OBJECTIVE RATING OF SIGNALS USING TEST AND SIMULATION RESPONSES

Christian Gehre
PDB – Partnership for Dummy Technology and Biomechanics
Germany
Heinrich Gades
VOLKSWAGEN AG
Germany
Philipp Wernicke
BMW Group
Germany

ABSTRACT

Today the numerical simulation is an inherent process of the development of the passive safety of vehicles. Robust and predictable computational models are the base of successful application of numerical simulations. The evaluation of the level of correlation of those models to the real world needs objective and reliable rating methods. In the past this rating was either done by engineering judgment or by analysing single peaks or zero-crossings of response curves in comparision with test data. Nowadays, it is common agreement that for an objective rating the complete curve data have to be taken into account.

In this paper, a new method is presented that provides an objective evaluation of whole response curves coming from test and simulation. The method combines two independent sub-methods, a corridor rating and a cross-correlation rating. The corridor rating evaluates the fitting of a response curve into user-defined or automatically calculated corridors. The cross-correlation method evaluates phase shift, shape and area below curves. It was found that the use of both of these two sub-methods is essential because the disadvantages of each sub-method are compensated by the other method. Both methods were implemented into a tool called CORA – correlation and analysis. The philosophy of this tool is to separate engineer's knowledge from the algorithms. External parameters to adjust the algorithms are representing this knowledge. So it is possible to tune the evaluation to the specific needs of the application.

The rating method was successfully used in a project on the improvement of Hybrid III 50th dummy models. It was possible to distinguish qualitatively and quantitatively between different releases of the model. In summary, the development of this rating method is a step forward to get an objective quality criterion of computational models.

In a next step the robustness of the rating will be analysed by varying the external parameters. Furthermore, the tool will also be used to analyse and evaluate results of physical tests.

INTRODUCTION

The analysis and comparison of signals coming from test and simulation is one of the major tasks of engineers working in vehicle safety business. A standardised method with reasonable scores is required to obtain an objective rating of the correlation of signals. So the requirements to a supplier regarding the level of validation of a computational model could be described more precisely in the future. The first step to an overall evaluation of the level of validation is the development of a metrics to compare a set of signals.

The evaluation of the correlation of two or more signals is not only relevant for the automotive industry. Hence, there are few rating tools on the market and even more published in the literature. For instance, the international standard ISO/TR9790 [1] uses a corridor method to compare cadaver and dummy responses for the biofidelity assessment of side impact dummies.

Each of the existing tools and algorithms has pros and cons but none seems to be both universal and adjustable to specific applications. Therefore, a new approach was developed to fulfil these requirements.

APPROACH

The existing rating algorithms suffer mainly from two reasons. Firstly, most of the algorithms are at least semi-universal but the settings, assumptions and simplifications made for implementation into software tools restrict the versatil-

ity. Secondly, many of the known algorithms are not robust enough to provide a reliable and applicable rating for good correlation as well as for poor correlation.

The new approach CORA avoids these issues. At first, the algorithms were consequently separated from the knowledge. Almost every parameter can be adjusted by the user. So it is possible to tune the algorithm to specific applications. Furthermore, completely different rating algorithms are integrated into the tool. Disadvantages of each algorithm are compensated by another algorithm. The user can decide about use and weighting of the different methods.

METHODS OF EVALUATION

CORA uses two different methods to calculate and assess the correlation of signals. While the corridor method calculates the deviation between the curves with the help of user-definied or automatically generated corridors, the cross correlation method analyses specific curve characteristics via parameters like phase shift or shape of the signals. The rating results ranges from "0" (no correlation) to "1" (perfect match).

The influence of the methods on the global rating is adjusted by user-defined weighting factors. Figure 1 shows the structure of the rating scheme in principle.
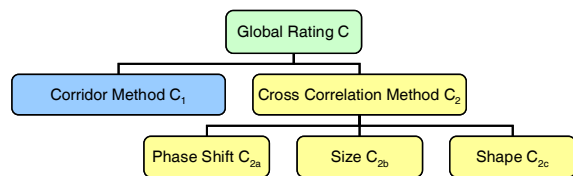

Figure 1.   Interaction of the sub-methods.

