Todo list

rework the discussion section . . . 14
add about probabilistic model . . . 14
rework the conclusion section . . . 16
add about probabilistic model . . . 16
add imu data plots . . . 18
Figure: 6DOF head motion for Slalom Drive . . . 18
Figure: 6DOF vehicle motion for car and simulator Drive . . . 18
Figure: 6DOF vehicle motion for Sickness with NDRT . . . 18
add plots of fitting of all participants . . . 18

Individualizing (or Personalising) Motion Sickness Models: Estimation and Distribution of Individual-Specific Parameters

Varun Kotian[1], Daan M. Pool[2], Riender Happee[1]
[1]Delft University of Technology, Faculty of Mechanical Engineering, Cognitive Robotics, 2628 CD Delft
[2]Delft University of Technology, Faculty of Aerospace Engineering, Control & Simulation, 2629 HS Delft

August 26, 2024

Abstract

Users of automated vehicles will move away from being drivers to passengers, preferably engaged in other activities such as reading or using laptops and smartphones, which will strongly increase susceptibility to motion sickness. Similarly, in driving simulators, the presented visual motion with scaled or even without any physical motion causes an illusion of passive motion, creating a conflict between perceived and expected motion, and eliciting motion sickness. Given the large differences in sickness susceptibility between individuals, it seems that any truly effective counter measure must consider sickness at an individual level. This paper combines a group-averaged sensory conflict model (as in Wada et al. (2020)) with an individualized accumulation model (as in Oman (1990); Irmak et al. (2020, 2022)) to capture individual differences in motion sickness susceptibility across various vision conditions. This consideration of the effect of vision is crucial in driving simulators where there is a strong contribution of visual cues. The model framework can be used to develop personalized models for users of automated vehicles and improve the design of new motion cueing algorithms for simulators. The feasibility and accuracy of this model framework are verified using two existing datasets with sickening conditions in 1) an experimental vehicle with and without outside vision (Irmak et al., 2020), and 2) comparing vehicle experiments with corresponding driving simulator experiments (Talsma et al., 2023). Both datasets involve passive motion, representative of being driven by an automated vehicle. The model is able to fit an individual's motion sickness responses using only two individualized parameters (gain K_1 and time constant T_1), as opposed to the 5 parameters in the original model. This ensures unique parameters for each individual. Better fits, on average by a factor of 1.7 (for AM2 model), of an individual's motion sickness levels, are achieved as compared to using only the group-averaged model (AM0 model). Furthermore, this model framework demonstrates robustness by accurately modelling various datasets with distinct motion and vision conditions. Thus, we find that models predicting group-averaged sickness incidence cannot be used to predict sickness at an individual level. On the other hand, the proposed combined model approach predicts individual motion sickness levels and thus can be used to control sickness.

Keywords— Motion Sickness, Simulator Sickness, Modeling, Driving Simulators, Automated Vehicles.

1 Introduction

Automated vehicles and driving simulators are very different technologies. However, they both share two common facts. The first is that they have become very popular in recent years, a trend that is expected to continue in the future. Secondly, they both share a common issue in *motion sickness*. Users of automated vehicles will move away from being drivers to passengers, preferably engaged in other activities such as reading or using laptops and smartphones, which strongly increases susceptibility to motion sickness. Similarly, in driving simulators, the apparent visual motion combined with scaled or without any physical motion causes an illusion of passive motion, creating a conflict between perceived and expected motion, and eliciting motion sickness (Bos et al., 2020). Though these two cases seem different, the inherent mechanism that causes motion sickness in both, i.e., *sensory-expectancy conflict*, is the same.

The mechanisms behind the development and evolution of motion sickness have been studied extensively, relying heavily on models that predict sensory conflicts based on inputs from the vestibular and visual sensory systems (Bos and Bles, 1998; Wada et al., 2020; Liu et al., 2022; Irmak et al., 2023; Kotian et al., 2024). These models are known as 'conflict generation' models. However, these models predict group-averaged sickness responses (conflicts), using Motion Sickness Incidence (MSI), and thus cannot be reliably used for predicting and controlling sickness, both in real vehicles and simulators, at an individual level.

