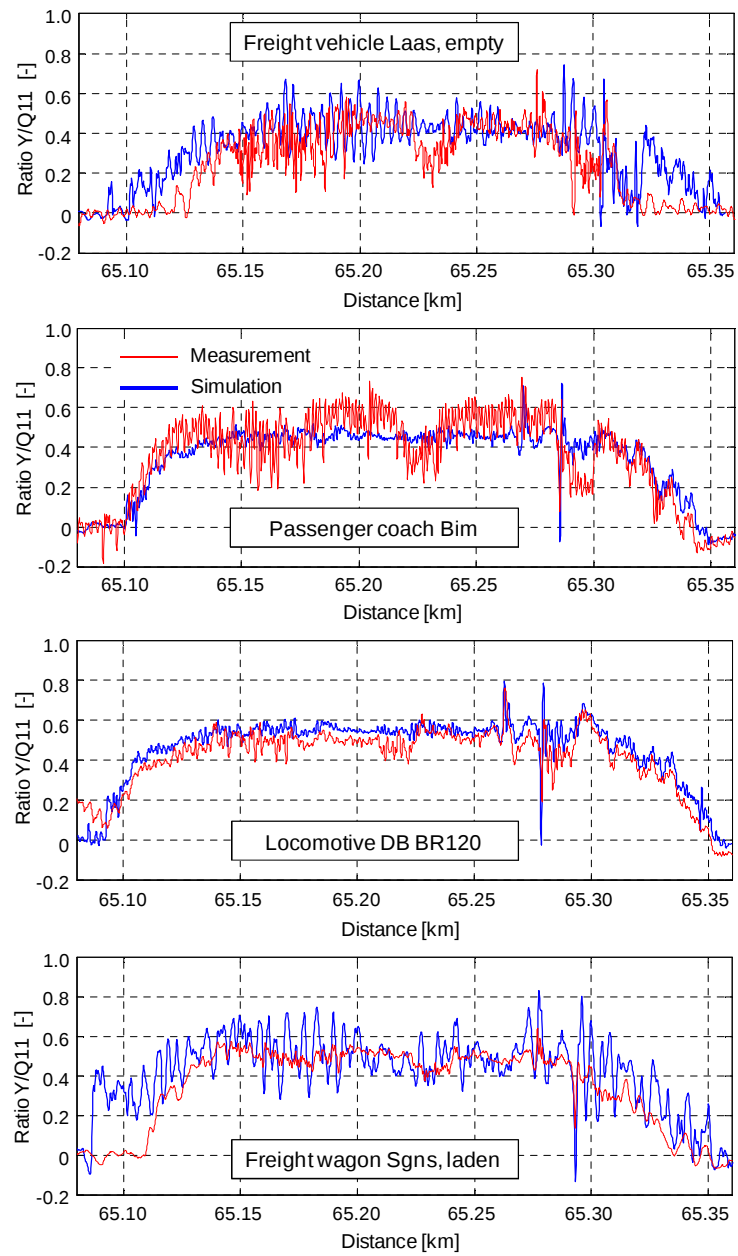


Figure 6. Validation examples: Ratio Y/Q on the outer wheel of leading wheelset, exercise 2, Germany, line Geislingen-Westerstetten, curve radius 312 m, cant 100 mm, speed 68 km/h.

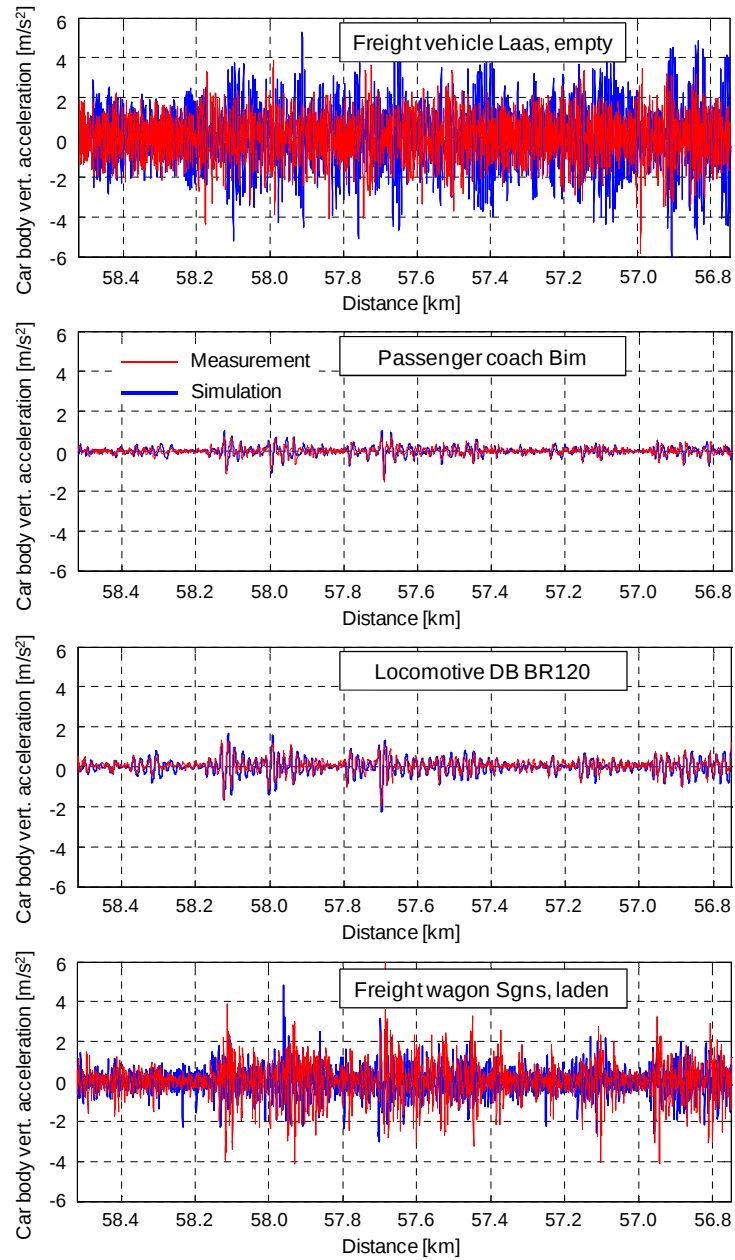


Figure 7. Validation examples: Vertical acceleration of vehicle body over the leading bogie (wheelset), exercise No. 8, Germany, line Uffenheim-Ansbach, straight track, speed 120 km/h.

3.4 Validation metrics

The validation assessment by comparisons of time histories between simulated and measured values generates questions caused by the subjectivity of this assessment as stated in the previous chapter. Validation metrics represent an approach in quantifying the comparisons of time history curves with the intent of minimising the subjectivity while still

maintaining a correlation with experts' opinion [22]. They are developed and used mainly for comparisons between simulation and measurement in the context of model validation. A possible metric used to compare the time domain diagrams is the integral approach introduced in 1984 by Geers. Integrals of two wave forms to be compared are computed and used to evaluate the difference in the magnitude and phase of the wave forms expressed by magnitude, phase and comprehensive error factors, with small values of error factors representing good agreement. The magnitude as well as phase form of error factors was later adapted by Russell [23]. The new phase form by Russell was combined with the 1984 Geers' metric by Sprague and Geers in [24]. By using the same sampling rate and the same length of time or distance interval for the compared measurement and simulation signals, the definitions of error factors proposed in [24] can be expressed by the following formulae [25].

