Quality comparison of motion cueing algorithms for urban driving simulations

Kolff, Maurice; Venrooij, Joost; Schwienbacher, Markus; Pool, Daan M.; Mulder, Max

Publication date
2021

Document Version
Final published version

Published in
Driving Simulation Proceedings

Citation (APA)
Kolff, M., Venrooij, J., Schwienbacher, M., Pool, D. M., & Mulder, M. (2021). Quality comparison of motion cueing algorithms for urban driving simulations. In A. Kemeny (Ed.), *Driving Simulation Proceedings*

Quality Comparison of Motion Cueing Algorithms for Urban Driving Simulations

Maurice Kolff[1,2], Joost Venrooij[1], Markus Schwienbacher[1], Daan M. Pool[2] and Max Mulder[2]

(1) BMW Group, Research, New Technologies, Innovations, 80788 Munich, Germany, e-mail : {Maurice.Kolff, Joost.Venrooij, Markus.Schwienbacher}@bmw.de
(2) Delft University of Technology, Faculty of Aerospace Engineering, section of Control & Simulation, 2629 HS Delft, The Netherlands, e-mail : {m.j.c.kolff, d.m.pool, m.mulder}@tudelft.nl

Abstract - *When designing driving simulation experiments with motion cueing, it is often necessary to make choices between Motion Cueing Algorithms (MCAs) without being fully able to know how well an MCA will perform during the experiment. Choices between MCAs can therefore be greatly supported by previous measurements or predictions of motion cueing quality. This paper describes a data collection experiment on a nine degree-of-freedom motion-base simulator, in which participants are asked to continuously rate the motion cueing quality during a pre-recorded drive through an urban environment. Three benchmark MCAs are compared: a Model-Predictive Control (MPC) algorithm with infinite prediction horizon, a Classical Washout Algorithm (CWA) tuned for the use-case, and the same algorithm (CWA), but with the tilt-coordination channels turned off. By comparing ratings for the whole scenario, as well as ratings for each maneuver individually, the results show a preference of the presence of tilt-coordination, as well as a preference for the optimization-based MPC algorithm over the CWA condition. The collected data will be used directly for modeling and predicting motion cueing quality for future experiments at BMW, such that the best-suited MCA and parameter setting can be selected before experiments.*

Keywords: *Motion cueing; quality comparison; urban simulations*

1. Introduction

By 2050, 68% of the global population will be living in urban areas (United Nations, Department of Economic and Social Affairs, Population Division, 2018). This is likely to impact the design and use of road vehicles. Driving simulation offers a unique possibility in guiding the development of new technologies in such environments by providing safe, emission-free and controllable alternatives to real vehicle testing.

A motion-base driving simulator is controlled by a Motion Cueing Algorithm (MCA), which maps the specific forces and rotational rates a driver would perceive in the measured or simulated vehicle onto the motion of the simulator. Several approaches for MCAs exist, such as the so-called Classical Washout Algorithm (CWA) (Conrad, Douvillier, and Schmidt, 1973; Reid and Nahon, 1985) and the more recently developed Model-Predictive Control (MPC) algorithms (first implemented for driving simulation by Dagdelen, et al., 2009). With growing computational power, the latter is becoming a suitable option for real-time applications (Beghi, Bruschetta, and Maran, 2012; Bruschetta, Cenedese, and Beghi, 2019). In order to be able to select and tune the best-suited MCA for each experiment, it would be ideal to be able to predict their motion cueing quality. This is also paramount to be able to justify the costs of other MCA quality factors, such as algorithm complexity, computational load and energy consumption (Kolff, et al., 2020).

In literature, recent developments have focused on the development of empirical models to predict motion cueing quality (such as Cleij, 2020), even without fully understanding the complete details of human perception. Such models have some predictive power to evaluate the subjective rating as given by humans based on a linear combination of the errors in the specific forces and rotational rates, or non-linear models to detect different types of cueing errors. Even though the models of Cleij show a high quality-of-fit, important questions are 1) whether these models still apply for combined, compound maneuvers of more realistic use-cases (as these were only tested for isolated maneuvers), 2) what the validity of the prediction is when applying different use-cases, such as urban scenarios, and 3) how such a model-based evaluation of human perceptual ratings can be used to perform a trade-off between MCAs, as well as their various configurations.

Currently, no continuously measured rating data are available for urban driving scenarios that can be used to answer these questions. Therefore, a data collection experiment is performed in the Driver-in-Motion simulator at BMW Group. The experiment makes use of the same paradigm followed by Cleij, 2020, by letting participants experience driving a pre-recorded route in an urban scenario, which resembles a realistic city driving environment. In the experiment, participants are asked to give continuous ratings on how well they think the motion of the simulator matches with what they expect from a real vehicle.

The paper is structured as follows. First, an introduction to measuring and modeling motion cueing quality, including an overview of the three compared

MCAs, is given. Then the experiment set-up is discussed in Section 3. Results and discussion are given in Section 4, followed by the conclusions.

2. Assessment of motion cueing quality

2.1. Design of experiments with motion cueing

When designing experiments with motion cueing for a given simulator, the best-suited Motion Cueing Algorithm (MCA) and its parameter settings must be selected, ideally based on the various advantages and disadvantages of the complete solution. In the design stage, however, many of these properties are often inherently unknown, such as which MCA approach is best suited for a use-case and how these MCAs would be perceived and rated by human drivers, which can only be evaluated after performing an evaluation study. This makes it difficult to systematically test MCA configurations for a given use-case.

