This is an Author's Accepted Manuscript of an article published in Vehicle System Dynamics, online since February 12, 2014, copyright Taylor & Francis

To cite this Article:
Polach, O. & Böttcher, A. (2014) A new approach to define criteria for rail vehicle model validation, Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility, 52:sup1, 125-141
To link to this Article: DOI: 10.1080/00423114.2014.881515
URL: http://dx.doi.org/10.1080/00423114.2014.881515

Vehicle System Dynamics, 2014
Vol. 52, Supplement, 125–141, http://dx.doi.org/10.1080/00423114.2014.881515

Taylor & Francis
Taylor & Francis Group

A new approach to define criteria for rail vehicle model validation

Oldrich Polach[a]* and Andreas Böttcher[b]

[a]*Bombardier Transportation (Switzerland) AG, Zürcherstrasse 39, CH-8401 Winterthur, Switzerland;* [b]*Alstom Transport Deutschland GmbH, Linke-Hofmann-Busch Str. 1, D-38239 Salzgitter, Germany*

(*Received 31 October 2013; accepted 4 January 2014*)

Validation of railway vehicle models is an important condition for the application of multi-body system simulations in the context of vehicle authorisation. Validation investigations carried out in the DynoTRAIN project represent a unique investigation. The measurements, received from a set of vehicles running through four European countries equipped with a simultaneous recording of track irregularities and rail profiles, were compared with simulations using vehicle models built in different simulation tools by several partners. The presented analyses resulted in a proposal of model validation based on 12 quantities 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 simulations and measurements are evaluated using values based on EN 14363 from at least 12 sections, which represent all 4 test zones according to EN 14363. The proposed method, criteria and limit values are intended for the validation of vehicle models applied in the vehicle acceptance process.

Keywords: validation; multi-body systems; railway vehicle; running dynamics; vehicle acceptance; authorisation

1. Introduction

Multi-body simulation tools are used in rolling stock design and development for several years to optimise the vehicle parameters and to conduct a wide range of investigations,[1] including the assessment of vehicle running behaviour and the prediction of test results. Computer simulations could be used to reduce the time and costs of testing for the acceptance of running characteristics of railway vehicles according to EN 14363:2005 [2] replacing a part of physical tests by 'virtual testing'. Applications of multi-body systems (MBSs) using vehicle dynamics simulations for vehicle acceptance purposes were introduced in UIC 518:2009 [3]; the recent draft of revision prEN 14363:2013 [4] also contains this option.

Good agreement between the behaviour of a real vehicle and its MBS model is the crucial requirement, when using MBS simulations. A model validation is used to approve this agreement by comparing simulation and measurement results.[5,6] The criteria applied to assess this agreement must consider that not only the vehicle model is not an exact representation of the reality because of the limits of general MBSs but also the measurement can deviate from the

*Corresponding author. Email: oldrich.polach@ch.transport.bombardier.com

© 2014 Taylor & Francis

reality due to the measurement errors and scatter of test conditions. The model validation in the context of vehicle acceptance should take into account not only the errors of the measurement of running dynamics quantities, but also that in the measurement of track layout and track irregularities, measurement of rail profiles and wheel profiles, scatter of the test conditions, e.g. friction coefficient between wheel and rail, as well as generally the stochastic character of the test results.

Unfortunately, no quantitative limits for a successful model validation are specified in [3,4]; an assessment by an independent reviewer is required instead. Publications on a common methodology for the validation of MBS vehicle models for simulations of running dynamics are rare too. The study by Jönsson et al. [7] can be mentioned as an example presenting a trial carried out during the preparation of UIC 518:2009.

This article presents a new approach to define measurable criteria and quantitative limits for the validation of railway vehicle models in the context of vehicle acceptance as investigated in Work Package 5 of the DynoTRAIN research project. This investigation is a unique approach analysing several types of vehicles. Most of the investigated vehicles were tested under the same conditions as in DynoTRAIN project. The simulations of these tests were carried out by different partners using different simulation software tools, compared to on-track measurements and assessed by project partners.

The article is structured as follows. Section 2 presents, briefly, the measurements carried out in DynoTRAIN and introduces the vehicle models as well as the selection of test sections and model configurations used for the investigations. Section 3 provides information about the assessments applied and validation approaches compared. These comparisons consist of the assessment based on quantities processed by analogy with EN 14363, the subjective engineering judgement as well as the so-called validation metrics intended to replace the subjective judgements by quantitative measures. Section 4 is dedicated to the evaluation of methods and criteria suited for model validation and the experience with the investigated assessments methods. The final proposal for the validation method, its criteria and limit values is described in Section 5, while Section 6 discusses advantages of the proposed method. Section 6 also presents the experience of how the usage of measured track data (track irregularities and rail profiles) affects the validation results. A summary and an outlook regarding further investigations are given in Section 7.

2. Validation exercises carried out in the DynoTRAIN project

The validation exercises used measurements carried out in DynoTRAIN Work Package 1 in October 2010. The test train with 4 types of tested vehicles was equipped with a total of 10 force measuring wheel sets and a number of other sensors recording over 300 measured signals. This train travelled for a total of 20 days of test runs through Germany, France, Italy and Switzerland reaching speeds up to 120 km/h with freight wagons connected and up to 200 km/h without the freight wagons. Measuring vehicles integrated in the test train continuously recorded the track irregularities and rail profile shapes during all test runs.

The following vehicle models were developed and assessed by comparisons with measurements:

- Locomotive DB BR 120 – modelled by Siemens in simulation tool Simpack.
- Locomotive DB BR 120 – modelled by IFSTTAR in simulation tool VOCO.
- DB passenger coach Bim – modelled by Bombardier Transportation in Simpack.
- DB passenger coach Bim – modelled by IFSTTAR in VOCO.

- Empty freight wagon Sgns with Y25 bogies – modelled by Technical University Berlin in Simpack.
- Empty freight wagon Sgns with Y25 bogies – modelled by IFSTTAR in VOCO.
- Laden freight wagon Sgns with Y25 bogies – modelled by Technical University Berlin using Simpack.
- Laas 2-axle flatbed wagon unit – modelled by Alstom in Simpack.

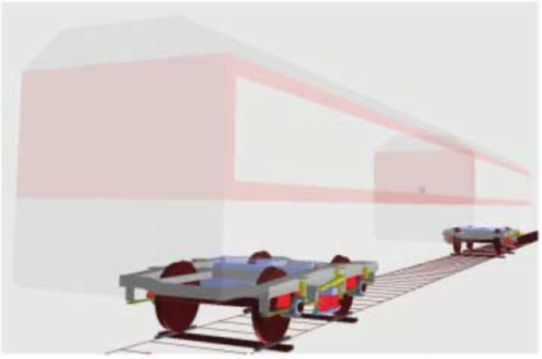
Furthermore, other models of two recently developed vehicle types were assessed using measurement results provided by the suppliers of the vehicles. A high-speed train delivered to Turkey by Construcciones y Auxiliar de Ferrocarrilles was modelled by this company using the simulation programme SIDIVE. The train DMU IC4 delivered by AnsaldoBreda to Denmark was modelled by the vehicle supplier in Simpack.