- Sprague and Geers magnitude error factor:

$$M_{SG} = \sqrt{\frac{\sum_{i=1}^n c_i^2}{\sum_{i=1}^n m_i^2}} - 1 \quad (1)$$

with c_i - simulated values
 m_i - measured values.

- Sprague and Geers phase error factor:

$$P_{SG} = \frac{1}{\pi} \cos^{-1} \left(\frac{\sum_{i=1}^n c_i m_i}{\sqrt{\sum_{i=1}^n c_i^2 \sum_{i=1}^n m_i^2}} \right) \quad (2)$$

- Sprague and Geers comprehensive error factor:

$$C_{SG} = \sqrt{M_{SG}^2 + P_{SG}^2} \quad (3)$$

The error factors of validation metric by Sprague and Geers according to equations (1)-(3) as well as the error factors by Russell [23] were calculated by the project partners for the comparisons between simulations and measurements provided in the time and distance domain plots and used for the subjective assessment by the partners.

4. Evaluation of validation method, criteria and limit values

4.1 Evaluation of the assessment based on EN 14363

The assessments based on quantities according to EN 14363 were carried out using a common preliminary set of validation limits, which were evaluated from the proposals provided by the project partners. These proposals deviated significantly not only in the proposed limit values but also in principle as shown schematically in Figure 8 displaying the areas fulfilling the proposed validation condition. If the simulated value S_v and measured value M_v are identical, the point is on the diagonal line. A deviation from this diagonal line represents the deviation between the simulation and measurement.

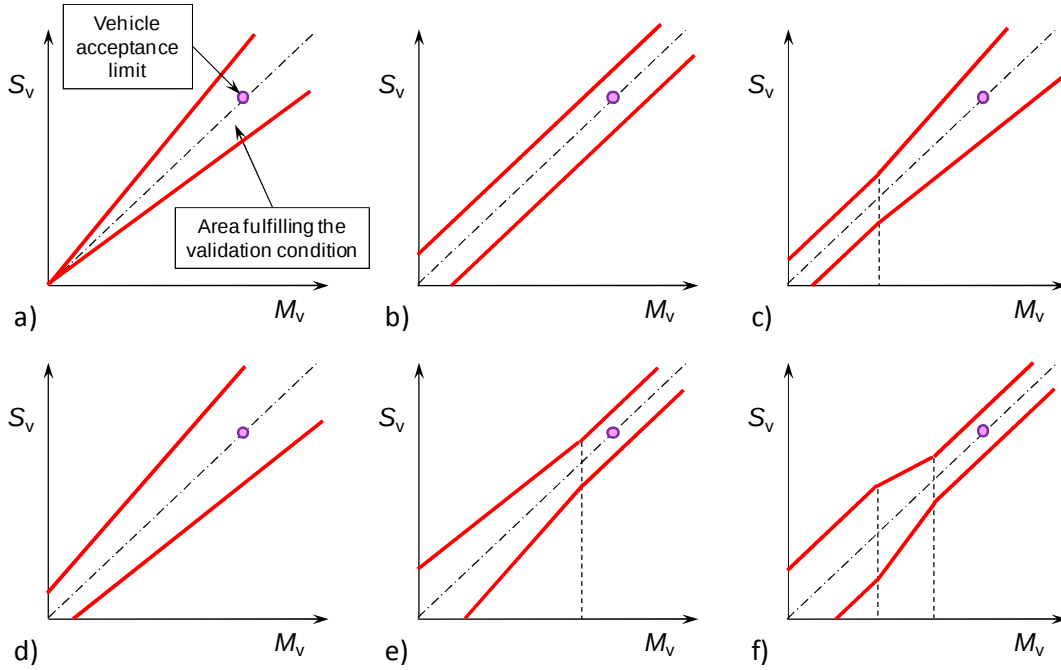


Figure 8. Principal differing definitions of the validation limit conditions proposed by project partners.

The following three principal differing definitions of the limit condition were proposed:

- Deviation limit as a percentage of the measured value (relative deviation limit) – diagram a) in Figure 8,
- Constant deviation limit (absolute deviation limit) – diagram b) in Figure 8,
- Deviation limit decreasing with the measured value increasing towards the limit for vehicle acceptance according to EN 14363, but not falling below a minimum absolute limit at high measured values, as shown in diagram e) in Figure 8.

Some partners proposed combinations of previous principles – a relative limit combined with absolute deviation limit as shown in diagram c), an addition of absolute and relative deviation limit as displayed in diagram d) or an absolute (constant) deviation limit changing with measured value according to the diagram f) in Figure 8.

A reasonable justification can be provided for each of the differing proposals. Any deviation or error is usually considered in regard to relative deviation thus supporting the approach a) in Figure 8. However, as the vehicle model is intended to be used for simulation of vehicle acceptance tests, it is important to achieve good agreement especially for the values, which are close to their limit values of vehicle acceptance, hence supporting the contradicting approach e) in Figure 8. Finally, it was agreed to use constant validation limit values (limits for absolute deviation simulation - measurement), which is quite simple and at the same time the most appropriate compromise for the proposals discussed during the investigations. A set of preliminary validation limits based on the partners' proposals was agreed and then applied for comparisons of model configurations and for the investigation of the possible approach for model validation.

4.2 Evaluation of subjective assessments

Comparisons of measurement and simulation of quantities presented in diagrams were assessed by the project partners using a simple “Yes/No”-method. Due to large amount of results presented in diagrams, only a part of all results could be assessed by project partners. The following model configurations of vehicles tested in DynoTRAIN were selected for this subjective assessment, all representing the initial vehicle models:

- Configurations F1 using measured data of wheel and rail profiles as well as measured track irregularities
- Configurations D1 using estimated (design) wheel and rail profiles and measured track irregularities
- Configurations E1 using measured wheel profiles, estimated (design) rail profiles, measured track irregularities
- Configurations C1 using measured wheel and rail profiles, but estimated track irregularities.

These subjective assessments totalled over 6,000 diagrams, each assessed by 7-10 project partners, which resulted in more than 50,000 single assessments.

Moreover, a workshop with invited experts dedicated to model validation was held on 7th November 2012, hosted by Siemens AG in Krefeld (Germany). A total of 26 workshop attendees (professors for railway vehicle dynamics, experts from industry, railway companies, testing and research institutes, members of the standardisation committee as well as DynoTRAIN project partners) participated in the subjective engineering judgement of diagrams. The assessments questionnaire contained 110 selected time or distance plots and 10 power spectral density diagrams. The workshop was intended to collect data about the visual assessment of diagrams containing comparisons between the simulation and measurement.

An assessment of a vehicle simulation model requires knowledge about the vehicle itself and about the boundary conditions of the comparison (i.e. kind and quality of available measurement data and vehicle model parameters). The information collected in the workshop was intended to be used to investigate the feasibility of replacing the subjective engineering assessment of a single diagram with an objective metric about the degree of similarity between simulation and measurement. For this reason the workshop procedure stressed the importance of focusing on each single diagram and the workshop attendees were asked to assess each diagram separately by a simple Yes/No-method under the following considerations:

- Assume that a sufficient number of diagrams have already been assessed, each one containing a comparison between simulation and measurement of the particular vehicle.
- Assume that until the current, last diagram, all previous diagrams were considered as satisfying the validation criteria; some of the previous diagrams, however, did not show a good agreement, so that there are still doubts, if this model can be confirmed as validated.
- Answer, if the current diagram confirms that the actual vehicle model can be considered as validated or if it confirms your doubts and this vehicle model thus cannot be validated.