For this reason, it is useful to make predictions of motion cueing quality on a perceptual level. This requires models of subjective ratings that are able to predict how well different MCAs are evaluated by humans, as subjective rating data are likely to be unavailable for experiments that are still to be performed. Such models can be either based on knowledge of the human-perceptual system, or empirical modeling based on the collection of rating data.

2.2. Measuring motion cueing quality

The work of Cleij, 2020; Cleij, et al., 2018 proposed a continuous rating method, in which participants give *continuous* motion incongruence ratings (MIRs) based on the perceived motion incongruences (PMIs). These ratings are indications of how accurately participants think the motion they perceive matches that of the real vehicle. Participants are driven through an environment passively, rather than driving themselves, such that they can fully focus on the rating method. The rating is measured on a scale from 0 to 10 and can be changed at any point in time, such that a continuous rating signal is obtained. A value of 0 indicates no mismatch at all, whereas the highest score of 10 indicates the worst motion cueing encountered in the experiment. The recorded continuous rating signals, having a high temporal resolution, can be used for modeling human quality measurements based on the errors in the motion cueing that are perceived by the subject.

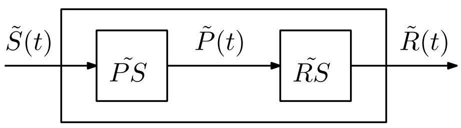


Figure 1: Block diagram of the human rating model, adapted from Cleij, 2020.

A block diagram of a human rating model as proposed in Cleij, 2020 is shown in Figure 1. Here, motion incongruences $\tilde{S}(t)$ are perceived by the human operator through a sensory system $\tilde{PS}$, which describes how humans perceive the motion, dominated by the vestibular and proprioceptive systems. The internal *perceived* motion incongruences $\tilde{P}(t)$ are used by the human as inputs for their rating dynamics $\tilde{RS}$. These rating dynamics include a time delay τ, but also elements such as the rating strategy, such as how actively participants rate. The rating dynamics result in a motion incongruence rating (MIR) signal $\tilde{R}(t)$ that can be directly measured through a rating interface, although ideally one would want to measure the signal $\tilde{P}(t)$.

This subjective rating method has since then been implemented in various studies (Cleij, et al., 2019; Cleij, et al., 2018; Ellensohn, et al., 2019; Ellensohn, et al., 2018; Van der Ploeg, et al., 2020). A benefit of this method is that it results in a continuous rating signal over time. Such a signal allows for detail-rich analysis of specific time periods or over larger time periods. Furthermore, the continuous rating data can be used for model fitting. Nevertheless, there are several limitations when working with continuous ratings.

First of all, it is assumed that any non-zero ratings given by the participants are caused by a perceived motion mismatch. This means that if participants would be asked to perform the same experiment while being driven around in a real vehicle, they should give perfect (zero) ratings all the time. However, participants are asked to actively rate the motion cueing quality, meaning that they might be inclined to more actively focus on incongruences than they would during real driving or any other driving experiment, as well as be more inclined to give out non-zero ratings, as they are specifically asked to rate. Measuring a quantity therefore has an effect on the measurement of the quantity itself, which is difficult or even impossible to completely mitigate, but can be reduced by proper instructions.

Furthermore, it is assumed that the continuous rating given by the humans at each point in time is also an accurate representation of what the human is perceiving at that point in time. Next to an inherent human time delay, as included in the models of Cleij, 2020, humans might anticipate for expected cueing errors of upcoming maneuvers based on *expectation* and scenario *experience*. The perception of incongruences of a past maneuver might also affect the given rating at the current time.

Nevertheless, with the drawbacks in mind, the continuous rating method still provides unique and useful information and is arguably currently the most detail-rich information source on motion cueing quality. Follow-up experiments in Ellensohn, et al., 2019; Ellensohn, et al., 2018; 2020 have extended the work of Cleij, 2020 towards more realistic, rural use-cases. The experiment described in this paper aims to bridge the gap in data collection by performing an urban driving data collection experiment, which is to be used for modeling in future research.

2.3. Motion Cueing Algorithms

Three MCAs were chosen for comparison in an urban driving scenario. The first algorithm is a Classical Washout Algorithm (CWA). The same MCA was also used in Ellensohn, et al., 2020, although with a different tuning configuration. Second, the same CWA without tilt-coordination is tested. Finally, a Model-Predictive Control (MPC) algorithm with perfect knowledge of the future states is used.

2.3.1. Classical Washout Algorithm (CWA)

The first MCA used in the experiment is a Classical Washout Algorithm (CWA). The algorithm is capable of real-time simulations, but limited in its use of the simulator's workspace, as it does not utilize optimization of the simulator's motion. This traditional approach (Conrad, Douvillier, and Schmidt, 1973; Reid and Nahon, 1985) uses high-pass filters in the inertial frame for the translational as well as the rotational channels, to wash out the simulator motion and ensure that the simulator remains in its workspace. The low-frequency translational accelerations are typically reproduced by tilt-coordination, by letting the gravity vector generate a sustained acceleration in longitudinal and lateral direction due to rotations in pitch and roll, respectively. The signal for tilt-coordination is low-pass filtered in the body frame, typically complementary to the high-pass filtered simulator translational accelerations. As the experiment described in this paper is performed on the nine degree-of-freedom (DoF) Driver-in-Motion simulator, with an additional tripod structure below its hexapod, additional filtering is applied to the acceleration channels, where a band-pass signal (with complementary filters to the hexapod translations and the tilt-coordination) is sent to the tripod. Tuning of the CWA was based on expert evaluations.