The vehicle model development was the responsibility of each partner using available documents, including the modelling and derivations of model input data as well as estimations of unknown parameters, such as parasitic stiffness of vehicle suspension. The vehicle models used in the investigations are three-dimensional fully nonlinear models as used today in the rolling stock engineering and applied research. The rigid bodies representing vehicle body, bogie frame, wheel set, 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 nonlinear wheel/rail contact models use Kalker's simplified theory (particular version of FASTSIM implemented in the applied simulation tool) with full Kalker's coefficients without reduction. The estimation of the friction coefficient between wheel and rail was the responsibility of each partner. The partners were advised to adjust the model mass parameters to achieve a good agreement between the static model wheel loads and the static wheel loads measured during the on-track tests, before starting the comparisons.

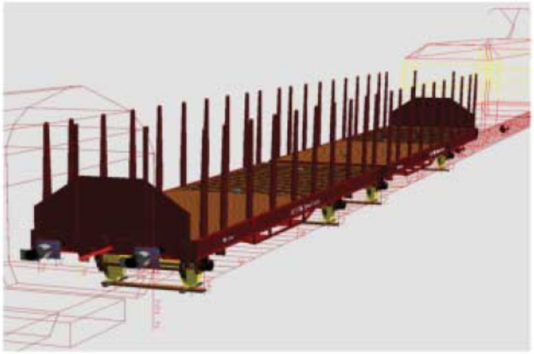
The comparisons between simulation and measurement were carried out for all vehicle models and model configurations under the same conditions over selected track sections of the test runs, called validation exercises; in this context the word 'section' does not mean a section according to the definition in EN 14363, but simply a part of track. One validation exercise consists either of a part of straight track or of one curve passing scenario, including both transitions and parts of straight track. The selected 17 validation exercises represented all 4 track zones according to EN 14363:2005 [2]: straight track and very large curves were represented by 5 sections, large radius curves ($R > 600\,\mathrm{m}$) by 2 sections; 4 sections were from small radius curves ($400\,\mathrm{m} \leq R \leq 600\,\mathrm{m}$) and 6 from very small radius curves ($250\,\mathrm{m} \leq R < 400\,\mathrm{m}$). They were from three countries: Germany (11), Italy (4) and Switzerland (2).

In order to assess the effect of using the actual measured infrastructure parameters, such as 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. In this context, the term 'estimated rail profile' means nominal, i.e. design rail profile and rail inclination of the particular country; similarly 'estimated wheel profile' represents design wheel profile S1002. The 'estimated track irregularity' data used by partners are either generated based on the power spectral density according to ORE B176 [8] or random track irregularities from other measurements. The selection of track irregularities to be used instead of the actual measured data was the responsibility of each partner.

The effect of using the results of stationary tests for the model validation with regard to the simulation of the on-track tests was investigated by comparing the simulation results using vehicle models before and after adjustments based on the comparisons with the stationary tests.



(a) Bim passenger coach by Bombardier



(b) Laas freight wagon unit by Alstom

Figure 1. Examples of vehicle models.

This article presents the investigations conducted by the whole project team and the proposed approach for the model validation. The validation results are illustrated with examples of two vehicle models, both developed in the MBS simulation tool Simpack:

- DB passenger coach Bim prepared by Bombardier Transportation, see Figure 1a. This vehicle is a single-deck passenger coach with two 2-axle bogies of Minden-Deutz design. The wheel set guidance is provided by leaf springs, and coil springs are used in the primary suspension. The secondary suspension is of swing bolster-type design with coil springs and friction damping between the bogie bolster and the car body. Yaw dampers are used to ensure the stability at speeds up to 200 km/h.
- Freight flatbed wagon Laas modelled by Alstom, see Figure 1b. This wagon unit consists of two 2-axle wagons with UIC double link suspension, connected by a short coupler to one vehicle unit. One wagon of this unit was tested in empty condition while the other one was laden. The simulation model represents the complete freight wagon unit, including the simplified models of neighbouring vehicles. The validation exercises concentrated on the empty wagon, which was equipped with force measuring wheel sets during the on-track test runs.

3. Validation assessments compared

3.1. *Assessments based on quantities processed by analogy with EN 14363*

The assessments by comparisons between the simulation and the measurement results contained:

- Assessment based on measured quantities, filtered and processed by analogy with EN 14363:2005.[2]
- Subjective engineering assessment using a simple 'Yes/No' method by project partners as well as during a project workshop with invited experts.
- Validation metrics, i.e. quantitative measures comparing simulation and measurement in the time histories with the aim to maintain agreement with engineering judgement.

The evaluation of quantities filtered and processed using the EN 14363 considers constant curvature track sections only, see Figure 2. The results used for comparisons consisted of 48 quantities per model configuration: quasi-static as well as dynamic wheel–rail forces and

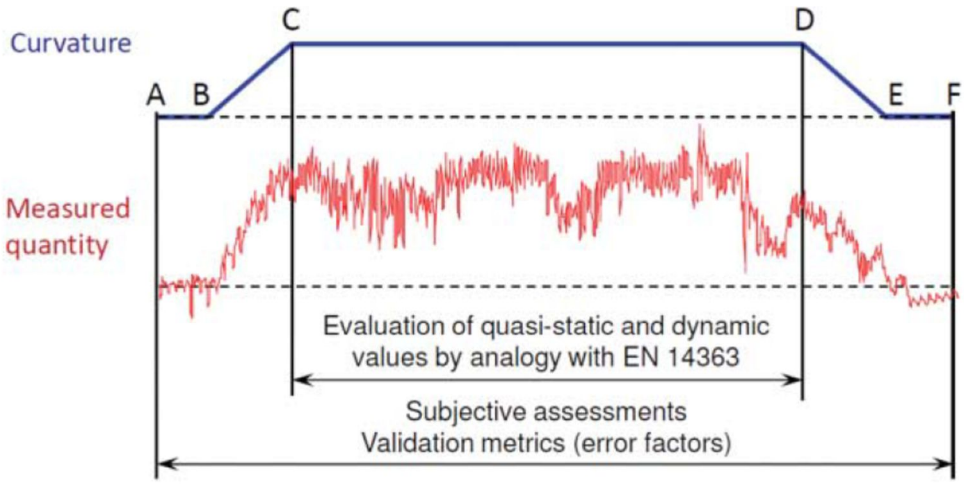


Figure 2. A validation exercise and its sections used for the assessments.

vehicle body as well as bogie frame accelerations. Moreover, 10 quantities based on power spectral densities of bogie and vehicle body accelerations were compared too.

3.2. *Subjective assessments by engineering judgement*

Comparisons between simulation results and measurements taking into account time-domain plots (time or distance diagrams) and power spectral density diagrams were carried out for a selection of vertical and lateral forces between wheel and rail, the derailment ratio Y/Q and for accelerations (20 plots per model configuration). The signals were post-processed by a low-pass filter of 20 Hz and compared for the full length of the validation exercise, i.e. including both curve transitions.