It was intended to ask for an engineering judgement based on a pure visual impression from the assessed diagram, not to be biased by any consideration about the actual boundary

condition of the simulation or any consideration about the reasons, why the signals show a particular behaviour. Thus, the requested judgement could be transformed in to a computable measure calculated using the data presented in the diagram without considering any other boundary condition.

The results of the workshop assessments showed strong variation. Only 6 from a total of 120 plots were assessed unequivocally; an equal assessment by more than 75% of attendees was provided for 54 plots (45%) of diagrams.

Although it was not possible to conclude about a replacement of the assessment results by computable values of investigated validation metrics, this workshop provided an interesting knowledge. The form of the presentation of diagrams comparing the simulation and measurement (scaling of diagrams, exchange of signals back/front) significantly influenced the assessment result. From 6 pairs of two plots presenting identical data using differing scale, only one set received the same assessment for both diagrams. The remaining 5 diagram pairs were assessed differently, see example in Figure 9.

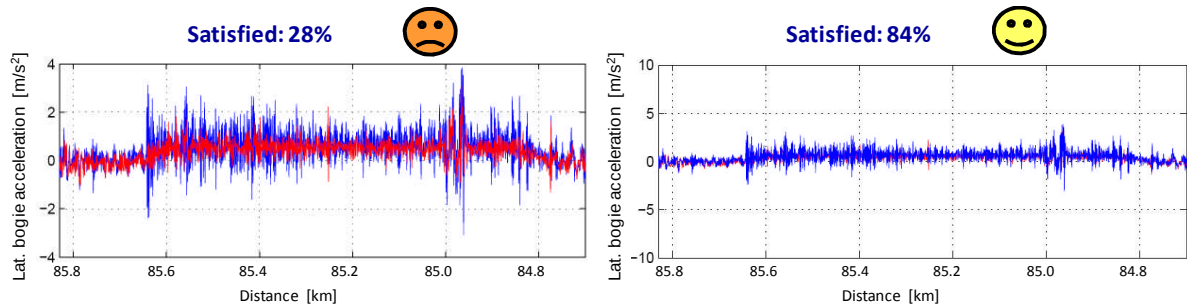


Figure 9. Example of workshop results displaying the effect of the form of diagrams on the assessment of plots presenting identical data. In the right diagram, the scale of vertical axis is enlarged and the forward and background signal exchanged (see on-line version for colours).

Furthermore, the workshop results showed large differences in the “level” of the strictness of the individual assessors. This can be seen in Figure 10 displaying the percentage of positive assessments in each of the six groups of plots provided by the particular attendee. The workshop attendees are ordered from more strict on the left to less strict on the right side. No correlation could be identified between the attendee’s strictness and any of the considered categories according to their affiliation or experience.

Although the workshop assessments were related to single diagrams only, without any background information about the vehicle type, test conditions and simulation procedure, and thus cannot be considered as representative validation assessments, they illustrate the weakness of subjective judgements. Therefore, it can be concluded, that a subjective assessment by engineering judgement is not ensuring the feasibility of an objective model validation.

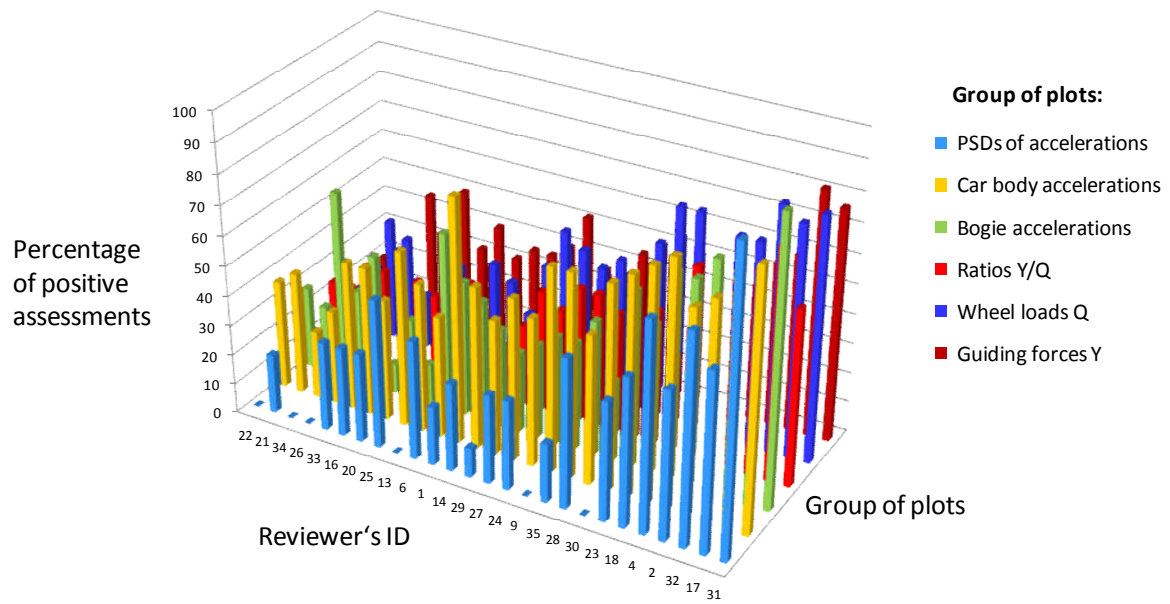


Figure 10. Workshop results: Percentage of positive assessments provided by particular attendee in each of the six groups of plots.

4.3 Evaluation of validation metrics

The investigations dedicated to validation metrics were introduced with the aim to replace a subjective engineering judgement of time or distance plots by a computable and thus objective measure. The previous discussion showed deviations between engineering judgements provided by different assessors, which will surely make a replacement of such judgement more difficult. Moreover, the judgement can further deviate dependent on the form and scaling of the diagrams in question as discussed in the chapter about the simulation output and comparisons with measurements.

These facts can partly explain the initially surprising effect of a missing correlation between the subjective assessments by project partners and the error factors of the investigated validation metrics by Sprague and Geers or by Russell, respectively. Nevertheless, the cases resulting in an unexpected disagreement between the validation metric and subjective assessments (high error factors for diagrams with high percentage of positive assessment and vice versa) were further analysed to understand and possibly modify the validation metrics. These analyses identified the following three possible reasons of disagreement between the subjective assessment and validation metrics as demonstrated on examples in [18].

- The validation metric error factors are based on a relative deviation, thus not considering the magnitude of the evaluated quantity. A relative deviation between simulation and measurement at a very low magnitude of measured quantity is usually neglected in engineering judgement; however, it can provide large error factor value suggesting large disagreement. Although the Russell's definition of the magnitude error factor aims to correct this effect, it is not well suited for the investigated application because of large differences in magnitudes of different quantities.