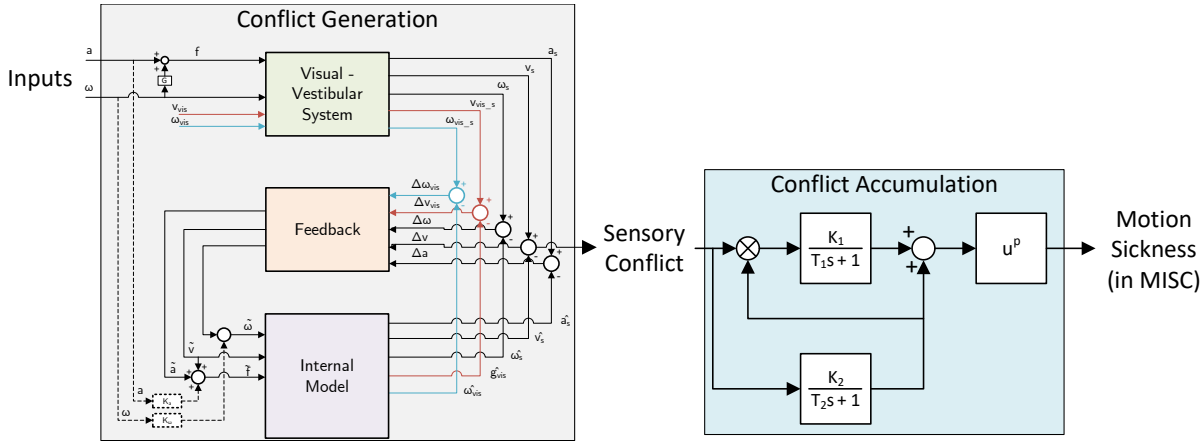


Figure 1: Framework for combined 'conflict generation' using the SVC model (Wada et al., 2020; Liu et al., 2022) and 'conflict accumulation' (Oman, 1990; Irmak et al., 2022) models to predict individual motion sickness, i.e., MIsery SCale (Reuten et al., 2020)

In this paper, we aim to combine the 'conflict generation' model with an 'individual conflict accumulation' model to improve the ability to capture the differences in individual susceptibility. This model framework is visualized in figure 1. 'Conflict generation' models, such as the SVC model by (Wada et al., 2020), calculate MSI using a conflict accumulator that integrates the generated conflicts in time. However, this is a single low-pass integrator and hence is not able to model the slow dynamics of motion sickness such as hypersensitivity after already being motion sick. One solution is to make use of the same 'conflict generation' model and replace the accumulation integrator with a much more complex integrator model (as in Oman (1990)), which will be referred to as the 'conflict accumulation' model in this paper (see figure 1). These 'conflict accumulation' models have been found very useful in motion sickness predictions for one degree of freedom motion (Irmak et al., 2020, 2022). In this paper, these will be personalized to each individual with a unique set of parameters. Further, a complex 'conflict accumulation' model will be able to capture crucial phenomena like recovery from motion sickness and hypersensitivity (Irmak et al., 2020, 2022). The personalization of parameters has already been shown to improve modeling accuracy by (Irmak et al., 2020) where it increased the accuracy by a factor of 2 compared to group-averaged models. Our work extends this to 6 degrees of freedom (DOF). This model has the additional benefit of allowing visual inputs. This is important when predicting motion sickness in simulators, i.e. simulator sickness, where there is a strong influence of vision. Additionally, individual motion sickness metrics, such as those obtained using the MISC scale (Reuten et al., 2020; De Winkel et al., 2022), instead of group-averaged metrics like Motion Sickness Incidence (MSI) (as in Wada et al. (2020); Kotian et al. (2024)), allow individualized fitting of the model resulting in better prediction in future motion paradigms for the individual.

The goal of the paper is to verify the feasibility of the combined motion sickness model approach, i.e., combining 'conflict generation' and 'conflict accumulation' models, to capture individual differences in motion sickness susceptibility across various vision conditions. To do this we have two further sub-goals:

- To verify the accuracy of the combined conflict generation/accumulation modelling approach
- To find a balance between model accuracy and the number of individually estimated accumulation model parameters

The second goal ensures a practical model with estimated parameter sets unique to each individual. We use two existing datasets with sickening conditions in 1) an experimental vehicle with and without outside vision (Irmak et al., 2020), and 2) comparing vehicle experiments with corresponding driving simulator experiments (Talsma et al., 2023). Both datasets involve passive motion, representative of being driven by an automated vehicle.

By achieving these goals, this framework can be used to develop personalized models for users of automated vehicles, which the automated vehicle can use to determine the comfort levels of individual users and subsequently best adapt its driving style, i.e., by limiting the acceleration and rotations of the car. This model framework can also improve and accelerate the design of new optimization-based motion cueing algorithms that aim to minimize sickening motions in simulators like in Hogerbrug et al. (2020); Baumann et al. (2021); Jain et al. (2023).