2.3.2. Classical Washout Algorithm without tilt-coordination (NTC)

The second algorithm under investigation is the same algorithm as the CWA discussed above, but with the tilt-coordination channel turned off, denoted as NTC. This was achieved by setting the gains of the tilt-coordination to zero. This creates large errors in the reproduced specific forces, which is expected to create less realistic motion cueing (Stratulat, et al., 2011). This allows for the comparison of the same algorithm with and without tilt-coordination, such that its effects can be explicitly quantified for urban scenarios. It also induces major cueing errors that are useful for understanding what exactly humans rate as bad, which are required for model fitting purposes. Furthermore, it provides a benchmark of (expected) "poor" motion cueing quality.

2.3.3. Oracle Model-Predictive Control (ORC)

The final MCA under investigation is an Model-Predictive Control (MPC) algorithm with infinite prediction horizon, therefore denoted as the 'oracle' (ORC). First implemented in driving simulation by Dagdelen, et al., 2009, MPC algorithms use predictions or knowledge of future states to optimize the simulator movement, by reducing cueing errors in the three specific force and three rotational rate signals through the minimization of a cost function:

$$J = \underbrace{\sum_{i=1}^{p} w_e e_{k+i}^2}_{\text{Minimizes cueing errors}} + \underbrace{\sum_{i=0}^{p-1} w_{\Delta u} \Delta u_{k+i}^2}_{\text{Minimizes control inputs}} + \underbrace{\sum_{i=0}^{p-1} w_x {x_{k+i}}^2}_{\text{Minimizes state excitations}}, \tag{1}$$

where e_{k+i} is the difference between the reference motion and motion of the simulator, which is used to minimize cueing errors. Δu_{k+i} is the control input vector and x_{k+1} the state vector. These state excitation and control input minimization terms prevent overshooting of the workspace and guarantee unique solutions for multiple DoF systems (Katliar, 2020). The weights of the errors in the six motion channels were set to $w_e = [1\ 1\ 1\ 10\ 10\ 10]^T$, similar as in Van der Ploeg, et al., 2020 and Katliar, 2020. In this case, the algorithm had perfect knowledge of the future states and is therefore able to calculate optimized simulator movement for the whole run. The algorithm is therefore expected to define the upper limit for motion cueing quality achievable on the simulator. More information regarding the optimization process is described in Ellensohn, et al., 2018.

3. Experiment set-up

3.1. Apparatus

The experiment was performed on the Driver-in-Motion simulator (Figure 2a), which is a nine degree-of-freedom structure, consisting of a traditional hexapod placed on a tripod. This simulator was chosen, as its additional lateral and yaw movement of the tripod are beneficial for urban environments, where these movements are required for cueing driving through sharp corners.

Participants were seated in the driver seat of the simulator. Four projectors presented the visuals on a 240° field-of-view screen. The velocity was visible on the traditional tachometer on the dashboard, as well as on the bottom of the main visuals, together with the driving direction and the current continuous rating given by the participant, for their own feedback.

3.2. Scenario

Driving data were collected in a virtual city scenario. The route consisted of multiple maneuvers that are typical for such environments. A top-down representation of the driven route is shown in Figure 3. The run started with an initial acceleration maneuver ('ACC') and was selected to have a balanced number of left- and right corners, indicated by 'CR'. Furthermore, it contained a roundabout ('RBT') and five lane changes ('LC'). 'DEC1' was a braking maneuver before a red traffic light, bringing the vehicle to a full stop and standing still for five seconds. After this, the vehicle performed a combined acceleration and corner maneuver ('ACR1'). 'DEC2' corresponds to the final deceleration of the run.

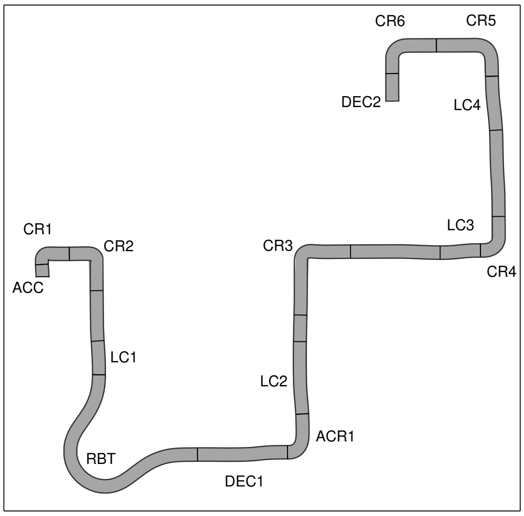


Figure 3: Top-down view of the driven route through the city, with the maneuvers indicated by their abbreviations.

Traffic was present to create a more realistic environment, but was configured to never come close to the direct surroundings of the simulated vehicle, such as by driving on the opposite lane, or passing by in front of the vehicle while it was standing still before the traffic light.

3.3. Independent variables

The experiment had a within-subject design with a single independent variable, which was the MCA. Three MCAs were tested, as described in Subsection 2.3. The resulting specific forces and rotational rates for all three MCAs are shown, together with the vehicle data, in Figure 4.