The project partners were asked to assess diagrams with comparisons of measurement and simulation signal quantities by using a simple 'Yes/No' method. Assessing a diagram with 'Yes' states that for the displayed signal quantity of the particular diagram the reviewer considers the model as validated. The complete simulation model would be confirmed as validated, if a large majority of diagrams are classified as validated. Overall, 16 time or distance plots and 4 power spectral density diagrams were provided for the subjective assessment of each model validation exercise.

Furthermore, a workshop was held in November 2012, during which 26 European running dynamics experts (professors for railway vehicle dynamics, experts from industry, railway companies, testing and research institutes, members of standardisation committees and DynoTRAIN project partners) assessed a selection of 120 diagram plots using the same Yes/No-method.

3.3. *Validation metrics*

As an alternative option to replace the subjective engineering judgement about the agreement between measurement and simulation, the application of so-called 'validation metrics' was investigated. Using these metrics, the curves of the simulation and measurement signals were compared with each other quantity by quantity in the time or distance domain by an integral approach, described by Magnitude, Phase and Comprehensive error factors proposed by Sprague and Geers.[9] By using the same sampling rate and length of time or distance interval for the compared measurement and simulation signals, the definitions of error factors proposed by Sprague and Geers in [9] can be expressed by the following formulae.[10]

Sprague and Geers Magnitude error factor:

$$M_{SG} = \sqrt{\frac{\sum_{i=1}^{n} c_i^2}{\sum_{i=1}^{n} m_i^2}} - 1, \tag{1}$$

where c_i are the simulated values and m_i the measured values.

Sprague and Geers Phase error factor:

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

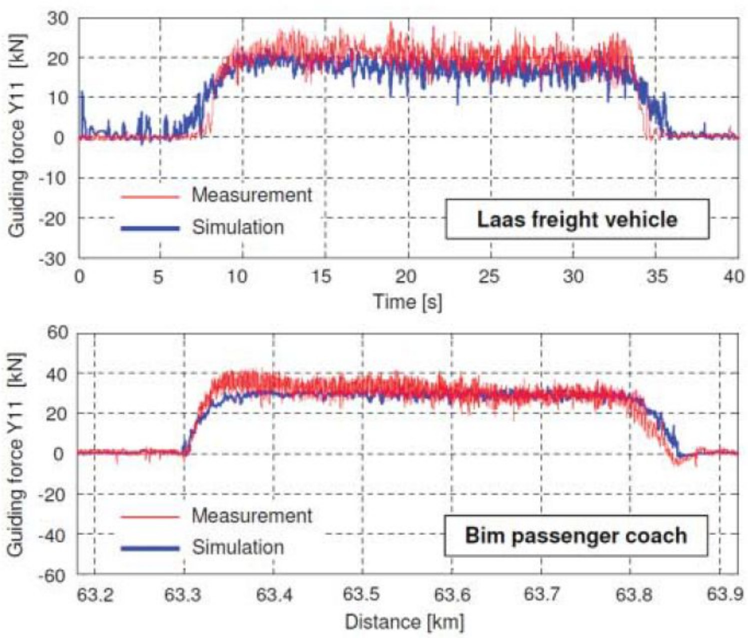
Sprague and Geers Comprehensive error factor:

$$C_{SG} = \sqrt{M_{SG}^2 + P_{SG}^2}. \tag{3}$$

Following the explanations of [11] and investigations carried out in the DynoTRAIN project, the better the agreement between simulation and measurement signals, the lower the error factor values shall be, i.e. error factor values near to zero demonstrate perfect agreement between simulation and measurement signals while high values represent disagreement. Since for vehicle homologation there are no applicable limits available to separate validated simulation models from non-validated ones, the evaluated validation error factors were combined with the partners' subjective assessments in order to check existing correlations for developing such validation limits.

3.4. *Examples of validation exercises*

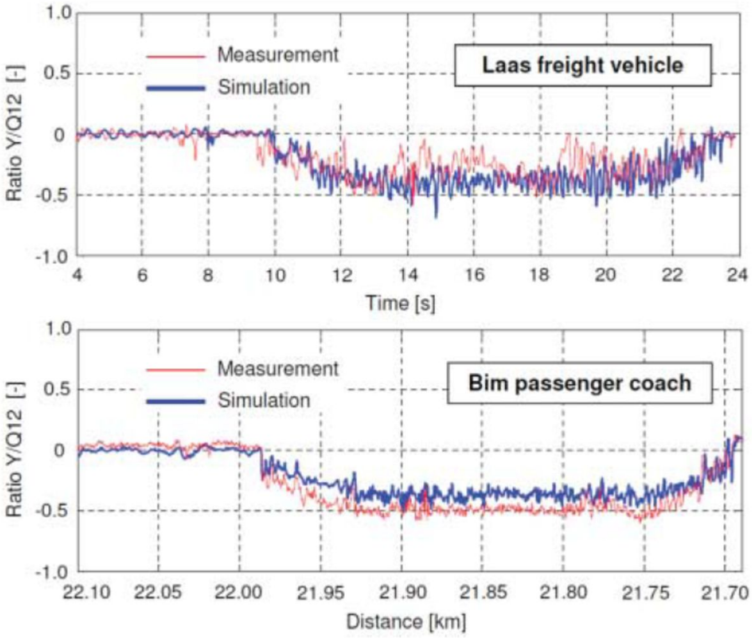
Examples of comparisons between simulation and measurement as well as assessment results for two tested vehicles – passenger coach Bim and freight wagon unit Laas – are shown in Figures 3–5. All three examples represent the test zone 4 (curves with very small radius).



Quasi-static guiding force Y_{qst}	Simulation	17.18 kN
	Measurement	20.55 kN
	Difference	-3.37 kN
	Rel. difference	-16.4%
Maximum guiding force Y_{max}	Simulation	21.87 kN
	Measurement	27.53 kN
	Difference	-5.66 kN
	Rel. difference	-20.6%
Positive assessment workshop		n.a.
Positive assessment partners		50%
Validation metric	M_{SG}	-0.135
	P_{SG}	0.074
	C_{SG}	0.155

Quasi-static guiding force Y_{qst}	Simulation	29.38 kN
	Measurement	31.01 kN
	Difference	-1.63 kN
	Rel. difference	-5.3%
Maximum guiding force Y_{max}	Simulation	33.05 kN
	Measurement	41.69 kN
	Difference	-8.64 kN
	Rel. difference	-20.7%
Positive assessment workshop		n.a.
Positive assessment partners		80%
Validation metric	M_{SG}	-0.072
	P_{SG}	0.057
	C_{SG}	0.092

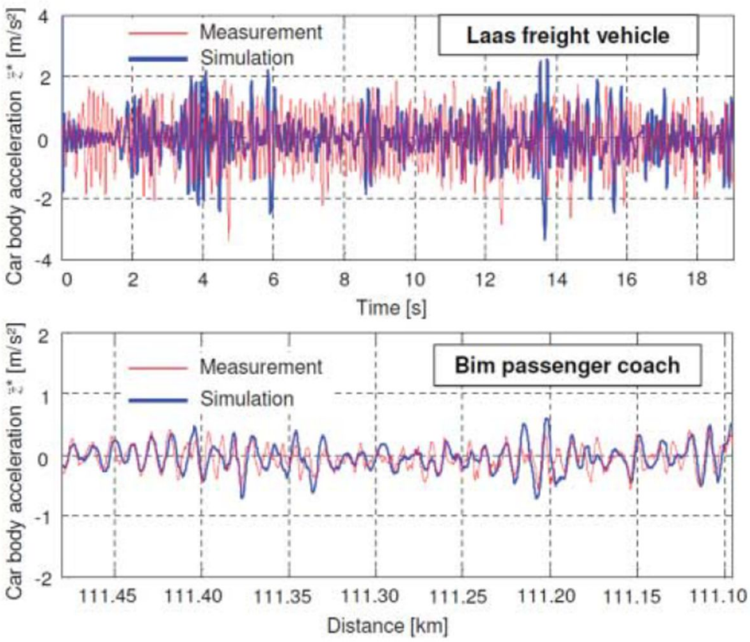
Figure 3. Validation examples Exercise F1.01, Germany, curve radius 282 m, speed 68 km/h.