The equations (1), (2) and (3) are showing the calculation of the global rating of a signal by using weighting factors for each sub-method.

$$C = w_1 \cdot C_1 + w_2 \cdot C_2 \qquad (1).$$

$$C_2 = w_{2a} \cdot C_{2a} + w_{2b} \cdot C_{2b} + w_{2c} \cdot C_{2c} \qquad (2).$$

$$1 = \sum w_i \qquad (3).$$

Other algorithms are analysing the global peak and the timing of peak. These parameters are captured by the metrics of CORA indirectly and can be controlled by a reduction of the corridor width around the peak. The direct

implementation of these features into CORA is not intended, because it would somehow overrate the evaluation of signal parts. It is emphasised that the focus of CORA is on the assessment of the overall correlation of signals and not on the evaluation of single characteristics.

Filtering of the signals is a crucial point of almost every method. Noisy signals are difficult to analyse. Especially mathematical methods like the cross correlation method may fail to calculate the right level of correlation. Therefore, the user has to filter the signals prior the start of the rating tool.

Corridor method

The corridor method calculates the deviation between two signals by means of corridor fitting.

At first the mean curve of the references signals, usually coming from tests, is calculated. Two corridors, the inner and the outer corridor, are defined along the mean curve. If the evaluated curve is within the inner bounds, a score of "1" is given. The assessment declines from "1" to "0" between the bounds of inner and outer corridor. This transition is user-defined. Usually it is linear, quadratic or cubical. The three different rating zones are shown in Figure 2.

The compliance with the corridors is calculated for each time step. The final rating of a signal is the average of all single time step ratings.
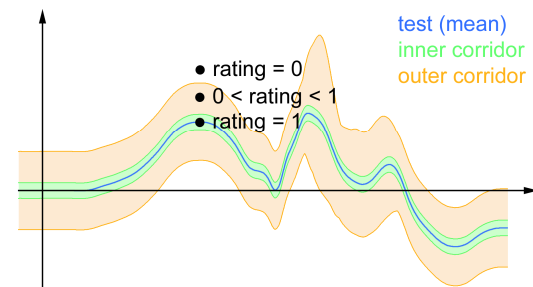

Figure 2.   Corridor method using corridors of constant width.

The philosophy is to use a narrow inner corridor and a wide outer corridor. It limits the number of "1" ratings to only good correlations and gives the opportunity to distinguish between poor and fair correlations. If the outer corridor is too narrow, too many curves of a fair or moderate correlation would get the same poor rating of "0" like signals of almost no correlation with the reference. The width of the corridors can be adjusted in order to reflect the specific signal characteristic.

One of the advantages of the corridor method is the simplicity and the clearness of the algorithm.

It reflects criteria which are used intuitively in engineering judgement. Especially the rating results are easy to understand because the method rates the hits of the rating zones. The simplicity is at the same time the disadvantage of the method. In case of little curve coincidence it may lead to an inadequate rating. As shown in Figure 12, the simulation curve scores 0.431 (corridor rating only) without any correlation to the reference signal. The scores are gathered by hitting the inner and the outer corridor arbitrarily. Therefore, a second method is required to counterbalance this disadvantage. Typically, the corridor method gives realistic results above ratings of approximately 0.5.

Different approaches to define the width of the corridors are implemented. The most common option is the use of constant corridor widths (Figure 2). Typically a share of the global absolute maximum is used as width.

The width can also be calculated by using the root mean square deviation (sigma). Sigma is added to baseline corridors of constant width. So it is possible to consider the scatter of the reference signals. If the balance between constant baseline width and variable sigma term is inadequate, the corridors could become either very small or large (Figure 3). The signals shown in Figure 2 and Figure 3 used the same data to calculate the corridors.

Finally, it is also possible to integrate user-defined corridors.
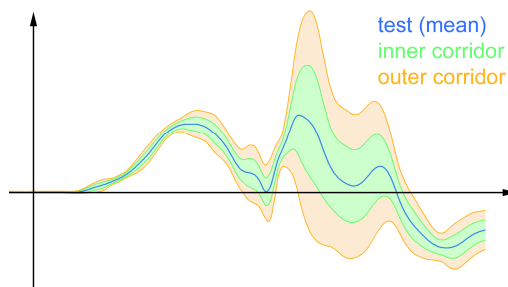

Figure 3. Corridor method using sigma-based corridor width.

Cross correlation method

The cross correlation method avoids the disadvantages of the corridor method by analysing the characteristics of signals. Three sub-methods with individual weightings factors are implemented. Similar to the corridor method, the mean curve of the reference signals is taken as base for the evaluation.