3.4. Dependent measures

By varying the MCA, the effect on the subjective ratings was measured through two mechanisms:

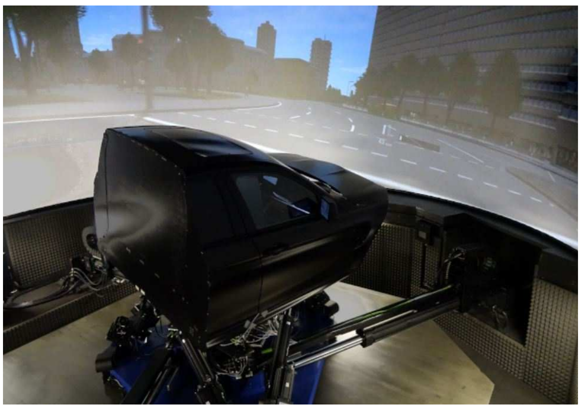
Continuous MIR The continuous MIR was given by the participants by rotating the BMW iDrive knob located on the central console of the vehicle (see Figure 2b). Participants were instructed to rate how well they thought the motion matched what they would expect from the real vehicle. Rotating the knob resulted in a step increase or decrease of the rating, meaning that only integer rating values could be given. The rating was on a scale from 0 to 10 and was visible on the screen.

Post-hoc rating Complementary to the continuous rating, participants were verbally asked to give a post-hoc rating at the end of each run, to rate the overall motion cueing quality. Although this is intended as a single rating for the whole run, it is possible that participants are more inclined to base their post-hoc rating on the worst motion cueing that was encountered during that run, or emphasize more on the initial and later segments of a run (serial-position effect).

3.5. Hypotheses

To analyze the quality of the three MCAs, the ratings are analyzed in two different ways. First, based on both the continuous MIRs and the post-hoc ratings, the scenario-based ratings are analyzed, which refer to how subjects rate the complete scenario. Secondly, the maneuver-based rating refers to analysing the ratings of each maneuver separately and is based on the continuous MIRs. For this purpose, the ratings corresponding to distinct maneuvers (as defined in Subsection 3.2) are analyzed.

By looking at the magnitude of the produced cues, the oracle algorithm (ORC) is best able to reproduce the specific forces and the yaw rate, as shown in Figure 4, followed by the Classical Washout Algorithm (CWA) and finally the CWA without tilt-coordination (NTC). Based on these observations, the following hypotheses are defined for the experiment:



(a) The Driver-in-Motion simulator while moving during the experiment.

(b) A participant using the rating knob inside the simulator to rate the Perceived Motion Incongruences (PMIs), resulting in the Motion Incongruence Rating (MIR), adapted from Ellensohn, et al., 2018.

Figure 2: The experiment set-up.