- Another drawback of the validation metric is a strong influence on phase error factor by the level of synchronisation between simulation and measurement signals. A perfect synchronisation is not easy to achieve and is usually not requested, which can lead to high values of the phase error as well as the comprehensive error factors suggesting disagreement between simulation and measurement in spite of positive visual judgement.
- The third identified drawback can occur in case of superposition of dynamic oscillation with a rather high constant quasi-static value. In this case, the resultant integrals will be given by the quasi-static value of the investigated quantity. Thus, if there is agreement of the quasi-static results between simulation and measurement, the error factors will be low and likewise for the case when there is a disagreement in dynamic values. This results in error factors suggesting very good agreement in spite of low subjective acceptance.

4.4 Evaluation of final validation method, criteria and limits

The variations of model input data, model adjustments and modelling depth together with variations of track input data resulted in a total of 78 model configurations of the investigated vehicles and more than 1,000 simulations of validation exercises. The correlations between the different groups of assessment (EN 14363 quantities, subjective assessments, validation metrics) as well as the relationship between the assessments and the achieved results were investigated as shown in Figure 11.

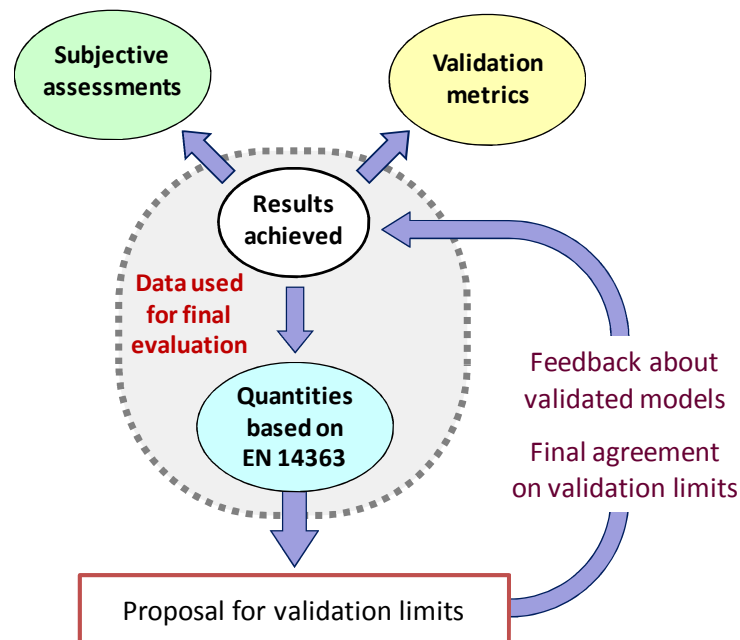


Figure 11. Schematic presentation of the process used to compare different kinds of validation assessment and to evaluate the final proposal.

Summarising the correlation analyses and other project results, it is believed that the comparisons of simulation and measurement using quantities based on EN 14363 represent the best suited methodology for the model validation in the context of vehicle acceptance. Subjective engineering judgement can vary dependent on the strictness of the reviewer, and the validation metric, which was considered as suitable for replacement of the subjective judgement, does not show any valuable improvement compared to the assessment using quantities based on EN 14363 and was therefore not considered in the final proposal.

The analyses of the deviation between simulation and measurement conducted in DynoTRAIN demonstrated that an assessment of a single particular quantity and single pairs of the simulated and measured values do not provide relevant information about the model quality. It is in fact more important to check the overall agreement instead of concentrate on single maximum differences between simulation and measurement. A single deviation between simulation and measurement can be related to a particular effect in the measurement or a particular deviation between conditions during the measurement and the input parameters used in the simulations. Moreover, it is left to chance, if such a single deviation will be identified, when selecting the test sections used for comparisons. Therefore, the model validation should approve the overall agreement of the deviations between compared pairs of simulation – measurement. Hence, a statistical approach has been selected to assess this overall agreement calculating the mean value and standard deviation of differences between the simulation value S_v and the measurement value M_v for each of 12 agreed quantities according to EN 14363 (e.g. for all Y_{qst} values) for a specified minimum of test sections representing the conditions for vehicle acceptance as it is described in detail in the next chapter. The minimum to be used for validation has been agreed as three sections from each test zone according to EN 14363, thus at least 12 sections, and a minimum of 2 different measuring signals per quantity. Using two force measuring wheelsets to fulfil the later requirement for quasi-static and maximum value of the sum of guiding forces, there are 48 pairs simulation - measurement for each of quantities Q , Y and Y/Q , which results in a total of 432 compared pairs between simulated and measured values, see Figure 12. The validation evaluations conducted in DynoTRAIN used even more compared pairs. They included 14 sections for freight vehicles and 17 sections for other vehicles, resulting in 504 or 612 compared pairs, respectively.

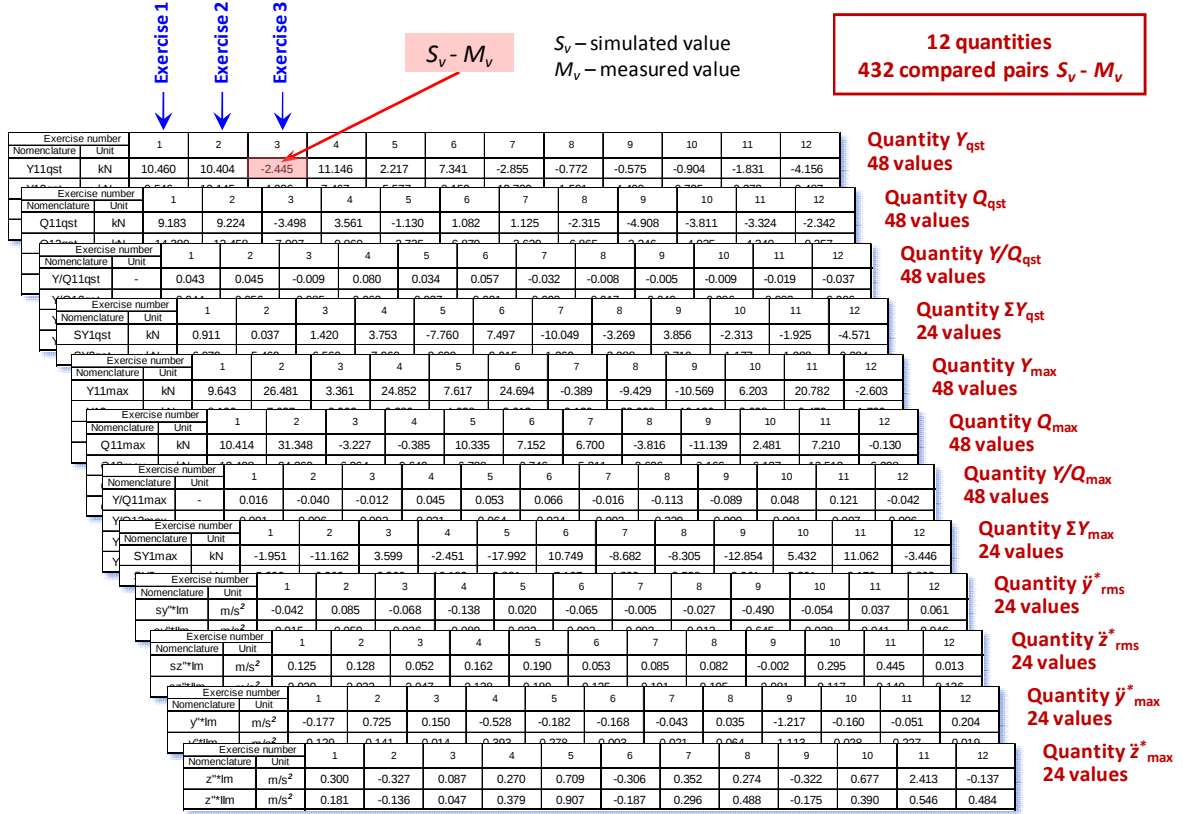


Figure 12. Example of a typical set of comparisons between simulation and measurement according to the proposed validation method.