Quasi-static derailment ratio Y/Q_{qst}	Simulation	0.39
	Measurement	0.27
	Difference	0.12
	Rel. difference	44.4%
Maximum derailment ratio Y/Q_{max}	Simulation	0.50
	Measurement	0.53
	Difference	-0.03
	Rel. difference	-5.7%
Positive assessment workshop		n.a
Positive assessment partners		20%
Validation metric	M_{SG}	0.249
	P_{SG}	0.150
	C_{SG}	0.291

Quasi-static derailment ratio Y/Q_{qst}	Simulation	0.37
	Measurement	0.49
	Difference	-0.12
	Rel. difference	-25.0%
Maximum derailment ratio Y/Q_{max}	Simulation	0.43
	Measurement	0.55
	Difference	-0.12
	Rel. difference	-22.3%
Positive assessment workshop		n.a.
Positive assessment partners		40%
Validation metric	M_{SG}	-0.257
	P_{SG}	0.061
	C_{SG}	0.265

Figure 4. Validation examples Exercise F1.12, Italy, curve radius 295 m, speed 76 km/h.



Car body acceleration, rms-value $\ddot{z}^*_{rms}$	Simulation	0.73 m/s²
	Measurement	0.83 m/s²
	Difference	-0.10 m/s²
	Rel. difference	-12.0%
Car body acceleration, maximum $\ddot{z}^*_{max}$	Simulation	2.87 m/s²
	Measurement	2.44 m/s²
	Difference	0.43 m/s²
	Rel. difference	17.6%
Positive assessment workshop		n.a.
Positive assessment partners		50%
Validation metric	M_{SG}	-0.138
	P_{SG}	0.494
	C_{SG}	0.513

Car body acceleration, rms-value $\ddot{z}^*_{rms}$	Simulation	0.23 m/s²
	Measurement	0.16 m/s²
	Difference	0.07 m/s²
	Rel. difference	40.1%
Car body acceleration, maximum $\ddot{z}^*_{max}$	Simulation	0.77 m/s²
	Measurement	0.48 m/s²
	Difference	0.29 m/s²
	Rel. difference	58.4%
Positive assessment workshop		100%
Positive assessment partners		80%
Validation metric	M_{SG}	0.117
	P_{SG}	0.528
	C_{SG}	0.541

Figure 5. Validation examples Exercise F1.15, Switzerland, curve radius 294 m, speed 74 km/h.

Figure 3 displays guiding force on the leading wheel from the validation section 1 in Germany. The diagrams show the time or distance plots, respectively. The tables on the right side present the values from simulation and measurement evaluated according to EN 14363 in the part of the section with constant curvature as well as their absolute and relative difference. Furthermore, the tables show the percentage of positive assessments during the workshop (if applicable) and the percentage of positive assessments by project partners. They also present the Magnitude, Phase and Comprehensive error factors of the validation metric by Sprague and Geers calculated for the displayed diagrams. Similarly, Figure 4 shows the ratio Y/Q in the validation section 12 (Italy) and Figure 5 the vertical vehicle body accelerations of the validation section 15 (Switzerland). The displayed results use actual measured wheel and rail profiles as well as measured track irregularities. The model configurations F1 represent the first development level of the vehicle models using available parameters but without any model adjustments by comparisons to stationary tests or to on-track test measurements.

4. Evaluation of validation criteria and limits

4.1. *Preliminary validation limits for quantities 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 was evaluated from the proposals provided by project partners. These proposals deviated significantly against each other as it is schematically shown in Figure 6 displaying the area satisfying the 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 characterises the deviation between simulation and measurement. A deviation acceptable for a successfully validated model is defined by the limit conditions displayed in Figure 6.

The following in principle differing definitions of the limit condition were proposed by the project partners:

- Deviation limit as a percentage of the measured value (relative deviation limit), Figure 6a.
- Deviation limit decreasing with the measured value increasing towards the limit for vehicle acceptance according to EN 14363 and constant for high measured values, Figure 6b.
- Constant deviation limit (absolute deviation limit), Figure 6c.

A reasonable justification can be provided for each of these differing proposals. Any deviation or error is usually considered with regard to the relative deviation hence supporting the first approach. However, as the vehicle is intended to be used for simulations of vehicle acceptance tests, it is important to achieve a good agreement for values, which are close to the limit values for vehicle acceptance, thus supporting the second, contradicting approach. Finally, it was agreed to use constant validation limit values for the deviation simulation–measurement, which is rather easy and at the same time the most appropriate compromise of the proposals discussed during the investigations.

A preliminary set of validation limits have been agreed based on the partners' proposals. These preliminary validation limits were then applied for the comparison of model configurations and for the investigation of a possible approach for validation. The effects of input data, like measured track irregularities, measured wheel and rail profiles, adjustment of model parameters by comparisons with stationary tests or differing depth of modelling, were compared and the quality of the model configuration was evaluated.

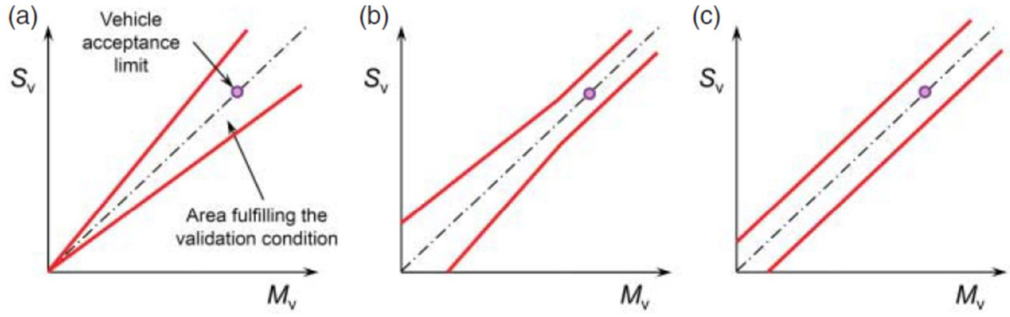


Figure 6. Principles of validation limit definitions proposed by project partners.