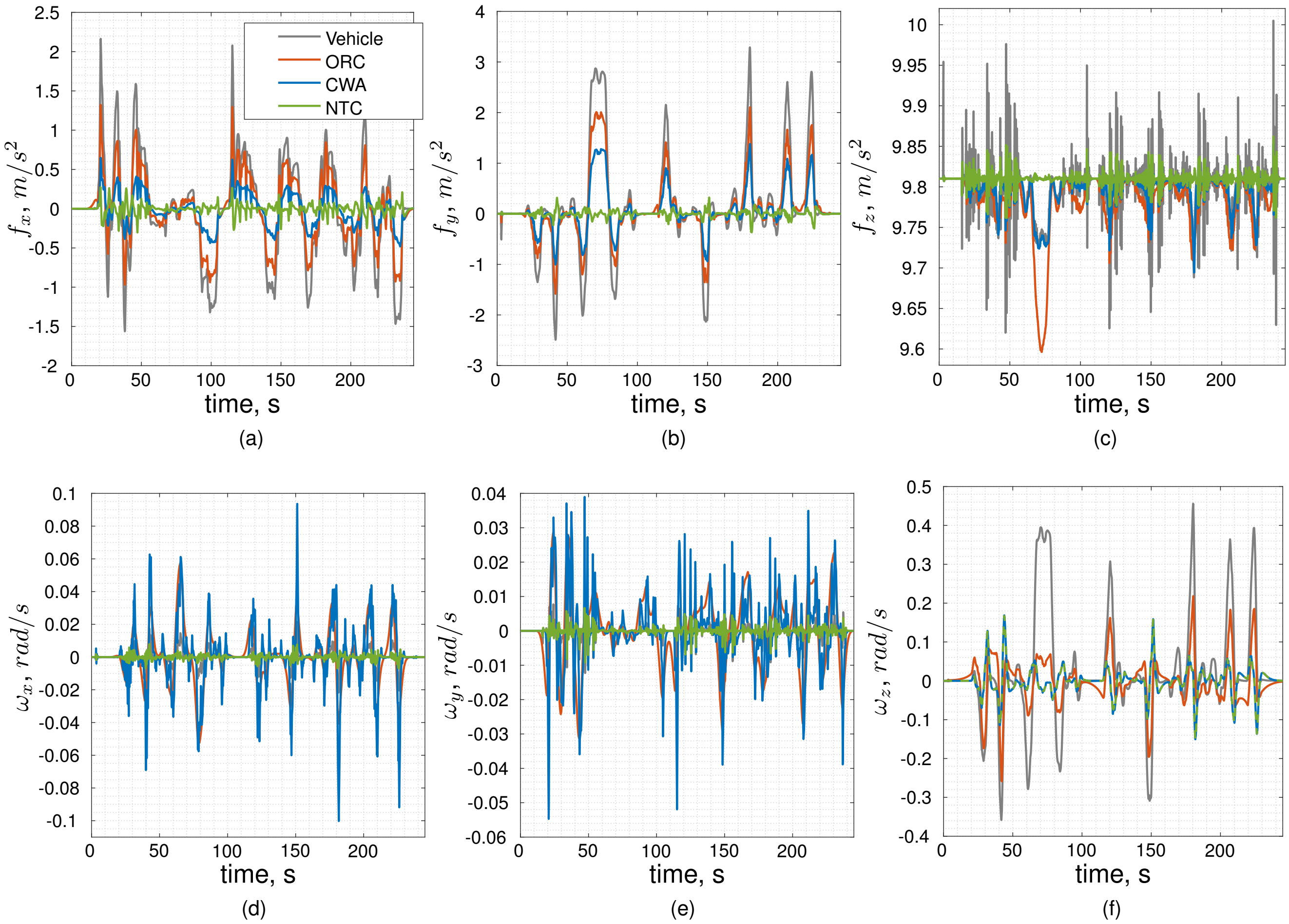


Figure 4: Specific forces (a-c) and rotational rates (d-f) for the measured vehicle data and the three MCA outputs.

H1: The scenario-based rating of the oracle (ORC) algorithm is lower than that of the Classical Washout Algorithm (CWA).
H2: The scenario-based rating of the Classical Washout Algorithm (CWA) is lower than that of the Classical Washout Algorithm without tilt-coordination (NTC).
H3: The scenario-based rating of the oracle (ORC) algorithm is lower than that of the Classical Washout Algorithm without tilt-coordination (NTC).
H4: The maneuver-based ratings of the oracle (ORC) algorithm are lower than those of the Classical Washout Algorithm (CWA).
H5: The maneuver-based ratings of the Classical Washout Algorithm (CWA) are lower than those of the Classical Washout Algorithm without tilt-coordination (NTC).
H6: The maneuver-based ratings of the oracle (ORC) algorithm are lower than those of the Classical Washout Algorithm without tilt-coordination (NTC).

3.6. Experiment procedure

A total of 60 participants performed the experiment, of which 50 were male and 10 female. All of them were employees of BMW Group with a European car driver's license B (μ = 22.38 years, σ = 10.16 years) and an average yearly driven distance of μ = 18833 km (σ = 13207 km). The average age was μ = 40.08 years with a standard deviation of σ = 10.12 years. 33 of the participants had previous experience with driving simulation. All participants provided informed consent before the experiment.

3.6.1. Training run

The experiment started with a training task, in which participants could familiarize with the experiment setup, such as the city environment, the sensation of simulator motion and the rating method. The simulator motion of the training run was controlled by the CWA algorithm, which used the same parameter setting as used in the measurement runs, including tilt-coordination. The training contained two lane changes and four corners, of which one contained lateral and yaw motion in the wrong direction, as a false cue, to present an example of 'bad' motion to the participants, of which they were informed before the corner. This cue was used as an anchoring for the maximum (10) rating. Furthermore, the braking cue before the last corner was inverted to create an acceleration cue, without informing the participants beforehand. This intentional false cue was used to check if they understood the rating task. The training run lasted four minutes and was repeated if it was observed that participants required some more time to get acquainted with the rating method.

3.6.2. Measurement runs

After the training, the measurement runs started. Participants performed a total of nine runs, with three repetitions of the same condition. The condition order was randomized by selecting one of the three sequences of runs (See Table 1), such that the condition sequence was balanced over all subjects. Participants were unaware that three repetitions of the same MCA were tested. Each run lasted four minutes, after which they were asked for the overall rating. Participants were required to take a break after runs 3 and 6. The experiment, including briefing and training, lasted approximately 90 minutes.

Table 1: The randomized order of the conditions as used in the experiment. Individual participants tested either the S1, S2 or S3 condition sequence for their nine runs.

Run	1	2	3	4	5	6	7	8	9
S1	CWA	NTC	ORC	NTC	ORC	CWA	ORC	CWA	NTC
S2	NTC	ORC	CWA	ORC	CWA	NTC	CWA	NTC	ORC
S3	ORC	CWA	NTC	CWA	NTC	ORC	NTC	ORC	CWA

3.6.3. Motion sickness

Due to the relative high visual and motion intensity of the use-case of urban driving, it was expected that motion sickness would be prevalent during the experiment, similar as in Hogerbrug, et al., 2020. The misery score (MISC), as introduced by Bos, MacKinnon, and Patterson, 2005, was used to keep track of motion sickness during the experiment. The MISC is a single-score evaluation method on a Likert scale, in which numbers ranging from zero to ten indicate increasing levels of discomfort.

The evaluation of the MISC served as a safety mechanism to protect the well-being of the participants. Participants were asked for a MISC-score after each run. Similar as in Hogerbrug, et al., 2020, the experiment was stopped if a MISC-score of 7 (medium nausea) or higher was reached for a single run, or if a MISC-score of 6 (some nausea) was attained two runs in a row. Participants were always free to abort the experiment themselves without specifying a reason, for example if they felt uncomfortable.