Firstly, the time shift of a signal to its reference is analysed (Figure 4). The maximum range of shifting is limited to avoid confusions of the

algorithms in case of sinus-like signals. The size and progression of a signal is evaluated after adjustment of the phase error.
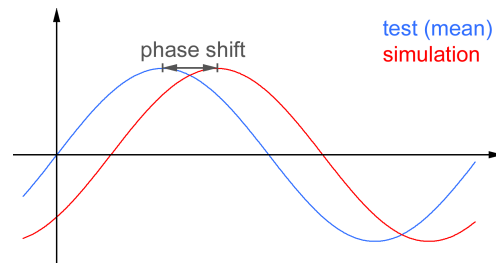

Figure 4. Evaluation of the phase shift.

Secondly, the size of the signals is analysed by comparing the area below the two curves globally. It is a helpful evaluation but is not sufficient to evaluate the real level of correlation. For instance, the area below a signal with high and narrow peak could be identical to the area of a curve with low but wide peak. The size method would evaluate this example with "1" although the shape of the signals is completely different.

Thirdly, the progression of the signal is calculated by means of the cross correlation function. This rating can be considered as a quantitative assessment of the shape of a signal.

The transition between ratings from "1" to "0" can be adjusted in all three sub-methods. Usually it is either linear or quadratic. Higher degrees are possible.

Compared to the corridor method, the cross correlation algorithms evaluate the level of correlation analytically. This kind of analysis is quite complex. Hence, the acceptance by users might be less than of the corridor method. The split of the cross correlation rating into three less complex sub-ratings improves the clearness and acceptance.

As shown in Figure 11, the cross correlation method is sensitive to noisy signals. The rating could be wrong if the noise interferes significantly the correct analysis of a signal.

Interval of evaluation

Most of the tools available on the market are analysing whole curves. The recording time of crash signals and simulation runs is typically slightly longer than required. So it would not be correct to simply use the whole signal for correlation analysis. Non-relevant parts of a signal may improve or worsen the rating. To avoid this, the response curve has to be cut properly before starting the rating. Figure 5 shows exemplarily the problem. The score of the

signal improves from 0.611 to 0.822 (approx. +35%) in this example when extending the interval of evaluation from the red tagged area to the whole curve, tagged in yellow. The differences may increase further if there are oscillations or peaks in the non-relevant parts of the signal.
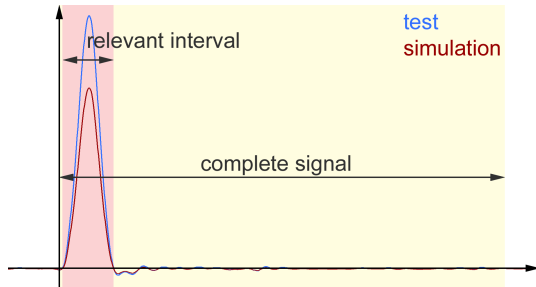

Figure 5.   Interval of evaluation.

Therefore, it is essential to restrict the interval of evaluation to the relevant part of a signal. Two options are implemented. Firstly, the user can define this interval for every signal manually and secondly, the software calculates this interval for each signal individually.

To define the start of the interval, the analysis starts from the first recorded value of the mean reference signal by forward scanning of the signal along the timeline. If the signal exceeds a pre-defined threshold, the start of the interval of evaluation is set. There are additional parameters available to modify the starting point slightly. This mechanism is very reliable.

Very similar to that, the signal is scanned backwards along the timeline to define the end of the interval of evaluation. The end is set if the signal falls below a given threshold.
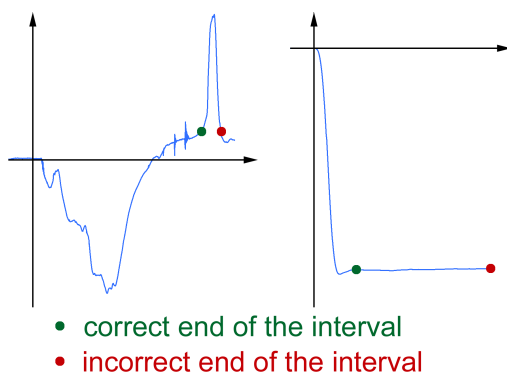

- correct end of the interval
- incorrect end of the interval

Figure 6.   Critical signals for automatically definition of the interval of evaluation.

Unfortunately, this mechanism does not work properly for all types of response curves. Signals which end at an almost constant high level (e.g. delta-v curves) or with a secondary impact (e.g.

peak of a rebound) are not treated in the right way (Figure 6).