The preliminary validation limits agreed in an earlier step of the project were used to assess the validation of the investigated 78 model configurations. The feedback about the validated models was then used for the final adjustment of the validation limits as can be seen in the schematic presentation of this process in Figure 11. Because the validation limits for wheel loads (both quasi-static as well as dynamic) are very sensitive to the static wheel load, a dependency on the static wheel load was introduced instead of constant validation limit values. The level of vehicle body accelerations of freight vehicles and vehicles without bogies or without secondary suspension is significantly larger than that of vehicles with typical soft secondary suspension; therefore, the validation limits for the vehicle body accelerations of those vehicles were doubled to account for this effect. The accelerations at the bogie frame were evaluated, but not proposed as a mandatory quantity for model validation. The dynamic behaviour of the bogie or running gear of the particular vehicle model is sufficiently approved by checking the quantities in the wheel/rail contact. Moreover, the investigations carried out showed that the application of bogie frame accelerations for model validation and the justification of suitable validation limits is rather difficult and not really necessary as the bogie dynamics is assessed by wheel/rail quantities anyway. The investigations dedicated to power spectral density (PSD) diagrams showed a large variety of results and of deviating assessments by partners as well as during the workshop. An introduction of criteria and quantitative limits in regard to PSD was not possible and would need further investigations.

5. Proposed validation method

The proposed validation process is based on a mathematical comparison between the results of on-track tests performed using the normal measuring method and the corresponding simulation results. The simulation and measurement results of the specified quantities have to be compared on at least 12 track sections, called validation exercises. A section can be either a test section according to EN 14363 or a part of a test track longer than the minimum length specified for track sections in the respective test zone. Moreover, these sections have to satisfy the other test section requirements such as constant curve radius. The selected validation exercises have to contain sections from all 4 test zones, at least 3 sections from each test zone. The track geometric irregularities have to represent the conditions of the on-track tests.

Each quantity has to be evaluated using at least two signals, e.g. vertical acceleration above the leading and trailing bogie, thus, at least 24 simulated values S_v compared to the corresponding measured values M_v of each quantity are required. Each compared simulated as well as measured quantity has to be filtered and processed according to the requirements in Table 4. The percentiles have to be calculated from the cumulative curve. For the maximum value calculated as 0.15% or 99.85%-value, the higher magnitude of the 0.15%- and 99.85%-values (absolute value) is used. The 50%-values (medians) are applied with their sign to show the agreement of both magnitude and direction of those quantities. The difference D_v between the simulated value S_v and the corresponding measured value M_v has to be evaluated for each value and each quantity, whereby this difference has to be transformed so that, if the magnitude of the simulation value is higher than the magnitude of the measurement (simulation overestimating the measurement), the difference is positive, and vice versa:

$$D_v = (S_v - M_v) \frac{M_v}{|M_v|} \quad \text{for } M_v \neq 0 \quad (4)$$
$$D_v = S_v \quad \text{for } M_v = 0$$

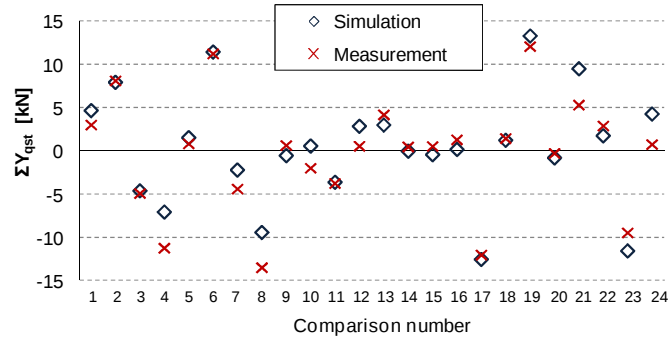
The following values have to be calculated for the whole set of differences D_v between the simulation and measurement for each quantity:

- Mean of differences between simulation value S_v and measurement value M_v
- Standard deviation of the same set of differences.

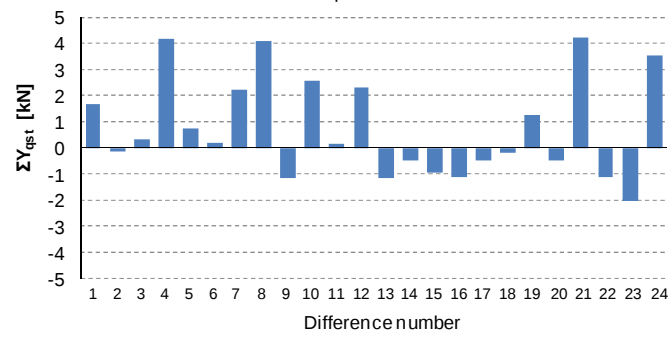
The standard deviation of the set of differences between simulation value S_v and the measurement value M_v for each individual quantity has to be not higher than its validation limit shown in Table 4. For each quantity the mean of the set of differences between the simulation value S_v and the measurement value M_v should not be higher than the validation limit equal to 2/3 of the respective limit for the standard deviation. The validation limits for accelerations (standard deviation as well as mean of differences) for freight vehicles or vehicles without secondary suspension are twice the relevant limit values for other vehicles.

As an example, Figure 13 explains the calculation of differences between the simulation value S_v and the measurement value M_v for the quasi-static values of the sum of guiding forces between wheelset and track, their transformation as well as calculation of mean value and standard deviation, which are used for comparison with the validation limits specified in Table 4.

Compared values from simulation and measurement



Differences between simulation and measurement



Transformed differences between simulation and measurement together with mean and standard deviation

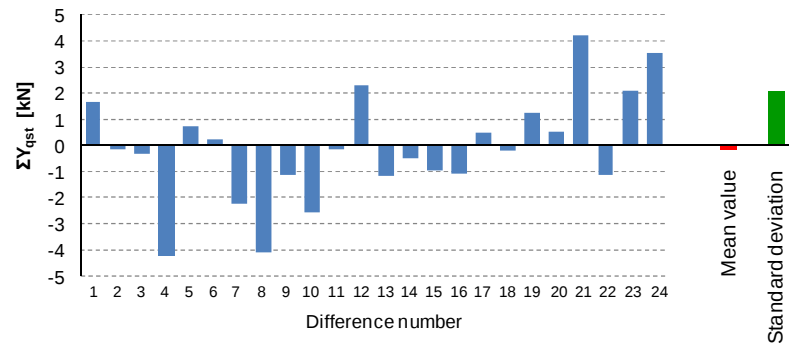


Figure 13. Example of data evaluation: Simulated and measured values, their differences, transformation of differences in regard to the sign of measured value and calculation of the mean value and standard deviation of the quasi-static sum of guiding forces ΣY .