3.7. Data collection

For the participants who suffered from motion sickness and could not finish the experiment, the data sets were not used for analysis, as these were incomplete. One data set was excluded from further analysis, as upon visual inspection it was found that this participant did not respond to the maneuvers at all and kept the continuous rating at a high value constantly. The three ratings as given by the participants per condition were averaged, resulting in a total of three data sets per participant.

4. Results and Discussion

Out of the total of 60 participants, five participants (8.33%) were not able to finish the experiment due to too high MISC-scores. Three additional participants (5.00%) indicated they felt uncomfortable and

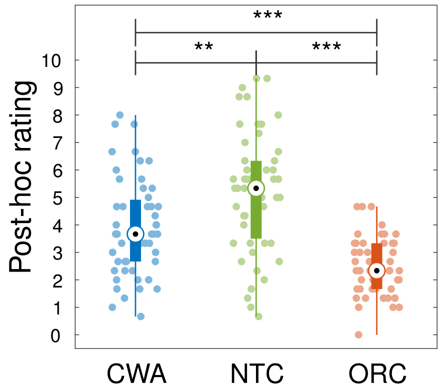


Figure 5: Box plots and data point distributions of the post-hoc ratings, per condition.

therefore quit the experiment on their initiative. In two cases the experiment could not be finished due to technical issues. This leaves a total of 50 complete data sets that are used for data analysis.

4.1. Scenario-based ratings

The scenario-based ratings (related to H1, H2 and H3) are calculated in two different ways. First, the overall post-hoc ratings for each MCA are analysed. Figure 5 shows the box plot distributions of the post-hoc ratings. Lilliefors tests ($p \leq 0.01$) showed that the data were normally distributed. When analyzing the means of the ratings, ORC (μ = 2.44, σ = 1.04) is rated better than CWA (μ = 3.89, σ = 1.79). The classical washout condition without tilt-coordination, NTC, is rated the worst with μ = 5.18 and σ = 2.17.

Furthermore, it was checked whether there were significant differences between the three MCA pairs (CWA-NTC, CWA-ORC and NTC-ORC). Within these pairs of independent variable settings, a repeated measures one-way ANOVA was performed. A post-hoc analysis with Bonferroni correction was used to correct for multiple comparisons. This shows very significant differences ($p \leq 0.01$) between the CWA-NTC pair and highly significant differences for the other two pairs ($p \leq 0.001$), indicated in Figure 5.

Complementary to the overall ratings, the average continuous ratings were calculated per condition. The continuous rating time signals are shown in Figure 6. The time-varying lines indicate the continuous MIR $\tilde{R}(t)$, averaged over all participants. The shaded areas indicate the standard deviation, and the horizontal lines indicate the averages of the continuous ratings for each MCA. Here, it is clear that the NTC algorithm performs the worst overall (highest rating), whereas ORC is the lowest rated. As the continuous rating includes segments where ratings are typically zero for all MCAs (such as standing still or driving perfectly straight), the average continuous ratings are lower than the post-hoc rating, as is visible in Figure 7. Nevertheless, CWA-NTC and NTC-ORC show highly significant differences, for CWA-ORC the difference is very significant. The means and standard deviations of the algorithms are for ORC: μ = 0.44, σ = 0.44, for CWA: μ = 0.72, σ = 0.67 and for NTC: μ

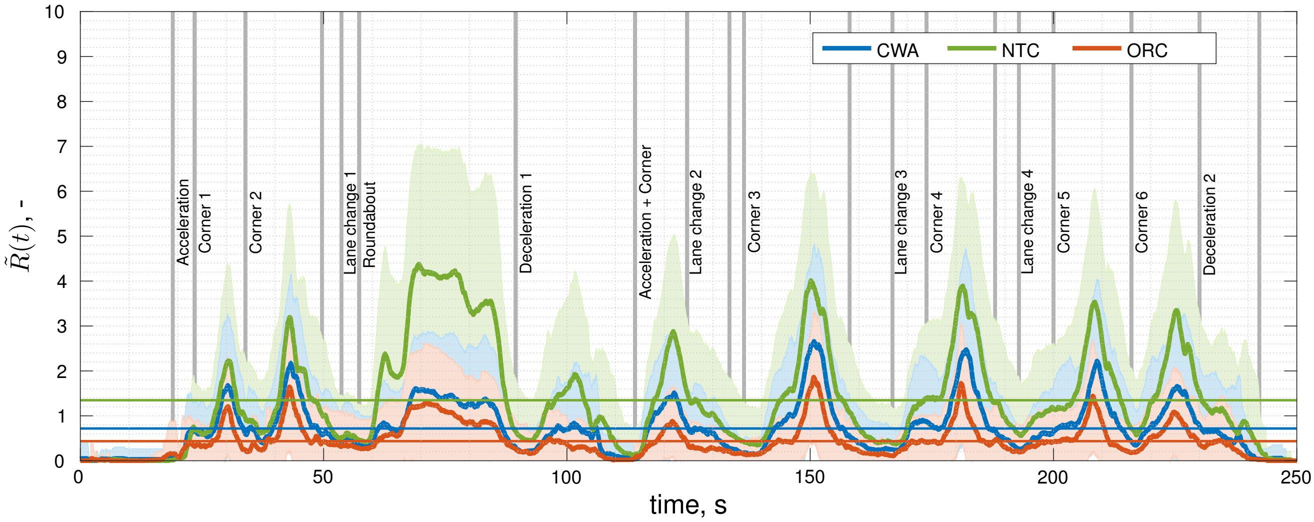


Figure 6: Continuous rating signals for the three MCAs, with the standard deviation displayed as shaded areas. The vertical lines indicate the distinct maneuvers, the horizontal lines show the means of the overall ratings.