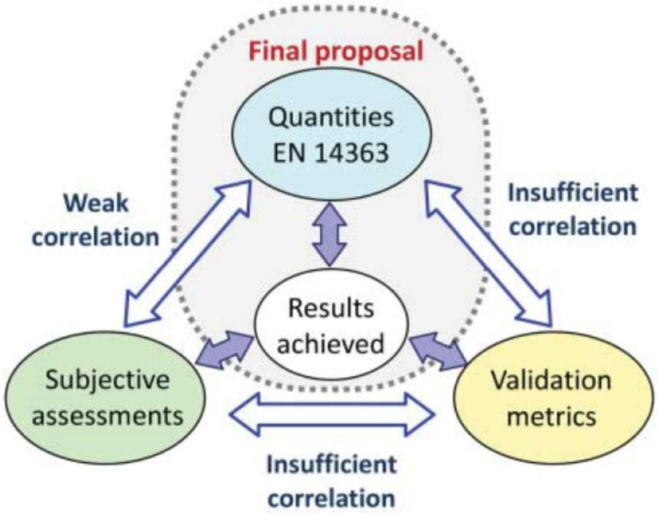


Figure 7. Correlation studies and investigations leading to the identification of the proposed validation limits.

4.2. *Evaluation of method and criteria suited for model validation*

Several 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 and more than 1000 simulations of validation exercises. The comparisons between simulation and measurement using values based on EN 14363 represent more than 50,000 single comparisons. The evaluation of time or distance as well as power spectral density diagrams comparing the simulation and measurement totalled more than 21,000 plots. About 6800 plots of selected model configurations were assessed by 7–10 project partners and 120 selected plots by 26 workshop attendees.

The correlations between the different groups of assessment (EN 14363 quantities, subjective assessments and validation metrics) as well as the relationship between the assessments and the achieved results were investigated, see Figure 7. They provided the following knowledge.

Subjective assessments by engineering judgement carried out by project partners as well as during the workshop demonstrated that these assessments vary significantly. Moreover, the presentation of the provided measurement and simulation results, such as different diagram axis scaling or change of front or back plane presentation of the signals, are influencing the reviewers' assessments.

Although some tendencies can be confirmed, the final assessment of each single comparison as well as the total assessment depends on the 'strictness' of each reviewer. The strictness deviation can be illustrated on the results from the workshop, in which a group of 26 workshop attendees assessed selected diagrams. The percentage of diagrams assessed positively by each reviewer varied significantly, see Table 1. The observed strictness variation was neither related to the set of plots nor to the reviewer's 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.

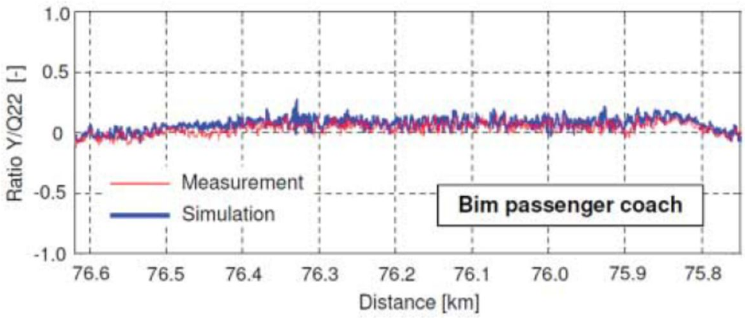
The investigations related to validation metrics were introduced with the intention of replacing subjective engineering judgement by quantitative, measurable criteria. This activity was motivated by information about the ongoing investigations on this topic carried out by TTCI in Colorado, USA, as reported in [12]. Unfortunately, based on the presented investigations, the

Table 1. Variation of the percentage of positive assessments in the workshop.

Set of plots	Number of plots	Percentage of positive assessments by a reviewer: Minimum	Maximum
Guiding forces Y	22	9.1	81.8
Vertical wheel forces Q	22	13.6	81.8
Ratios Y/Q	22	18.2	77.3
Bogie accelerations	22	10.0	95.5
Vehicle body accelerations	22	22.7	86.4
PSDs of bogie and body accelerations	10	0.0	100.0

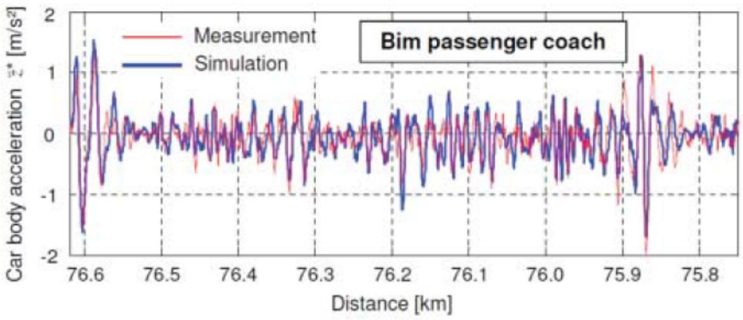
evaluated correlation between subjective assessments and validation metrics is not suited to establish limits distinguishing between validated and not validated simulation models. Deviations between simulation results and measurements are often neglected using engineering judgement, if these deviations occur at very small values close to zero or well within the limits for vehicle acceptance according to EN 14363. Since the validation metric error factors are based on a relative deviation, they do not regard effects like this, see Figure 8.

Another drawback of the validation metric is the strong influence on Phase error factor by the level of synchronisation between simulation and measurement signals, see Figure 9. An importance of the synchronisation of the compared signals is stressed in the investigations regarding the validation metrics presented in [13]. However, a perfect synchronisation is not easy to achieve and is usually not requested, which can lead to high values of the Phase as well as the Comprehensive error factors suggesting disagreement between simulation and measurement.



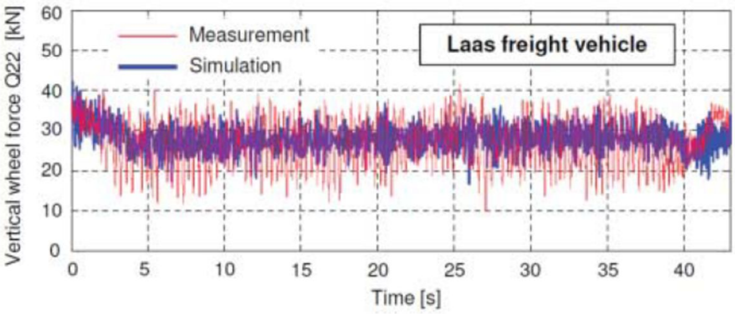
Quasi-static derailment ratio Y/Q_{qst}	Simulation	0.08
	Measurement	0.05
	Difference	0.031
	Rel. difference	60.1%
Maximum derailment ratio Y/Q_{max}	Simulation	0.17
	Measurement	0.13
	Difference	0.04
	Rel. difference	30.4%
Positive assessment workshop		n.a.
Positive assessment partners		100%
Validation metric	M_{SG}	0.447
	P_{SG}	0.246
	C_{SG}	0.510

Figure 8. Exercise F1.04, Germany, curve radius 580 m, speed 110 km/h: positive assessment by partners in contrast to high error factors.