Table 4 Quantities and limits for model validation in regard to simulation of on-track test

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 value S_v as well as measured value M_v

**) For freight vehicles and vehicles without bogies or without secondary suspension, these limits have to be doubled

6. Discussion

6.1 Advantages of the proposed validation method

The proposed final set of validation limits was applied to assess the validity of all the investigated model configurations. From a total of 78 model configurations evaluated, only 20 fulfil the proposed model validation limits:

- 8 from 24 models of locomotive BR 120 by Siemens
- 10 from 13 models of Bim coach by Bombardier Transportation
- 2 from 4 models of Bim coach by IFSTTAR.

The validated models are only the models of vehicles tested in DynoTRAIN, validated using measured track irregularities as well as measured wheel and rail profiles. The models successfully validated are only the models of locomotive BR 120 and Bim coach. Not one model configuration of the freight vehicles could be validated. Furthermore, not one model configuration without measured track irregularities could be validated.

The contributions of quantities leading to the failure of 58 from a total of 78 model configurations are displayed in Figure 14. The failure of validation can be caused either by one quantity or by more quantities at the same time; the exceedance can be due to either the standard deviation of differences, or the mean value of differences or both values. The

most frequent cause was an exceedance of the maximum value of car body vertical acceleration. Other rather common causes were Y/Q (quasi-static as well as maximum value), Y_{qst} and ΣY_{max} . The wheel loads seldom cause the limits to be exceeded and there was no exceeding of the validation limit for the mean value of the differences of Q_{qst} . Thus, it seems that expected model exactness using the proposed validation method can be easily achieved for the vertical wheel forces whereas it is rather difficult to achieve for the quantities as ratio Y/Q or vertical vehicle body accelerations.

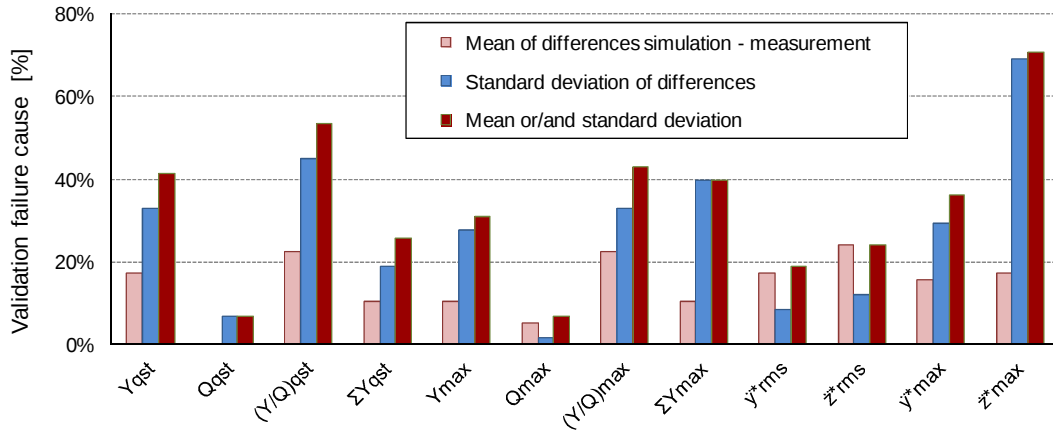


Figure 14. Contributions of individual quantities to the validation failure for the 58 from a total of 78 investigated model configurations.

An important advantage of the proposed validation procedure is that this assessment represents an overall assessment of a large number of data which is impossible to carry out by using engineering judgement of the plots, as it is not practically possible to display, check and document the approval of such a large number of plots. The calculation of characteristic parameters of mean and standard deviation of differences between the simulation values S_v and the measurement values M_v , and their comparison with the validation limits, however, allows a fast identification of quantities with large deviation. The data of a particular quantity can be easily checked in detail to identify the validation exercise (section) and the signal (sensor position) which provides a large deviation between the simulation and measurement. The specified set of 12 quantities covers the quasi-static as well as dynamic behaviour of the vehicle in regard to the vehicle acceptance, which is the intended range of the application for a validated model. The signal processing is carried out by analogy with EN 14363:2005 for both the measurement and simulation, thus allowing direct use of the acceptance tests data.

The weakness of the model in question can be identified by a normalisation of the validation criteria, dividing them by the validation limits as it can be seen in Figure 15. The model is validated, if the absolute magnitudes of all displayed values are not higher than 1. The vehicle models in Figure 15 were prepared by using the available parameter data, before any model adjustments by comparisons with either stationary or on-track tests. This figure shows normalised values of mean (left) and standard deviation (right) of the differences between simulation and measurement for two vehicle models of the locomotive BR 120 by Siemens. The initial model F1 does not fulfil the proposed validation limits.

This model used measured track irregularities as well as measured wheel and rail profiles, but has not been adjusted based on the measurements. The improved model T2 after adjustments by comparisons with on-track tests and with stationary tests fulfils the validation limits.

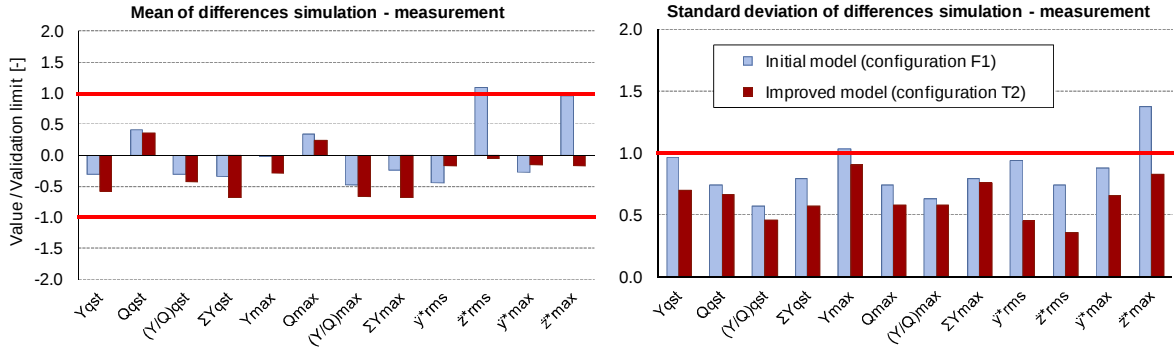


Figure 15. Example of validation results using the proposed method. Normalised values of mean (left) and standard deviation (right) of the differences between simulation and measurement for two vehicle models of the locomotive BR 120 by Siemens: Initial model F1 and improved model T2 after comparisons with on-track as well as stationary tests.

The results confirm the proposed validation criteria and limits as a suitable and robust methodology for validation of railway vehicle models. The proposed validation method allows not only an objective assessment, but also a clear identification of the model weaknesses, see also [26].

6.2 Effect of model adjustment using stationary tests on the simulation of on-track tests

Static and low speed tests can be used to identify missing or uncertain vehicle model parameters and to support the vehicle model validation. A comparison of simulations with available stationary measurements is required as a part of the model validation process in prEN 14363:2013 [13]. The simulation and measurements of the stationary tests are compared and the uncertain model parameters adjusted if necessary [27]. But what is the effect of a model adjustment based on the comparison with stationary tests on the agreement between simulation and measurement of the on-track test (ride test)? This is typically neither presented nor investigated; it is believed that an improved agreement with stationary tests will implicitly improve the exactness of the on-track test simulation. In order to investigate this effect, the validation exercises with on-track tests were repeated with several versions of the same model, either before the comparison with stationary tests or after comparison and adjustment in order to fit the stationary tests results, respectively. The stationary tests used during the validation evaluations of simulation models differed dependent on the availability of the test results. An overview of the stationary tests used for comparisons and model adjustments of vehicles tested in DynoTRAIN is shown in Table 5. Not all comparisons resulted in a proposal of model adjustment. Particularly the tests in a flat curve with a radius of 150 m representing a part of the test of safety against derailment

on twisted track according to Method 2 in EN 14363:2005 [20] did not provide any support regarding the parameter identification of the investigated vehicles.