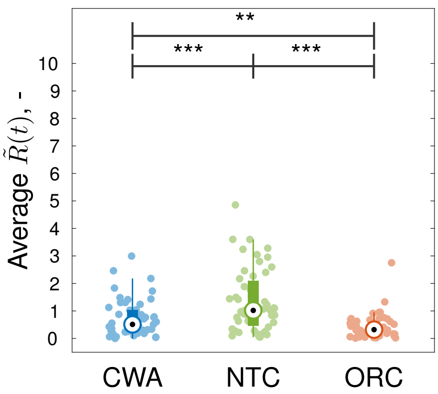


Figure 7: Box plots and data point distributions of the average MIRs, per condition.

= 1.35, σ = 1.14, which are in line with the post-hoc ratings. Based on both of these findings, hypotheses H1, H2 and H3 are confirmed.

4.2. Maneuver-based ratings

Figure 6 also shows the division of the maneuvers. The maneuver-based ratings are calculated by averaging the rating in that section. As these data were not normally distributed for any of the maneuvers except for 'Corner 4', the Friedman's test was used for all maneuvers to check for significance. Similar to the scenario-based ratings, the mean values are complemented with the three MCA pairs to look for significance, where a Bonferroni correction is applied. These results are shown in Table 2.

These results show that when analyzing the pair 'CWA-NTC', the condition NTC is the worst rated of the two, as all other mean values are higher than those of the CWA conditions. Only for the initial acceleration the NTC has a slightly lower mean value, but this is not a significant difference. Furthermore, it was found that only the maneuvers at the start of the run (ACC, CR1, CR2 and LC1), which are relatively weak compared to the rest of the run, do not have significant differences between the two conditions, whereas all other maneuvers do. This indicates that participants generally do prefer the addition of tilt-coordination, therefore (partially) confirming hypothesis H5. Especially the roundabout, where constant tilt-coordination is the only method to produce the sustained lateral acceleration, sees a strong beneficial effect of adding tilt-coordination.

Furthermore, when comparing CWA with ORC, the latter results in lower average ratings compared to the classical washout strategy. In all maneuvers that contained major longitudinal cues (ACC, DEC1, ACR1 and DEC2), no significant differences were found. However, all cornering maneuvers are rated significantly better for ORC than for CWA, partially confirming hypothesis H4. This indicates that for the highly dynamic 90° corners, typical for urban environments, there is an advantage when using an optimization-based algorithm. Even though the latter is not possible in real-time, human-in-the-loop driving, optimization-based algorithms, also those that *are* capable of supporting human-in-the-loop driving, might significantly improve cueing quality in urban simulations compared to washout algorithms.

It is clear from the average values that NTC and ORC are the worst- and best-rated conditions, respectively, meaning that their respective differences compared to the CWA algorithm are most interesting. Nevertheless, Table 2 also includes this pair for completeness. Significant differences for almost all maneuvers are found, confirming hypothesis H6.

When considering the three conditions, the findings indicate a benefit of tilt-coordination in urban environments, as well as a benefit of using optimization-based strategies, mainly for lateral maneuvers. The deterministic nature of lateral maneuvers, based on the road geometry, can ensure that real-time capable optimization-based algorithms are able to reach a similar performance for lateral maneuvers compared to oracle algorithms. These findings can be used as general guidelines for MCA selection, as well as serve as the basis for predictive models for choosing the best-suited MCA settings for an urban simulation.

Table 2: Average MIR per MCA and significance levels per maneuver between each MCA pair. A * is significant ($p<0.05$), ** is very significant ($p<0.01$) and *** is highly significant ($p<0.001$).

		Average $\tilde{R}(t)$			Significance pairs		
Maneuver	Abbreviation	CWA	NTC	ORC	CWA-NTC	CWA-ORC	NTC-ORC
Acceleration	ACC	0.31	0.24	0.25	-	-	-
Corner 1	CR1	0.95	1.22	0.56	-	***	***
Corner 2	CR2	1.02	1.48	0.63	-	*	***
Lane change 1	LC1	0.44	0.54	0.32	-	-	*
Roundabout	RBT	1.05	2.85	0.79	***	-	***
Deceleration 1	DEC1	0.40	0.89	0.22	***	*	***
Acceleration + Corner	ACR1	0.94	1.58	0.49	**	-	***
Lane change 2	LC2	0.57	1.04	0.26	**	***	***
Corner 3	CR3	1.08	1.76	0.62	***	*	***
Lane change 3	LC3	0.60	0.87	0.32	**	-	***
Corner 4	CR4	1.30	2.32	0.71	***	***	***
Lane change 4	LC4	0.48	0.91	0.34	*	-	***
Corner 5	CR5	1.09	1.86	0.60	**	**	***
Corner 6	CR6	1.01	1.85	0.60	**	**	***
Deceleration 2	DEC2	0.46	0.91	0.26	**	-	***

5. Conclusion

For the purpose of motion cueing quality modeling, an experiment was performed where participants were driven through an urban environment. By letting the participants continuously rate their perceived motion incongruence, the quality of motion cueing for urban driving with three MCAs was compared: an optimization-based Model-Predictive Control algorithm with infinite prediction horizon, a Classical Washout Algorithm (CWA) and the same algorithm without tilt-coordination active. Both overall (scenario-based) and maneuver-based rating data show that participants prefer the optimization-based strategy over the CWA, and that the presence of tilt-coordination has a positive effect on the motion cueing quality. These data will be essential for modeling and predicting human rating behavior in urban driving scenario's, required for motion cueing quality predictions of untested MCA configurations.