2 Method

Experimental Datasets

Table 1: Experimental datasets used in this study

Datasets	Details	Reference	No. of Participants
Slalom Drive Dataset	Slalom with Internal and External vision Motion sickness responses for hypersensitivity	Irmak et al. (2020)	16
Car and Simulator Dataset	Naturalistic drive in vehicle and moving base simulator Only External vision	Talsma et al. (2023)	24
Sickness with NDRT Dataset	Naturalistic driving data with varying NDRTs External and Internal Vision	Metzulat et al. (2024)	20
Efficient recreation of Motion Sickness Dataset	Naturalistic driving data on road recreated on a smaller track Internal Vision	[cite]	47

For the analysis in this paper, we make use of an existing real-world 'Slalom Drive' dataset (16 participants) with varying vision conditions, i.e., with and without an outside view, where mean MISC levels at the end of motion exposure were 5.3 (severe symptoms) with an outside view and 3.3 (some symptoms) without an outside view (Experiment 1 in (Irmak et al., 2020)), see figure 2. This experiment compared the motion sickness development with and without an outside view from the car. This dataset is ideal for proving that our new model framework can predict individual motion sickness for various vision conditions in real vehicles. The model framework is also validated using the 'Car and Simulator' dataset (24 participants) by Talsma et al. (2023) and 'Sickness with NDRT' dataset by Metzulat et al. (2024). The 'Car and Simulator' experiment contains motion sickness responses from real-world driving and its (matched) simulation on a moving-base driving simulator (Talsma et al., 2023). For this experiment, the mean MISC levels at the end of motion exposure were around 5.5 (severe symptoms) in the car and 1.5 (slight discomfort or vague symptoms) in the simulator (see figure 3). The 'Sickness with NDRT' experiment is real-world sickening drive around a fixed track with three different non-driving related tasks (NDRT) - visual dynamic, visual static and auditory. The mean MISC levels at the end of each conditions were around 4.5 (medium symptoms) for the auditory, 5 (severe symptoms) for the visual static and 6.5 (some nausea) for the visual dynamic condition. Using the 'Car and Simulator' and 'Sickness with NDRT' datasets, we demonstrate the capability of the model framework to generalize well for different types of passive motion, whether it originates from a car or a simulator while carrying out visual or auditory tasks. Furthermore, all datasets are used to study the number of parameters needed to accurately simulate the motion sickness development in different participants.
All three datasets capture sickness responses to passive motion, representative of being driven by an automated vehicle. In these datasets individual motion sickness levels (using MISC) were reported as a function of time with motion stimuli varying in time, showing major individual differences in sickness susceptibility, making this data suitable to model sickness accumulation at the individual level.

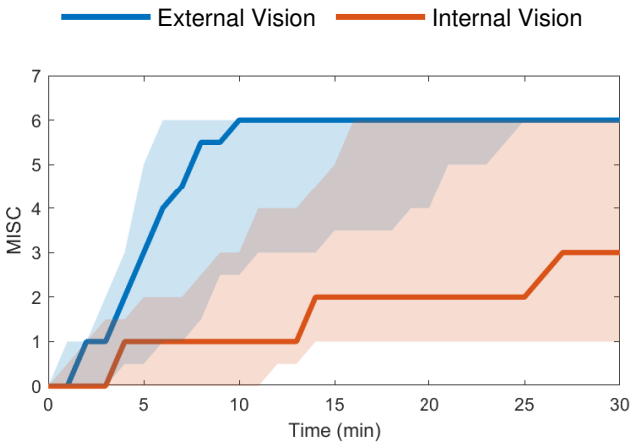


Figure 2: Slalom Drive dataset (Irmak et al., 2020) group-averaged MISC levels versus time in the conditions of external (blue) and internal vision (red)

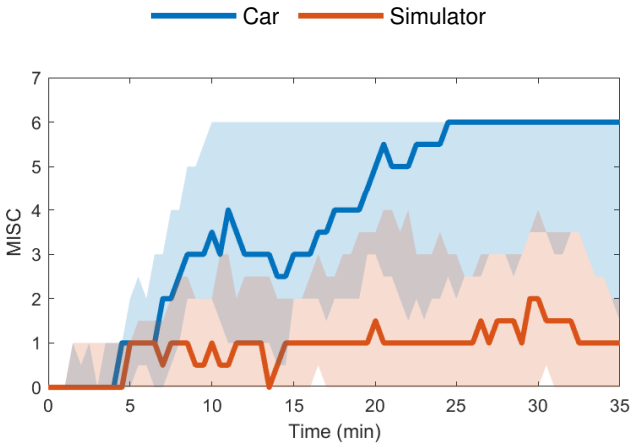


Figure 3: Car and Simulator dataset (Talsma et al., 2023) group-averaged MISC levels versus time in the car (blue) and simulator (red)