Car body acceleration, rms-value $\ddot{z}^*_{rms}$	Simulation	0.31 m/s²
	Measurement	0.25 m/s²
	Difference	0.06 m/s²
	Rel. difference	24.0%
Car body acceleration, maximum $\ddot{z}^*_{max}$	Simulation	0.96 m/s²
	Measurement	0.75 m/s²
	Difference	0.21 m/s²
	Rel. difference	28.0%
Positive assessment workshop		n.a.
Positive assessment partners		100%
Validation metric	M_{SG}	0.078
	P_{SG}	0.789
	C_{SG}	0.793

Figure 9. Exercise F1.04, Germany, curve radius 580 m, speed 110 km/h: positive assessment by partners in contrast to high Phase and Comprehensive error factors.



Quasi-static wheel load Q_{qst}	Simulation	27.84 kN
	Measurement	26.87 kN
	Difference	0.97 kN
	Rel. difference	3.6%
Maximum wheel load Q_{max}	Simulation	36.02 kN
	Measurement	40.44 kN
	Difference	-4.42 kN
	Rel. difference	-10.9%
Positive assessment workshop		n.a.
Positive assessment partners		20%
Validation metric	M_{SG}	0.015
	P_{SG}	0.076
	C_{SG}	0.077

Figure 10. Exercise F1.13, Italy, curve radius 292 m, speed 76 km/h: rather negative assessment by partners in contrast to low error factors.

Moreover, combined deviations, such as the simulation signal's size of its quasi-static value combined with its dynamic amplitude deviation from the measured signal, can result in low validation metric error factors suggesting a good agreement (Figure 10), although a subjective acceptance for positive validation is small.

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 means for a more objective assessment. The validation metric, which was considered as suitable for the replacement of subjective assessment, does not show any valuable improvement and is thus not used in the proposal. Further investigations and modifications of the validation metrics would be required for its future applications in the context of the validation of railway vehicle models.

Moreover, the project investigations showed that assessments of single comparisons between simulation and measurement do not provide relevant information about the model quality with regard to validation. To approve the model validity, it is more important to assess an overall agreement between simulation and measurement, than to concentrate on differences in a few individual comparisons. An assessment of a large set of comparisons between simulation and measurement values evaluated by analogy with EN 14363 was finally selected as the best way to an objective and reliable model validation. The preliminary validation limits agreed in an earlier step of the project were used to assess the validation of the investigated model configurations. The feedback about the validated models was then used for the final adjustment of the validation limits; see schematic of this process in Figure 7.

From a total of 78 model configurations evaluated in this project, only 20 model configurations fulfilled the proposed model validation limits. The validated models were only those vehicle models, which were tested in the framework of DynoTRAIN WP1, where the measured track layout and track irregularity data were available and could be used for the model validation. The models successfully validated were only the models of locomotive DB BR 120 and DB passenger coach Bim. Neither a model configuration of the freight vehicles nor a model configuration without the usage of measured track irregularities could be validated. This result demonstrates the difficulties of modelling freight vehicles with large uncertainties of friction suspension parameters and stresses the importance of the actual measured infrastructure data for a successful model validation.

5. Proposed validation criteria and limits

The investigations resulted in a proposal of analysing 12 quantities covering the quasi-static and dynamic wheel–rail force measurements and vertical and lateral accelerations in the car body, filtered and processed by analogy with EN 14363, to be compared on at least 12 test

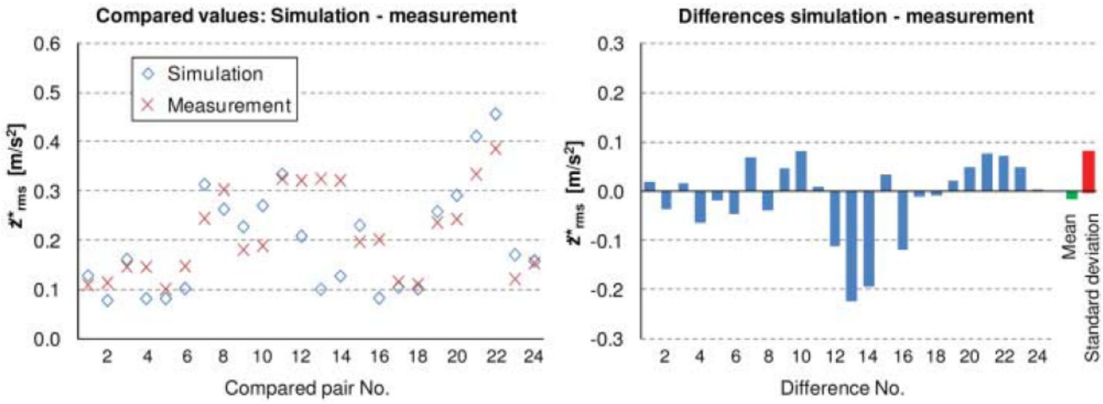


Figure 11. Example of evaluation of differences between simulated and measured values.

sections. A ‘section’ means either a test section according to EN 14363 or a part of test track longer than the minimum length specified for test sections in the particular test zone according to EN 14363. The selected validation exercises should contain sections from all 4 test zones, at least 3 sections from each test zone.

Each quantity is evaluated using at least two signals, e.g. the vertical accelerations above the leading and the trailing bogies, thus at least 24 simulated values can be compared with their measured counterparts for each quantity, see Figure 11. Each compared simulated as well as measured quantity is filtered and processed according to the requirements given within Table 2. The percentiles (0.15%-, 50%- and 99.85%-values) are calculated from the cumulative curve as described in EN 14363. The definitions of simulated values S_v and the corresponding measured values M_v are given in Table 2. 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%-value (median) is applied with its sign to approve the agreement of both magnitude and direction of the particular quantity.

The difference D_v between the simulated value S_v and the corresponding measured value M_v is evaluated for each value and quantity, whereby this difference is transformed depending on the sign of the measured value M_v so that, if the magnitude of the simulation value is higher than that of the measurement (simulation overestimating the measurement), the difference is positive, and vice versa:

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

$$D_v = S_v \quad \text{for } M_v = 0.$$

The following values shall be calculated for the whole set of differences D_v between the simulation and measurement for each quantity (e.g. for all Y_{qst} values) and compared with the validation limits:

- Mean value.
- Standard deviation.

Table 2 presents the proposed validation limits for the standard deviation of the differences simulation–measurement. The validation limits for the mean values are equal to two-thirds of the limits for the standard deviation. The validation limits for accelerations (mean as well as standard deviation) for freight vehicles or vehicles without secondary suspension are twice the relevant limit values stated in Table 2.

Table 2. Quantities and limits proposed for model validation with regard to the simulation of on-track test.