Table 5: Comparisons with stationary tests and adjustments of the vehicle model. The tests used for comparisons are marked by cross (X); “no adjustment” is stated if the comparison with the respective stationary test did not provide any suggestion for model adjustment.

Vehicle model	Wheel unloading test (twist test) according to EN 14363:2005, Chapter 4.1.3.3	Test in flat curve R = 150 m according to EN 14363:2005, Chapter 4.1.3.3	Bogie rotational resistance test according to EN 14363:2005, Chapter 4.4	Bogie lateral resistance test to measure the characteristic of the lateral suspension	Sway test - measurement of roll coefficient
Locomotive DB BR 120, Siemens	X	X (no adjustment)	X	X	X
Locomotive DB BR 120, IFSTTAR	–	–	X	X	X
DB passenger coach Bim, Bombardier Transportation	X	X (no adjustment)	X	X	–
DB passenger coach Bim, IFSTTAR	–	–	X (no adjustment)	X	–
Freight wagon Sgns, empty, Technical University Berlin	–	X (no adjustment)	X	–	–
Freight wagon Sgns, empty, IFSTTAR	–	–	X	X	–
Freight wagon Sgns, laden, Technical University Berlin	–	X (no adjustment)	X	–	–

Figure 16 shows comparisons of the validation results using the proposed validation method for six models of vehicles tested in DynoTRAIN WP1. The figure presents comparisons of the initial model configurations F1 using measured track irregularities and measured wheel as well as rail profiles however, without any model adjustments based on comparisons with stationary or on-track tests, and model configurations T1 after adjustments based on comparisons with stationary tests stated in Table 5 (in case of laden freight wagon Sgns by TU Berlin the compared configurations are F2 and T2 with modified suspension modelling).

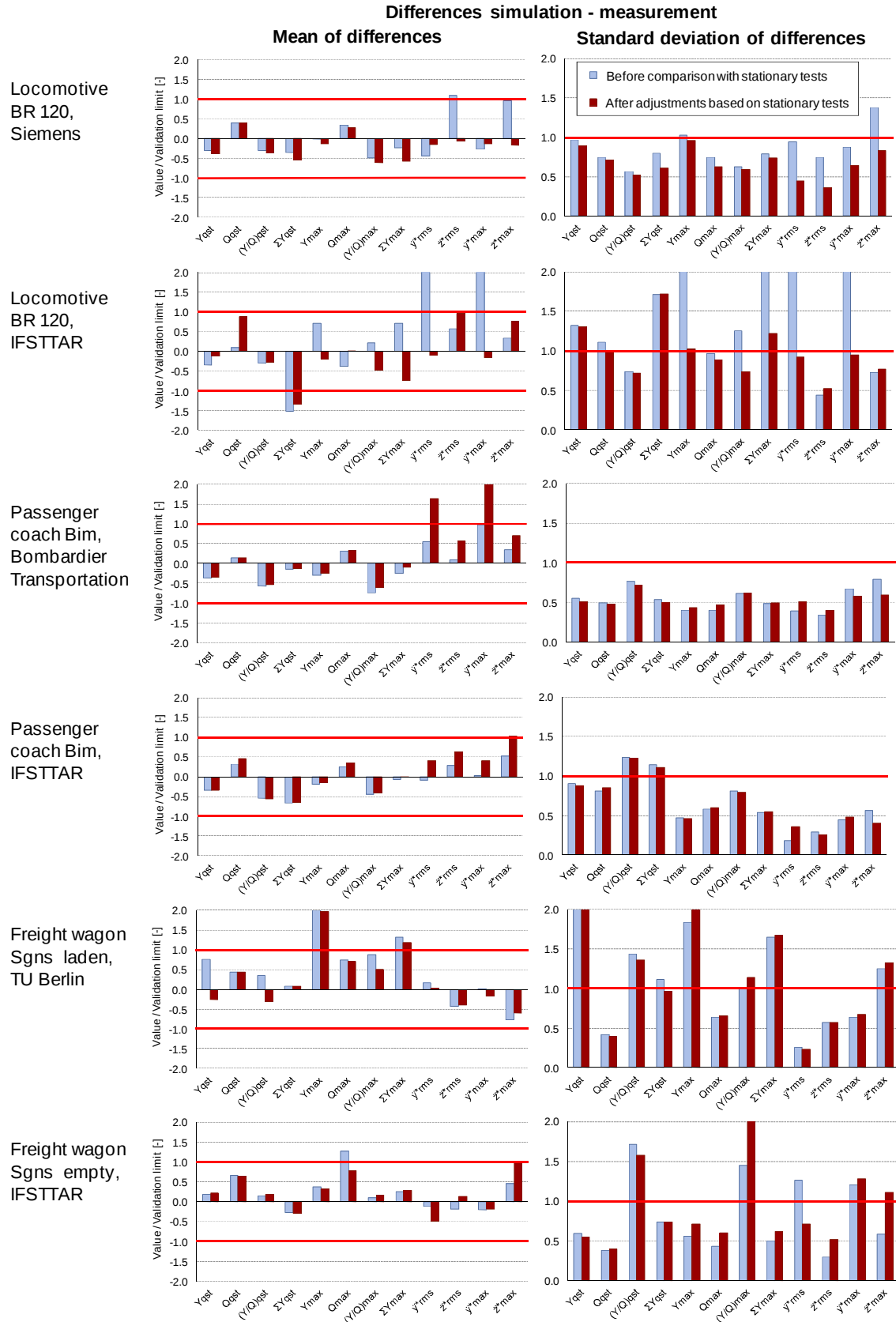


Figure 16. Effect of comparisons with stationary tests on the validation results. Normalised values of mean (left) and standard deviation (right) of the differences between simulation and measurement for the initial vehicle models and models after adjustments based on comparisons with stationary tests.

The model adjustments by comparisons with stationary tests did not lead to expected improvements of the results regarding the simulations of the on-track tests. Only the investigations by Siemens and IFSTTAR regarding the locomotive BR 120 provided clearly better results for the models after the comparison and adjustment due to the stationary tests. In other cases, the model adjustments introduced using the stationary tests did not significantly affect the agreement between the simulation and measurement concerning the on-track tests or provided even worse results. For example, in the model of Bim coach by Bombardier Transportation, an implementation of friction elements intended to model a rather small hysteresis in the secondary lateral suspension resulted in the failure of the model validation due to significantly higher lateral car body accelerations compared with the values measured during the on-track test.

These investigations did not confirm the traditional opinion regarding the positive effect of the model adjustments by comparisons with stationary tests on the simulation of on-track tests. A possible explanation is that focussing on the static and low speed behaviour can result in model adjustments which are less accurate in regard to dynamic behaviour. The stationary tests can support the identification of model parameters which were unknown or uncertain. A good agreement between simulation and measurement of stationary tests, however, does not guarantee a good agreement between simulation and measurement of on-track tests. An adequate amount of comparisons between simulations and on-track measurements represents the only suitable and reliable model validation method in regard to the simulation of on-track test.