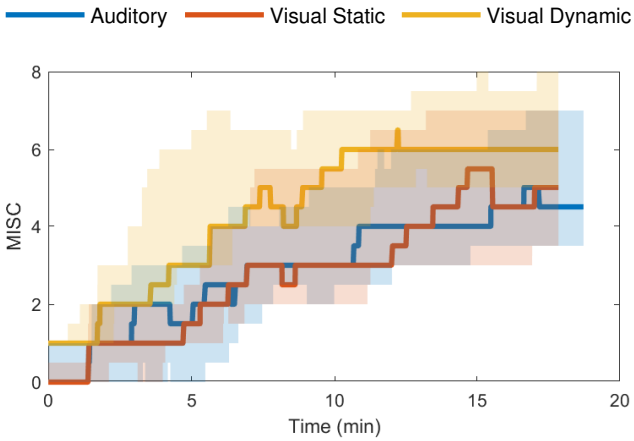


Figure 4: Sickness with NDRT dataset (Metzulat et al., 2024) group-averaged MISC levels versus time in the three different tasks - auditory (blue), visual static (red) and visual dynamic (yellow)

Inputs and Outputs

Inertial inputs, such as acceleration and angular velocity, and vision inputs, such as visual verticality (orientation) and visual rotation (rotational velocity), will be input into the 'conflict generation' part of the model which is the SVC model by Wada et al. (2020) and Liu et al. (2022). These inputs are shown on the left side of figure 1.

We applied 6D motion (3D translation, and 3D rotation), using the recorded seat and platform motion for the Car and Simulator dataset, and recorded head motion for the Slalom Drive dataset. This is shown in appendix A.

The human eye estimates motion through vision by measuring the rotation of visual cues between the current and previous states, a process known as optic flow. We assume the visually perceived rotations to be equivalent to head (or vehicle, as in the case of the Car and Simulator dataset) rotations when observing the external environment. For internal vision, we select a zero visual input assuming no visual head motion relative to the vehicle.

The output of the model is in form of a score on the MIsery SCale (MISC) by Bos et al. (2005), which measures the progression of symptoms Reuten et al. (2020) and has positive relation to discomfort De Winkel et al. (2022). It is a 11-point symptom-based scale that measures in discrete symptoms running from 0 to 10 where zero means no problem and ten indicates vomiting. In contrast, the model simulations predict the score on a continuous scale. These predictions will be compared with experimentally reported MISC scores. The parameters of the model, including time constants as well as gains (which are explained extensively in subsequent section) shall be tuned in 'conflict accumulation' model so as to have best fit for individual responses.

Selection of models

As previously discussed, we combine two models in our proposed model framework. The first component of the model is the 'conflict generation' model (see the left part in figure 1), which models the integration of sensory inputs and generates a one-dimensional sensory conflict signal. This model should be reliable in forecasting conflict signals consistently, especially while being driven. Additionally, the model must be capable of incorporating visual inputs so that the effects of different vision conditions can be predicted. To select an appropriate model, we refer to our study by Kotian et al. (2024) where various models and implementations of vision inputs were compared.

Based on their findings, we conclude that the 6DOF Subjective Vertical Conflict (SVC) model with Visual Rotational Velocity input (Wada et al., 2020) is ideal for our purpose of predicting motion sickness. This model has been shown to accurately replicate the frequency and amplitude dynamics of motion sickness as reported in numerous previous studies (McCauley et al., 1976; Golding and Markey, 1996; Griffin and Mills, 2002; Howarth and Griffin, 2003; Irmak et al., 2021, 2022). It is shown in Kotian et al. (2024) that the Visual Vertical input (red pathways in figure 1) as implemented in the SVC model was not effective in predicting sickness and hence will not be used in the current study (the corresponding gains were set to zero).

The second part of the model is the 'conflict accumulation' model. As the name suggests, this model accumulates (integrates) the sensory conflicts generated by the first part of the model framework to estimate the level of motion sickness experienced by an individual as it builds up over time. Usually, the conflict is integrated using a hill function followed by a leaky integrator. However, this method is insufficient to capture the unique dynamics of motion sickness, such as hypersensitivity, which is the faster than normal increase of motion sickness after already being subjected to motion sickness before. Thus, we adopt a more advanced Oman model (Oman, 1990) for this. The Oman model is a five-parameter non-linear model that integrates the conflict (see the right part in figure 1), with fast and slow pathways combined with leakage. The fast path has a low time constant (T_1) and models the direct response to a sickening stimulus. The slow path has a high time constant (T_2) and captures the slower secondary effects of sickening stimuli, such as recovery and hypersensitivity. This is also relevant in simulators where a sudden increase in motion incongruence can make participants hypersensitive. This has been observed by Cleij et al. (2018); Kolff et al. (2022), where it was shown that motion tends to be momentarily bad, but not continuously. Both pathways have a gain (K_1 and K_2) to control the contribution of each path. Additionally, there is a power law (p) at the model's output to account for nonlinear scaling effects.