Therefore, an additional method is introduced to handle signal endings of delta-v-like curves. At first, the end of the interval is defined by using the standard mechanism. Afterwards, a narrow corridor is defined for the remaining curve. The new end of the interval is set to that point when the signal leaves the corridor. This mechanism works fairly well but needs to be improved. This corridor is not linked to the corridors of the corridor method.

The second type of crucial signals cannot be treated right for the moment. The user has to set the end of interval of evaluation manually to exclude secondary impacts from the rating.

APPLICATION

CORA is a command line tool for several computer platforms that realises the developed approaches. It is solely a demonstrator of the proposed algorithms. There is no intension to use CORA commercially. A description of the methods used as well as the software itself will be available to the research community.

Program flow

Figure 7 shows the flow of the rating process in principle. At first the reference data set (input data 1) is defined. The data can be obtained from test or simulation. If more than one curve per channel is defined, the software calculates a mean curve because the rating algorithms are only able to compare and assess two signals.

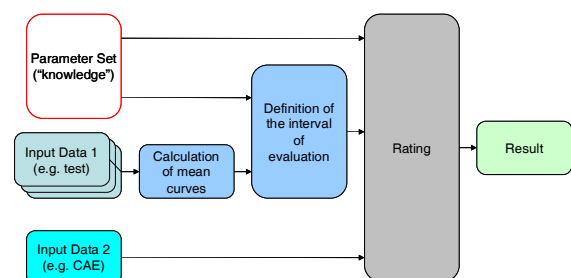Afterwards, the interval of evaluation is defined by using the information provided with the parameter set.

Figure 7.   Process of preparation and evaluation of the data.

In a next step, the software imports the second data set which has to be compared with the reference data. The information of setting up the rating algorithms is taken again from the parameter file. Additionally, the validity of the interval of evaluation is checked. If the recording

time of input data 2 is shorter, the interval is adjusted automatically.

The results of the rating are stored in several formats. A plain text file contains a brief summary of the rating. All control parameters, weighting factors, calculated parameters (e.g. interval of evaluation) as well as the rating results are included in this file. It can be processed by other software like spreadsheet programs. Furthermore, a detailed HTML report is provided with graphs of the signals. The calculated mean curves and corridors are also exported. They can be processed by any CAE postprocessor.

Hierarchy of the data

CORA provides not only a rating of a single pair of curves. The rating of signals can also be extended to sub-ratings (sub-load case) and global ratings. All sub-load cases are merged to a load case rating. Finally, the rating of the load cases is combined to the global rating (Figure 8).

This four level structure was mainly introduced to cover the requirements of the ISO/TR9790 biofidelity rating. The sub-load cases contain all biofidelity tests of a body region such as head, neck or plevis. The combination of all load cases (body regions) is the total biofidelity rating of a dummy. Each signal, sub-load case as well as load case has individual weighting factors. So the influence of unimportant signals or tests on the rating of the higher level can be reduced.

The intruduced hierachy is also applicable for any other application.
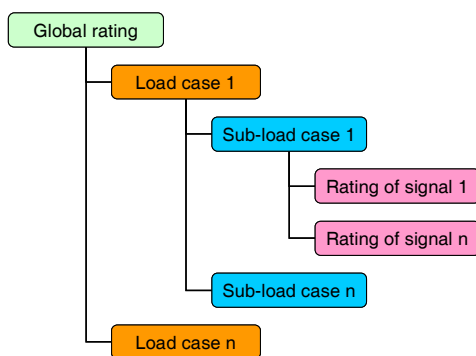

Figure 8.   Hierachy of the data.

Biofidelity rating according ISO/TR9790

The biofidelity rating according the ISO standard was implemented in the latest version of the tool. A pre-defined template of the control parameter set is provided with the software to simplify the rating process. The user has to define the source of the test data only. All biofidelity corridors and weighting factors are already provided with the template.

Examples

The pros and cons of the described rating method are explained in the following examples.

Figure 9 shows the chest deflection of a Hybrid III 50th dummy and of two different models of it. Simulation 2 correlates better with the test curve. The corridor rating is 0.648 and the cross correlation rating is 0.681 (total rating of 0.665). The response of simulation 1 has got a corridor rating of 0.400 and a cross correlation rating of 0.507 (total rating of 0.454). All results correspond with subjective expectations on the rating.
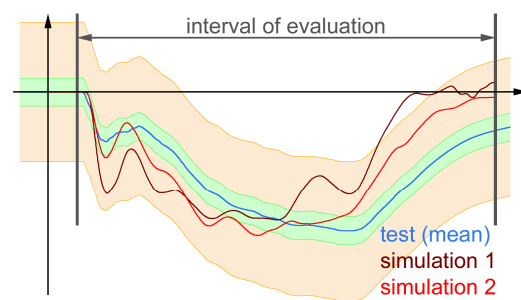

Figure 9.   Evaluation of chest deflection, constant corridor width.