References

Beghi, A., Bruschetta, M., and Maran, F., Dec. 2012. A real time implementation of MPC based Motion Cueing strategy for driving simulators. In: *Proceedings of the 51st IEEE Conference on Decision and Control, Maui, Hawaii*.

Bos, J. E., MacKinnon, S. N., and Patterson, A., Dec. 2005. Motion Sickness Symptoms in a Ship Motion Simulator: Effects of Inside, Outside, and No View. *Aviation, Space, and Environmental Medicine*, 76(12), pp. 1111–1118.

Bruschetta, M., Cenedese, C., and Beghi, A., 2019. A real-time, MPC-based Motion Cueing Algorithm with Look-Ahead and driver characterization. *Transportation Research Part F: Traffic Psychology and Behaviour*, 61, pp. 38–52.

Cleij, D., Feb. 2020. Measuring, modelling and minimizing perceived motion incongruence. PhD Dissertation. Delft University of Technology.

Cleij, D., Venrooij, J., Pretto, P., Katliar, M., Bülthoff, H., Steffen, D., Hoffmeyer, F., and Schöner, H.-P., 2019. Comparison between filter- and optimization-based motion cueing algorithms for driving simulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 61, pp. 53 –68.

Cleij, D., Venrooij, J., Pretto, P., Pool, D. M., Mulder, M., and Bülthoff, H. H., Feb. 2018. Continuous Subjective Rating of Perceived Motion Incongruence During Driving Simulation. *IEEE Transactions on Human-Machine Systems*, 48(1), pp. 17–29.

Conrad, B., Douvillier, J. G., and Schmidt, S. F., 1973. Washout Circuit Design for Multi-Degrees-Of-Freedom Moving Base Simulators. In: *Proceedings of the AIAA Visual and Motion Simulation Conference, Palo Alto (CA), United States, September 10-12*. AIAA-1973-929.

Dagdelen, M., Reymond, G., Kemeny, A., Bordier, M., and Maïzi, N., Apr. 2009. Model-based predictive motion cueing strategy for vehicle driving simulators. *Control Engineering Practice*, 17(2009), pp. 995–1003.

Ellensohn, F., Hristakiev, D., Schwienbacher, M., Venrooij, J., and Rixen, D., 2019. Evaluation of an Optimization Based Motion Cueing Algorithm Suitable for Online Application. In: *Proceedings of the Driving Simulation Conference 2019 Europe, Strasbourg, France*, pp. 1–8.

Ellensohn, F., Schwienbacher, M., Venrooij, J., and Rixen, D., 2018. Motion Cueing Algorithm for a 9 DoF Driving Simulator: MPC with Linearized Actuator Constraints. *SAE International Journal of Connected and Automated Vehicles*, 2(3), pp. 145–155.

Ellensohn, F., Spannagl, M., Agabekov, S., Venrooij, J., Schwienbacher, M., and Rixen, D., 2020. A hybrid motion cueing algorithm. *Control Engineering Practice*, 97, p. 104342.

Hogerbrug, M., Venrooij, J., Pool, D. M., and Mulder, M., 2020. Simulator Sickness Ratings Reduce with Simulator Motion when Driven Through Urban Environments. In: *Proceedings of the Driving Simulation Conference 2020 Europe, Antibes, France*.

Katliar, M., Dec. 2020. Optimal control of motion simulators. PhD Dissertation. Albert-Ludwigs-Universität Freiburg.

Kolff, M., Venrooij, J., Pool, D. M., and Mulder, M., 2020. Comparison of Quality Metrics between Motion Cueing Algorithm in a Virtual Test Environment. In: *Proceedings of the Driving Simulation Conference 2020 Europe, Antibes, France*.

Reid, L. and Nahon, M., 1985. *Flight Simulation Motion-Base Drive Algorithms: Part 1. Developing and Testing the Equations*. Tech. rep. UTIAS 296. University of Toronto, Institute for Aerospace Studies.

Stratulat, A., Roussarie, V., Vercher, J., and Bourdin, C., 2011. Does tilt/translation ratio affect perception of deceleration in driving simulators? *Journal of vestibular research: equilibrium & Orientation*, 21(3), pp. 127–139.

United Nations, Department of Economic and Social Affairs, Population Division, 2018. *World Urbanization Prospects: The 2018 Revision*. Tech. rep. ST/ESA/SER.A/420. New York: United Nations.

Van der Ploeg, J. R., Cleij, D., Pool, D. M., Mulder, M., and Bülthoff, H. H., 2020. Sensitivity Analysis of an MPC-based Motion Cueing Algorithm for a Curve Driving Scenario. In: *Proceedings of the Driving Simulation Conference 2020 Europe, Antibes, France*.