In this paper, we use the combined model to predict sickness as rated with the MIsery SCale (MISC) (Reuten et al., 2020; De Winkel et al., 2022). These predictions will be compared against the experimentally reported MISC scores and the parameters in the 'conflict accumulation' model will be tuned for the best fit for each individual, keeping the parameters in the 'conflict generation' model constant (shown in table 2). This assumes that the conflict generated is same for all individuals and the difference in motion sickness is due to the difference in the accumulation of the conflict.

We make this assumption because the initial stage of conflict generation primarily involves the basic sensory mechanisms that all humans share. However, the degree to which different individuals accumulate and tolerate this conflict can indeed vary widely depending on genetic and psychological factors. People with heightened sensitivity may accumulate conflict more quickly and therefore experience symptoms more intensely. Previous experiences and conditioning can also influence susceptibility. Additionally, individuals who have frequently experienced motion sickness may anticipate symptoms more anxiously, thereby exacerbating the response.

Table 2: SVC model parameters

	Parameter Symbol	Value	Explanation
Anticipation Gains	K_a K_ω	0 0	Fully passive motion assumed
Vestibular Feedback Gains	K_{ac} K_{vc} $K_{\omega c}$	1 5 10	as in Wada et al. (2020)
Visual Feedback Gains	K_{gvis} $K_{\omega vis}$	0 10	VV gain set to zero VR gain as in Wada et al. (2020)
Perception Time Constants	τ τ_{scc}	5 7	as in Liu et al. (2022)

Parameter reduction in the Accumulation model

In addition to studying the accuracy of the Accumulation model with all 5 parameters fitted individually, we reduce the number of individually fitted parameters while ensuring goodness of fit. It is essential to simplify the model for several key reasons: Firstly, simplifying the model enhances generalisability, reducing the risk of over-fitting and ensuring reliability with new data. Secondly, a more simplified model is computationally efficient, saving resources and time, which is vital for real-time applications. Thirdly, simplified models are easier to interpret and help to identify the most significant contributors, allowing for targeted research. Finally, achieving a good fit while reducing parameters balances simplicity and accuracy, creating models that are both scientifically sound and practically applicable. In summary, simplifying parameters enhances generalisability, efficiency, interpretability, and maintains the crucial balance of model performance.
As a starting point, empirical relations and median values of parameters, as observed in previous studies and current simulations are used.

- $K_2 = 5K_1$; from Oman (1990)
- $T_2 = 7T_1$; from Irmak et al. (2020)
- $p = 0.4$; from Irmak et al. (2022)
- $T_1 = 60$ s; from Oman (1990)
- K_1 best fit within datasets (median values from individual estimates)
  - $K_1 = 2$; for Slalom Drive dataset
  - $K_1 = 18$; for Car and Simulator dataset
  - $K_1 = 9$; for Sickness with NDRT dataset

We use combinations of these assumptions to test if the number of parameters can be reduced. Table 3 summarizes the different cases of the accumulation (Oman) model with the parameters that are estimated (marked with a ✓) and the parameters that are held constant or as an empirical relation to the estimated parameters. K_1 and K_2 are the gains, T_1 and T_2 are the time constants and p is the power term in the accumulation model in figure 1. AM5 is the original version of the accumulation model from Oman (1990). AM0 is the model with group-averaged parameters. The a and b version of the AM1 and AM4 are models with same number of parameters (1 for AM1 and 4 for AM4) but have different set of parameters estimated. The K_1 is different for each dataset as the motion captured in each dataset is different. In the 'Slalom Drive' head motion was recorded, however, in the 'Car and Simulator' only vehicle motion was recorded. Hence, there is a difference in the magnitude of rotations which explains the difference in estimated gains (K_1). This could also be due to the difference in average motion sickness susceptibility of the participants in each experiment where the 'Sickness with NDRT' dataset (with $K_1 = 9$) would have less motion sickness susceptible participants than in 'Car and Simulator' dataset (with $K_1 = 18$).

Table 3: Parameter reduction. Simplified versions of the 'conflict accumulation' model with a different number of parameters. Check marks indicate that the parameter is estimated for each individual.