Quantity	Notation	Unit	Filtering	Processing	Validation limit for standard deviation of the differences between simulation and measurement
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)	$4\,(1+0.01Q_0)$, 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[a]	9
Vertical wheel force, maximum	Q_{max}	kN	Low-pass filter 20 Hz	99.85%-value[a]	$6\,(1+0.01Q_0)$, 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[a]	0.10
Sum of guiding forces, maximum	ΣY_{max}	kN	Sliding mean (2 m window, step 0.5 m)	0.15%/99.85%-value[a]	9
Car body lateral acceleration, rms value	$\ddot{y}^*_{rms}$	m/s^2	Band-pass filter 0.4–10 Hz	rms value	0.15
Car body vertical acceleration, rms value	$\ddot{z}^*_{rms}$	m/s^2	Band-pass filter 0.4–10 Hz	rms value	0.15
Car body lateral acceleration, maximum	$\ddot{y}^*_{max}$	m/s^2	Band-pass filter 0.4–10 Hz	0.15%/99.85%-value[a]	0.40
Car body vertical acceleration, maximum	$\ddot{z}^*_{max}$	m/s^2	Band-pass filter 0.4–10 Hz	0.15%/99.85%-value[a]	0.40

[a] Absolute values of simulated value S_v as well as measured value M_v.

6. Experience with the proposed model validation method

6.1. *Advantages of the proposed validation method*

The proposed validation method represents an overall assessment of a large number of data, which are not practical to carry out by using engineering judgement of the plots, as it would involve having to display, check and document the approval of such a large number of plots. This method allows a fast identification of quantities with the highest deviation; data of a particular quantity can be easily checked in detail to identify the validation exercise (section) and the signal (sensor position) that provides large deviations between the simulation and measurement. The specified set of 12 quantities to be evaluated covers the quasi-static as well as dynamic behaviour of the vehicle with regard to the vehicle acceptance, which is the intended range of the application for a validated model. The vehicle's safety relevant behaviour and the track loading results are validated by comparing quantities measured using force measuring wheel sets. The representation of vehicle ride is validated by comparing the root mean square (rms) values and maximum values of the car body accelerations.

The signal processing is carried out by analogy with EN 14363 for both the measurement and simulation, thus allowing the direct use of the acceptance tests data. Only the quantities Y_{max}, ΣY_{qst} and $(Y/Q)_{qst}$ proposed for model validation are not requested according to EN 14363:2005,[2] however; they are recommended for evaluation in recent draft prEN 14363:2013.[4] Thus, there are almost no additional requirements on measurement with regard to the proposed model validation.

The DynoTRAIN investigations confirmed that the proposed validation criteria and limits represent a suitable and robust methodology for the validation approval of railway vehicle models. A large number of compared pairs of values simulation–measurement ensure an approval of the overall vehicle's behaviour; it is difficult to improve the validation results by a simple modification of uncertain parameters or by a certain selection of validation sections. The evaluated values of the mean and the standard deviation of differences between simulation and measurement can be normalised by the proposed validation limits to visualise the weaknesses of the investigated models and to compare the quality of models against each other; see examples of such an assessment for the Bim passenger coach and the Laas freight wagon unit in the next section.

The quality of vehicle model can be improved by adjustments of uncertain or estimated input parameters. The identification of model parameters can be supported using comparisons of stationary tests measurements with simulations of those tests; however, this may be accompanied by some problems and pitfalls as reported by Evans.[14] Moreover, the project investigations showed that an improvement of the overall model quality regarding the simulations of on-track tests using comparisons with stationary tests is often marginal.[15]

6.2. *Experience regarding the effect of the usage of measured track data on the validation results*

The model configurations with measured track parameters and with estimated or nominal parameters, respectively, were evaluated using the proposed validation method and criteria in order to assess the importance of the usage of measured track data concerning a successful model validation. Figure 12 shows such an assessment for the Bim passenger coach and the Laas freight wagon unit. Three model configurations of each vehicle with differing knowledge of wheel and rail profiles are compared using non-dimensional values. The passenger vehicle Bim provides values with lower magnitude than the Laas freight vehicle. As the Bim coach's magnitudes of all values are below the validation limit, all three model configurations are considered as validated. The validation of the Bim coach model could thus be demonstrated, also without measured rail and wheel profiles. The Lass unit is not validated, because of it exceeding the Y_{qst}, $(Y/Q)_{qst}$ and $(Y/Q)_{max}$ and the vehicle body accelerations. In spite of the fact that with measured wheel and rail profiles the model configuration of the Laas unit provides better results than other Laas vehicle model configurations, the exceeding of validation limits occur for the same quantities. Thus, the Laas wagon unit cannot be confirmed as validated.

Figure 13 shows validation results without the actual measured track irregularities (i.e. using estimated track irregularity data) in comparison to the initial model configurations F1 with the actual measured track irregularities. The profiles of rails and wheels use the actual measured data in both cases. As expected, the comparisons show better agreement when the actual measured track irregularities are applied. But, it is interesting to see that the differences are smaller than one probably would expect. The model of the Bim coach fulfils the proposed validation limits using measured track irregularities and fails without the measured data, however, the exceeding of the validation limit is only

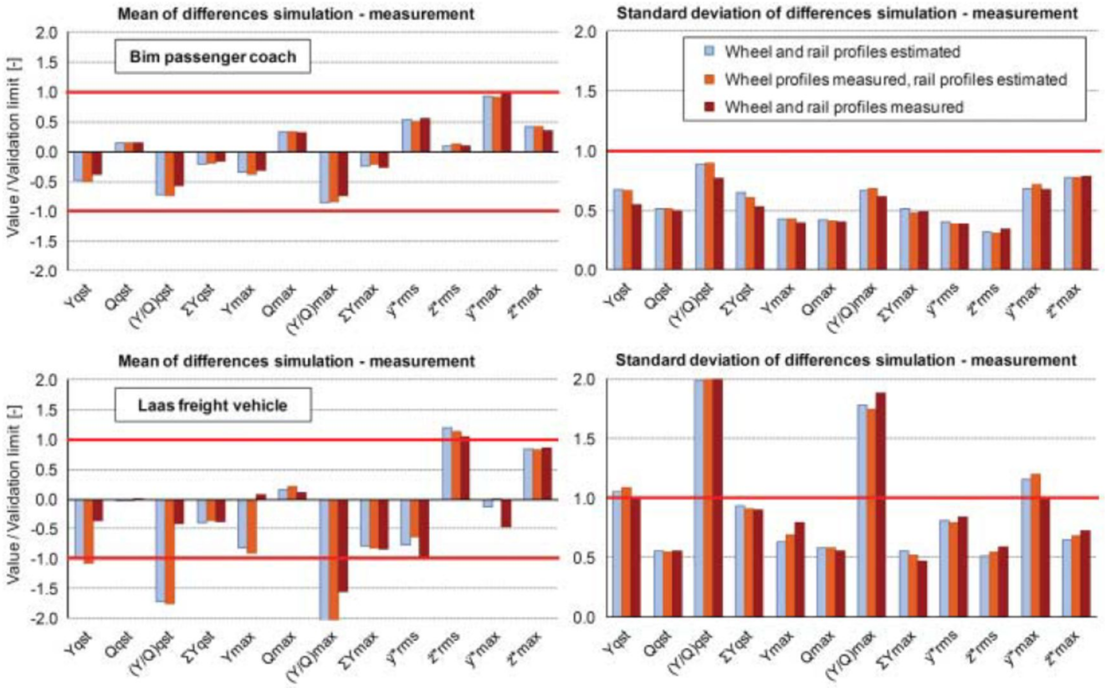


Figure 12. Normalised mean and standard deviation of differences between simulated and measured values for the validation of two vehicle models using actual measured wheel and rail profiles as opposed to estimated profiles.