The same signals are shown in Figure 10. Solely the calculation of the corridors was changed. Inner and outer corridors are using the standard deviation to calculate the corridor width. Furthermore, the width of the corridors is significantly reduced. The rating using the corridor method drops from 0.648 to 0.251 (simulation 1) and from 0.400 to 0.168 (simulation 2). The cross correlation rating remains unchanged.
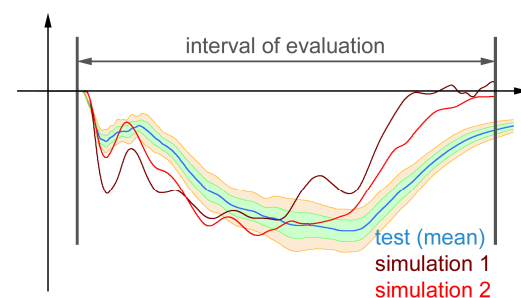

Figure 10. Evaluation of chest deflection, variable corridor width.

This scoring does not reflect subjective expectations of users. Due to the improper adjustment of the outer corridor, which is actually too narrow, the signals are very often outside of this corridor and obtain scores of "0".

Furthermore, the absolute difference between both signals is reduced and the rating does not reflect the real level of correlation to the reference signal.

Another example demonstrates the influence of filtering on the rating. Figure 11 shows the neck tension force of a Hybrid III 50th dummy in a frontal test. The signals simulation 1 and simulation 2 are taken from the same simulation run but using different filter classes.
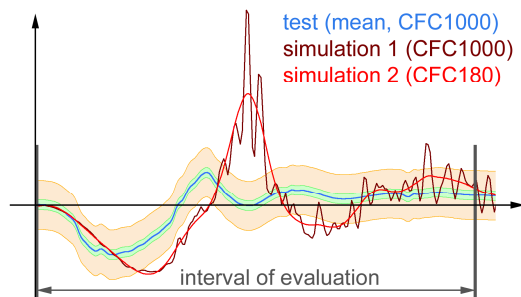

Figure 11. Evaluation of neck tension forces, influence of filtering on the rating.

The corridor rating is not sensitive to the applied filter class in this example. The difference of the rating is less than 0.005 (relative change of approx. 2%). Compared to that, the sub-ratings of the cross correlation method improve up to 0.07 (relative change of 14%) if the signal is filtered with CFC180. The algorithms of the cross correlation method are sensitive to noise. The spikes makes it is difficult to identify the right phase shift and progression.

The total rating of the simulation improves from 0.366 (simulation 1) to 0.389 (simulation 2) by using CFC180 instead of CFC1000. However, ratings should not be improved by applying higher filter classes. If there are noisy signals in the computational model, the cause of the noise should be fixed first.

The last example points out the limitations of rating methods. Figure 12 shows the knee displacement (knee slider) of a Hybrid III 50th measured in a frontal sled test without knee contact. The signal is caused by the inertia of the knee slider only. The maximum displacement is less than 2 mm. So the signal might not be relevant when evaluating a model in this specific load case.
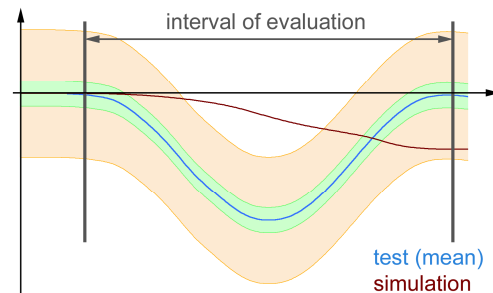

Figure 12. Evaluation of knee slider, limitations of the rating.

The total rating of the simulation curve is 0.391 (0.431 corridor method, 0.351 cross correlation method). This rating does not correlate to any subjective rating of users because there is no correlation between test and simulation. The rating of the previous example (Figure 11) is similar but the correlation between test and simulation is clearly better.

The disproportion is not necessarily a problem of the applied rating methods and their control parameters. The user has to define weighting factors for each channel. So the rating of the knee slider signal (Figure 12) might be acceptable if the influence on the global rating is smaller than that of the neck force (Figure 11). The definition of those weighting factors cannot be handled with a rating tool automatically.