Label	# parameters	K_1	K_2	T_1	T_2	p
AM5	5	✓	✓	✓	✓	✓
AM4a	4	✓	$5K_1$	✓	✓	✓
AM4b	4	✓	✓	✓	$7T_1$	✓
AM3	3	✓	$5K_1$	✓	$7T_1$	✓
AM2	2	✓	$5K_1$	✓	$7T_1$	0.4
AM1a	1	2/18/9	$5K_1$	60	✓	0.4
AM1b	1	2/18/9	✓	60	$7T_1$	0.4
AM0	0	2/18/9	$5K_1$	60	$7T_1$	0.4

Parameter estimation in the accumulation model

To effectively capture individual differences in motion sickness susceptibility with the 'conflict accumulation' model, we have made a concise choice to model both conditions together using a single set of parameters for each individual. This approach not only enhances generalisability but also enables the model to simulate and predict motion sickness across a wide range of conditions. This simplifies the modeling process while maintaining robustness, ultimately making the model more applicable and practical for real-world scenarios.
To fit these parameters a constrained optimization problem is formed. This estimation is carried out in *MATLAB* with the *fmincon* solver using the *sqp* algorithm. In addition to this, *multistart* was used to simultaneously find 16 local minima and then find the lowest out of them. This way the chance that we find the global minimum is improved. The Root-Mean-Squared Error (RMSE) between the actual and predicted MISC responses, as a function of the parameter vector x, is chosen as the cost function for the estimation. As both conditions are fitted together to give one set of parameters, the cost function is the sum of RMSE for both conditions. The optimization problem along with the constraints are shown in equation (1).

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \textstyle\sum_{i=1}^{n} RMS(MISC_{meas} - MISC_{pred}(\boldsymbol{x})) \\ \text{where,} \quad & x = (K_1, K_2, T_1, T_2, p)^T, \\ & n = \text{number of conditions}, \\ & RMS = \text{Root Mean Square} \end{aligned} \tag{1}$$

3 Results

In this paper, we propose a new model framework that combines the 'conflict generation' model with an individualized 'conflict accumulation' model to capture differences in individual susceptibility. We fitted and validated this model using the methods described in the previous section, and the results are presented below. These results aim to justify our goal, i.e., verify the feasibility and accuracy of the new model framework. Our findings indicate that individualizing only the accumulation part of the model may be sufficient for achieving accurate predictions. This implies that customization of the 'conflict generation' part may not be necessary for our applications. By demonstrating that a focus on 'conflict accumulation' model parameters can yield robust results, we present an efficient model framework that accurately captures individual differences in motion sickness susceptibility.
In the first part, results are shown for various datasets to show the performance of the model with varying number of parameters that are estimated. This will be used to determine the optimal number of parameters required for recreation of motion sickness.
In the second part, a probabilistic model for the observed variation in accumulation model parameters across individuals is proposed. This model can facilitate the generation of diverse group of parameter sets with different susceptibility to motion sickness. This enables the simulation of various motion scenarios and observation of the rate of development of motion sickness in different population groups. This approach serves as a benchmark for evaluating motion cueing algorithms in simulators and motion planning algorithms in autonomous vehicles. The insights gained could lead to optimized algorithms that enhance user experience while minimizing motion sickness.

3.1 Part 1: Model Performance & Parameter selection

First, results are presented for the Slalom Drive dataset by Irmak et al. (2020) demonstrating the performance of the model with varying vision conditions. This is followed by results for the Car and Simulator dataset by Talsma et al. (2023) where the adaptability of the model to real-world driving and driving simulators is shown.

Fits of the model framework to the actual MISC responses with Motion Sickness Incidence (MSI) predictions overlayed are shown first, followed by their RMSE with the actual MISC responses. In addition to this, the parameter study is also conducted. The fits and RMSEs are shown for all version of the 'AM' model and this will be used to determine the optimal number of parameters.

3.1.1 Model performance

Figure 5 shows the comparison of the various models with experimental recorded motion sickness responses (MISC) for a representative selection of participants from Irmak et al. (2020)'s experiment for the conditions of internal (left column) and external (right column) vision. The fitting for the rest of the participants in appendix B. Experimentally reported MISC is shown in black and MSI predictions using a Hill function and integration (from Kamiji et al. (2007)) in orange. Model predictions of only 3 out of the 8 models in table 3 are shown here: AM2 in solid green, AM5 in dashed violet and AM0 in dotted grey. The AM0 model and MSI predictions have same parameters for all participants. Any difference in their predictions is solely due to difference in vehicle motion which are used as inputs to these models.
It is evident that our approach of estimating parameters for each individual (in particular for AM2 and AM5 models) offers improved accuracy in predicting MISC responses compared to the use of group-averaged parameters (AM0). The average RMSE reduces from 1.94 (AM0) to 1.13 (AM2) and 0.74 (AM5). This proves that using parameters estimated for each individual (AM2 and AM5) is, 40% and 60% more accurate respectively than using group-averaged parameters, as in the AM0 model.