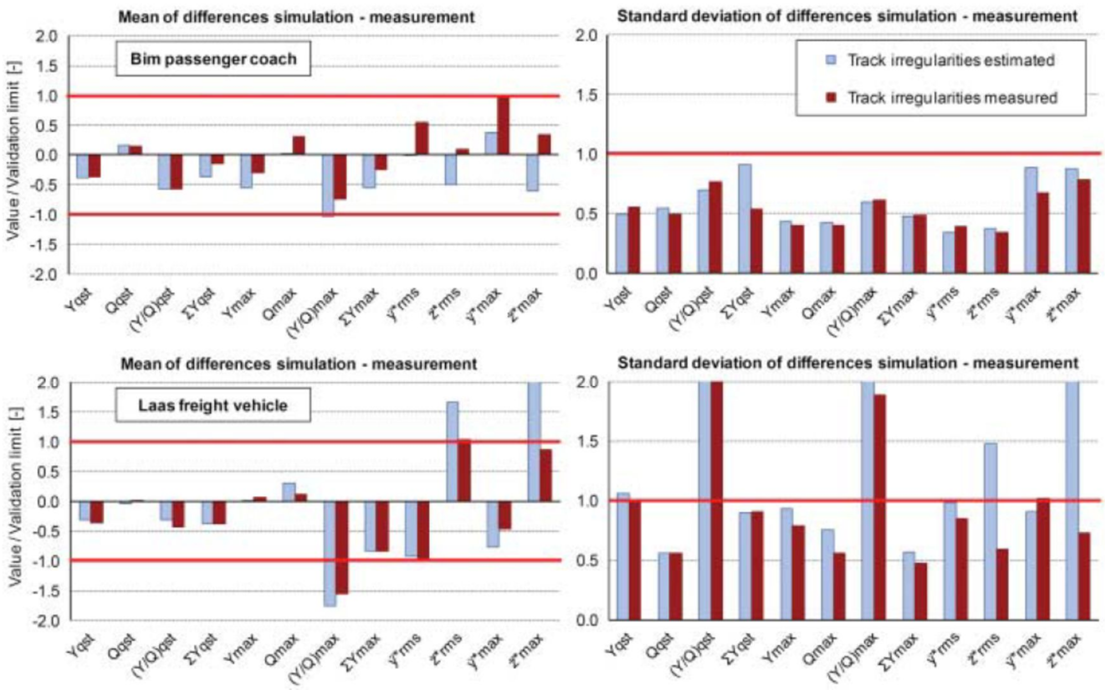


Figure 13. Normalised mean and standard deviation of differences between simulated and measured values for validation of two vehicle models using the actual measured track irregularities as opposed to random ones.

marginal. The Laas freight vehicle fails in both model configurations because of exceeding the Y_{qst}, $(Y/Q)_{qst}$ and $(Y/Q)_{max}$ and the vehicle body accelerations. As expected, larger deviations between simulations and measurements occur without the measured track irregularities.

7. Summary and outlook

This article presents investigations of the validation process, the criteria and the limits for the validation of MBS vehicle models with regard to simulations of on-track acceptance tests carried out in the DynoTRAIN project. These investigations represent unique work with regard to both simulations as well as measurements. The analyses are carried out using measurements with a test train consisting of several vehicles and using 10 force measuring wheel sets, running over 20 days through 4 European countries and being equipped with a simultaneous recording of track irregularities and rail profiles. The simulations, comparisons with measurements and evaluations were conducted using vehicle models built in different simulation tools by several partners.

The proposed criteria and validation limits are based on 12 quantities 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 are evaluated using values based on EN 14363 from at least 12 sections, which represent all 4 test zones according to EN 14363, from straight to curves with very small radius. It is intended to use these criteria for model validation in the context of vehicle acceptance according to EN 14363 and gain experience with this method in future projects.

The assessments using error factors of validation metrics are analysed too, but do not provide better and more reliable assessment than subjective assessments. This can be explained by the identified drawbacks of the validation metrics. Future investigations could remove these drawbacks by a modification of the validation metrics with regard to railway vehicle dynamic behaviour.

Acknowledgements

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 (http://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. The authors thank all partners involved in the presented investigations for their contribution in developing the proposal for the model validation process and criteria presented in this article.

References

[1] Polach O, Berg M, Iwnicki S. Chapter 12: Simulation. In: Iwnicki S, editor. Handbook of railway vehicle dynamics. London: Taylor & Francis; 2006:359–421.
[2] EN 14363:2005. Railway applications—testing for the acceptance of running characteristics of railway vehicles—testing of running behaviour and stationary tests. Brussels: CEN; 2005.
[3] UIC Code 518:2009. Testing and approval of railway vehicles from the point of view of their dynamic behaviour – safety – track fatigue – running behaviour. Paris: International Union of Railways; 2009.
[4] prEN 14363:2013. Railway applications—testing and simulation for the acceptance of running characteristics of railway vehicles—running behaviour and stationary tests. Brussels: CEN; 2013.
[5] Evans J, Berg M. Challenges in simulation of rail vehicle dynamics. Veh Syst Dyn. 2009;47:1023–1048.
[6] Bruni S, Vinolas J, Berg M, Polach O, Stichel S. Modelling of suspension components in rail vehicle dynamics context. Veh Syst Dyn. 2011;49:1021–1072.
[7] Jönsson L-O, Nilstam N, Persson I. Using simulations for approval of railway vehicles: a comparison between measured and simulated track forces. Veh Syst Dyn. 2008;46(Suppl.):869–881.
[8] 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. Utrecht: ORE; 1989.
[9] Sprague MA, Geers TL. A spectral-element method for modelling cavitation in transient fluid-structure interaction. Int J Numer Methods Eng. 2004;60:2467–2499.
[10] Mongiardini M, Ray MH, Anghileri M. Development of software for the comparison of curves during the verification and validation of numerical models. 7th European LS-DYNA Conference; 2009 May 14–15; Salzburg, Austria: DYNAmore GmbH; 2009.

[11] Schwer LE. Validation metrics for response histories: perspectives and case studies. Eng Comput. 2007;23:295–309.
[12] Fries R, Walker R, Wilson N. Validation of dynamic rail vehicle models. 23rd International Symposium on Dynamics of Vehicles on Roads and Tracks; 2013 August 19–23; Qingdao, China, Paper No. 10.1, ID416.
[13] Sarin H, Barbat S, Yang R-J, Kokkolaras M, Hulbert G, Papalambros, P. Comparing time histories for validation of simulation models: Error measures and metrics. J Dyn Syst Meas Contr. 2010;132:061401-1–061401-10.
[14] Evans J. Validation of vehicle dynamic modelling – some practical experience. 23rd International Symposium on Dynamics of Vehicles on Roads and Tracks; 2013 August 19–23; Qingdao, China, Paper No. 2.3, ID483.
[15] 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 multi-body models for simulations in authorisation of rail vehicles. Proceedings of the 9th International Conference on Railway Bogies and Running Gears BOGIE '13; 2013 September 9–12; Budapest.