7. Summary and conclusion

The presented part of the investigations in the framework of the DynoTRAIN project was dedicated to evaluation of the validation method suited for simulations in the context of vehicle acceptance. It represents a unique activity of complex testing, simulations, comparisons with measurements and evaluations. The on-track measurements included several vehicles, tested using 10 force measuring wheelsets in four European countries in a test train equipped with simultaneous recording of track irregularities and rail profiles. The simulations were performed using several vehicle models, built with the use of different simulation tools by different partners. The comparisons between simulation and measurement were conducted in a large number of simulations using a set of the same test sections. The results were assessed by three different validation approaches: by comparisons based on values according to EN 14363, by subjective engineering judgement and by using computable measures, so called validation metrics.

The proposed model validation criteria and limits are based on 12 quantities evaluated by analogy with EN 14363, covering the quasi-static and dynamic wheel/rail force measurements and vertical as well as lateral vehicle body accelerations. For each quantity, a set of at least 24 comparisons between simulation and measurement were evaluated using values based on EN 14363 from at least 12 sections which represent all 4 test zones according to EN 14363 from straight track to curves with very small radius. The agreement between simulation and measurement was assessed comparing the mean value and standard deviation for a set of differences between simulated and measured values of each quantity with the proposed validation limit.

The investigations did not confirm the traditional opinion about the positive effect of the model adjustments by comparisons with stationary tests on the simulation of on-track tests.

Comparisons between simulations and on-track measurements represent the only suitable and reliable model validation method in regard to the simulation of on-track tests. The proposed method, criteria and limits represent a suitable and robust methodology for the validation of railway vehicle models. It represents an overall assessment of a large number of data which is impossible to carry out by using engineering judgement, as it is not practically possible to display, check and document the approval of such a large number of plots. This validation process not only allows an objective assessment, but also supports an identification of the model weaknesses. The presented methodology is proposed to be implemented in to the revised standard EN 14363. Feedback from future applications of this method in common projects will help to further improve and develop the model validation which is the crucial condition for successful use of simulation to reduce the amount and cost of physical testing in railway vehicle acceptance.

Acknowledgement

This article describes work undertaken in the context of the DynoTRAIN project, Railway Vehicle Dynamics and Track Interactions: Total Regulatory Acceptance for the Interoperable Network (www.triotrain.eu). DynoTRAIN is a collaborative project – medium-scale focused research project supported by the European 7th Framework Programme, contract number: 234079 and is led by UNIFE.

References:

- [1] ASME V&V 10-2006, Guide for verification and validation in computational solid mechanics, New York, 2006.
- [2] Cooperrider NK and Law EH. A survey of rail vehicle testing for validation of theoretical dynamic analyses. *Journal of Dynamic Systems, Measurement and Control*, 100, December 1978; 238-251.
- [3] Gostling RJ and Cooperrider N. Validation of railway vehicle lateral dynamics models, *Vehicle System Dynamics* 1983; 12: 179-202.
- [4] Cheli F, Corradi R, Diana G and Facchinetti A. Validation of a numerical model for the simulation of tramcar vehicle dynamics by means of comparison with experimental data. *Journal of Computational and Nonlinear Dynamics* 2007; 2: 299-307.
- [5] Alfi S, Bruni S and Mazzola L. Numerical methodology for the evaluation of high speed vehicle stability in curved track. *Proceedings of the 8th International Conference on Railway Bogies and Running Gears BOGIE 2010*, Budapest, 13-16 September 2010, pp. 247-256.
- [6] Kuka N, Ariaudo C and Verardi R. Modelling and simulation of tilting trains. Paper No. 135, in *Proceedings of 22nd International Symposium on Dynamics of Vehicles on Roads and Tracks*, Manchester, August 14-19, 2011.
- [7] Evans J and Berg M. Challenges in simulation of rail vehicle dynamics, *Vehicle System Dynamics* 2009; 47: 1023-1048.
- [8] Bruni S, Vinolas J, Berg M, Polach O and Stichel S. Modelling of suspension components in rail vehicle dynamics context. *Vehicle System Dynamics* 2011; 49: 1021-1072.
- [9] GM/RC2641, Issue Two. Recommendations for vehicle static testing. RSSB, London, June 2009.
- [10] GM/RT2141, Issue 3. Resistance of Railway Vehicles to Derailment and Roll-Over. RSSB, London, June 2009.
- [11] EN 15273-2:2013 Railway applications - Gauges - Part 2: Rolling stock gauge, May 2013
- [12] 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.

- [13] prEN 14363:2013 Railway applications—Testing and simulation for the acceptance of running characteristics of railway vehicles—Running behaviour and stationary tests. CEN, Brussels, June 2013.
- [14] Jönsson LO, Nilstam N and Persson I. Using simulations for approval of railway vehicles: a comparison between measured and simulated track forces. *Vehicle System Dynamics*, Suppl. 2008; 46: 869–881.
- [15] Mazzola, L. The Influence of modelling of the suspension components on the virtual homologation of a railway vehicle. In: *Proceedings of the First International Conference on Railway Technology: Research, Development and Maintenance*, La Palma, April 2012. Paper No. 75. Civil-Comp Press, Stirlingshire, Scotland, 2012.
- [16] Mazzola L and Berg, M. Secondary suspension of railway vehicles -air spring modelling: Performance and critical issues. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 2014; 228: 225-241.
- [17] Zacher M and Kratochwille R. **WP1stationary and on track tests (?)**. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit (in this Special Issue)*
- [18] Polach O and Böttcher A. A new approach to define criteria for rail vehicle model validation, *Vehicle System Dynamics* (2014) – Latest articles, available at www.tandfonline.com.
- [19] 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.
- [20] EN 14363:2005 Railway applications—Testing for the acceptance of running characteristics of railway vehicles—Testing of running behaviour and stationary tests. CEN, Brussels, 2005.
- [21] Fries R, Walker R and 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, 2013.
- [22] Schwer LE. Validation metrics for response histories: perspectives and case studies, *Engineering with Computers* 2007; 23: 295–309.
- [23] Russell DM. Error measures for comparing transient data: Part I: Development of a comprehensive error measure, Part II: Error measures case study. In: *Proceedings of the 68th Shock and Vibration Symposium*, Hunt Valley, Maryland, 1997.
- [24] Sprague MA and 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.
- [25] Mongiardini M, Ray MH and Anghileri M. Development of software for the comparison of curves during the verification and validation of numerical models. In: *7th European LS-DYNA Conference*, Salzburg, Austria, 14th – 15th May, 2009. DYNAmore GmbH, 2009.
- [26] Polach O, Böttcher A, Vannucci D, Sima J, Schelle H, Chollet H, Götz G, Garcia Prada M, Nicklisch D, Mazzola L, Berg M and Osman M. Validation of multi-body models for simulations in authorisation of rail vehicles. In: *Proceedings of the 9th International Conference on Railway Bogies and Running Gears BOGIE '13*, Budapest, 9-12 September, 2013 (in print).
- [27] Evans J. Validation of vehicle dynamic modelling – some practical experience. In: *23rd International Symposium on Dynamics of Vehicles on Roads and Tracks*, August 19-23, 2013, Qingdao, China, Paper No. 2.3, ID483.