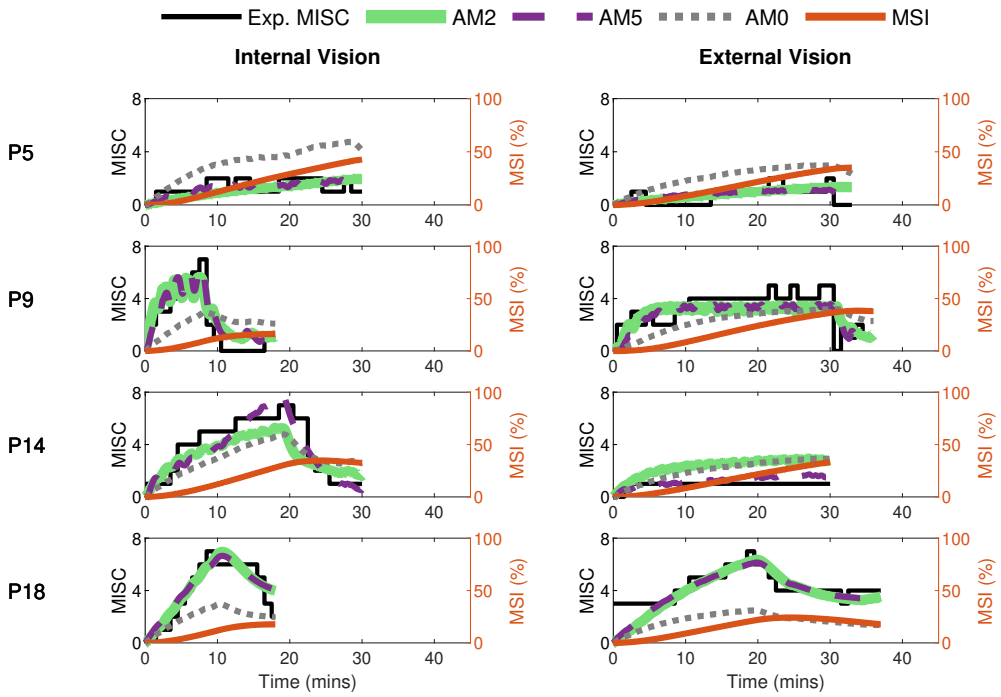


Figure 5: Slalom Drive vehicle tests. Motion sickness responses (MISC) in experiments by Irmak et al. (2020) in black, fitted AM2 model predictions (MISC) in green, fitted AM5 model predictions (MISC) in dashed violet, fitted AM0 model predictions (MISC) in dotted grey, and MSI predictions from Hill function in orange for 4 participants (participant label shown on the left) for the conditions of internal (left column) and external (right column) vision.

Another important observation is that all the accumulation models capture the recovery from motion sickness. This recovery occurs when the sickening stimuli are stopped and the participant is allowed to rest. This is more evident in participants 9 and 14 (second and third row). The Hill function predicting MSI (in orange) is not able to capture this reduction of motion sickness. MSI reports the percentage of people who will get motion sick but does not give information about any individual. Hence the MSI predictions are similar for participants, with some variation due to the actual vehicle acceleration and the moment of stopping.
Furthermore, we validated these models on the car and simulator dataset by Talsma et al. (2023) and sickness with NDRT dataset by Metzulat et al. (2024). The car and simulator dataset had 24 participants, each experiencing motion with external vision in a real-world car and in a simulator. The sickness with NDRT had 20

participants, each experiencing 3 different Non-Driving Related Tasks (NDRT). Figures 6 and 7 shows the experimental recorded motion sickness responses (MISC) in black, MSI predictions in orange, and fitted model predictions in green, violet, and grey for 4 out of the 24 participants for both cases. It can be seen that even with 2 parameters, the fits of the AM2 model are very close to the actual MISC responses in both datasets. The MSI prediction values are very low, especially in the car and simulator dataset. This is due to the low levels of conflict generated in these experiments as compared to the other datasets such as the slalom drive by (Irmak et al., 2020). Another reason may be that the participants sampled in Talsma et al. (2023) may be highly susceptible to motion sickness. The 'conflict Accumulation' model is able to account for these combinations of low levels of conflict and high susceptibility of the participants adequately. While MSI being a population metric is not able to capture this sampling bias.