LIMITATIONS

The development of the tool and the methods is not completed yet. The current version offers the opportunity to compare and evaluate signals coming from any source and groups single ratings to combined ratings. It is a first step to the objective rating of computational models.

Limitations of the algorithms

The introduced mechanism to define the interval of evaluation needs to be improved. There are certain types of signals that treated not correctly. So the user has to ensure manually that the right interval is defined. Otherwise the rating would not meet the expectations.

Furthermore, the algorithms are not able to handle signals with hysteresis. So it is impossible to evaluate force-deflection characteristics. The user has to evaluate the time history signals of force and deflection separately.

Classification of the rating

One of the uncertainties of the approach is the global meaning of the calculated scores. In case of using the same parameter set it is possible to distinguish between different variants of a model. Most of the results correlate with the subjective rating of users.

Changes of the parameter set do not change the ranking totally. The distances between models may vary and also the ranking of very similar signals may change. However, it is impossible to get the best rating for the worst model and the other way round just by modifying the control parameters.

The global meaning of the score has to be clarified in the future. There is a reference needed to be able to assign 0.6, 0.8 or any other score to a certain level of validation.

Evaluation of models

The tool offers the opportunity to evaluate whole models including different load cases and signals. Every signal and load case can be assigned with specific weighting factors. However, some general decision must be made before starting this global rating. At first, relevant load cases for the evaluation of a model have to be selected and prioritised. Afterwards, weighting factors for every signal of the load cases must be defined. This has to be done by the user. The CORA tool is not able to do this or to recommend baseline settings. Furthermore, the user has to adjust the control parameters of the algorithms. Only these settings can be derived from previous evaluations.

The difficulties of the selection and right weighting of load cases and signals are exemplarily explained on a dummy model.

Usually, there are extensive databases of certification tests available. The validation of many dummy models started with these data. It is assumend that the performance of the computational model in these tests is very well. A rating tool would probably calculate high scores.

However, a dummy model is typically used in vehicle environments. The load paths as well as the load levels may differ significantly to the conditions of certification tests. It is not guaranteed that the performance of the model in vehicle load cases is similar to that in certification tests. Therefore, the scoring in certification tests is probably not relevant in vehicle tests.

A global rating of a model must include all relevant loading conditions including the right balance between them (weighting factors).

The selection and evaluation of the right signals of a load case is sensitive too. There are important and less important signals measured with a dummy. Especially the handling of secondary axes has to be considered (e.g. transverse acceleration in a frontal crash). Firstly, these axes could be treated like the main axes. Acceptable deviation would become then a disproportional influence on the final rating. Secondly, they could also be ignored by using small weighting factors but then any problem with one of these axes would probably not be noticed in the global rating. Thus, there is an optional mechanism in the CORA tool that offers a special treatment of secondary axes. The main parameters, such as width of the corridors are taken from the main axis. So the rating of secondary axes is more tolerant but relevant deviations are reflected in the global rating.

However, more investigation is needed to extend the rating of single signals to a global rating of computational models.

SUMMARY

A new approach of the evaluation of the correlation of signals coming from test or simulation was developed. The algorithms and its control parameters are separated. That offers the opportunity to adjust the rating scheme to specific needs of each application and the requirements of users.

In total four methods are implemented to counterbalance the disadvantages of each sub-method. The influence of each method on the global rating can be influenced by control parameters.

Besides these features, there is an algorithm implemented that calculates the interval of evaluation automatically. It ensures that only the relevant parts of a signal are considered for the rating. All introduced automatisms can be switched off and replaced by user-defined settings.

The latest feature of CORA is the integration of the biofidelity rating of side impact dummies according ISO/TR9790.

The development of the software is not completed yet. More investigation is needed to define a valid and robust parameter set. This set could become the baseline or starting parameter set. Additionally, more investigation is needed to improve the implemented automatisms like the calculation of the interval of evaluations.

The software and the description of the methods used will be available for the research community. The most likely option is the publication as freeware.

OUTLOOK

As mentioned above the development is not completed yet. It is intended to use this approach as part of the evaluation of computational models in the future.

Furthermore, the work of ISO TC22 SC10/12 Working Group 4 is supported. This group is working on international standards to evaluate the correlation of response curves with the focus on the comparison of test and simulation.

REFERENCES

[1]   International Organization for Standardization. Technical Report 9790; Road Vehicles; Anthropomorphic side impact dummy – lateral impact response requirements to assess the biofidelity of the dummy; Geneva; 1999.