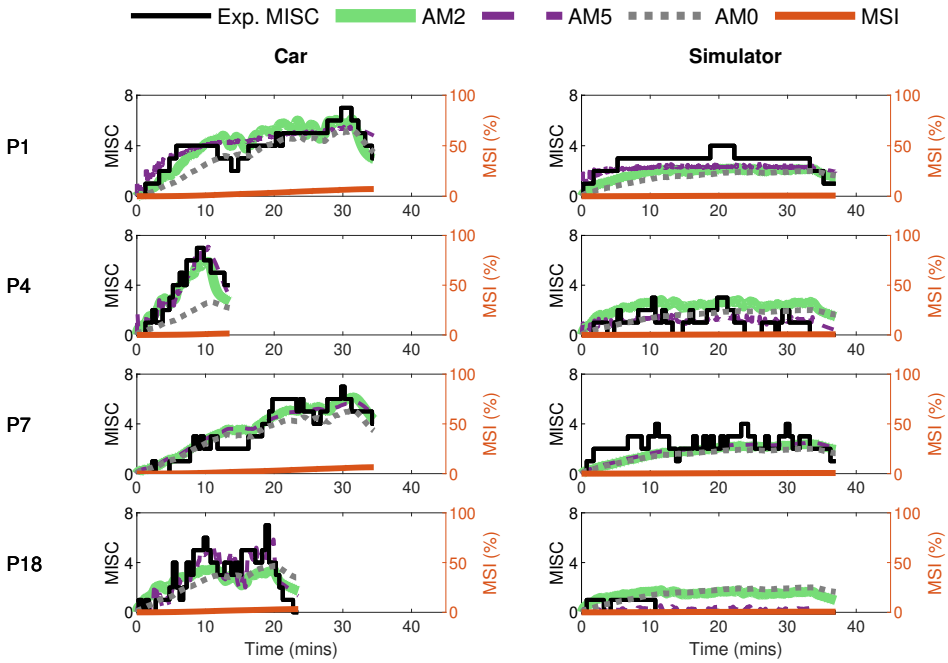


Figure 6: Car and Simulator tests. Motion sickness responses (MISC) in experiments by Talsma et al. (2023) in black, fitted AM2 model predictions (MISC) in green, fitted AM5 model predictions (MISC) in dashed violet, fitted AM0 model predictions (MISC) in dotted grey, and MSI predictions from Hill function in orange for 4 participants (participant label shown on the left) for the case in the car (left column) and the simulator (right column).

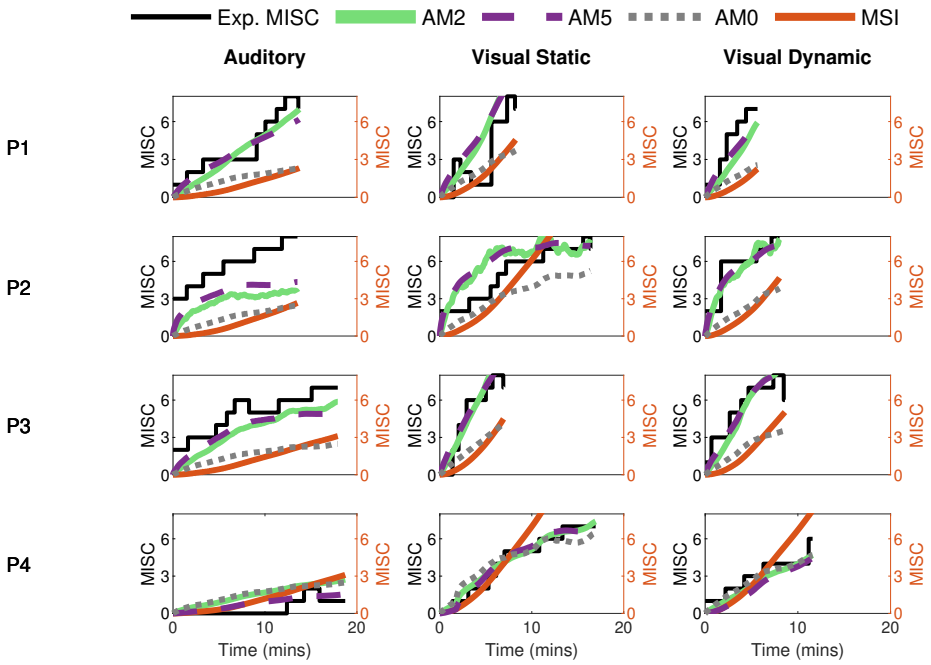
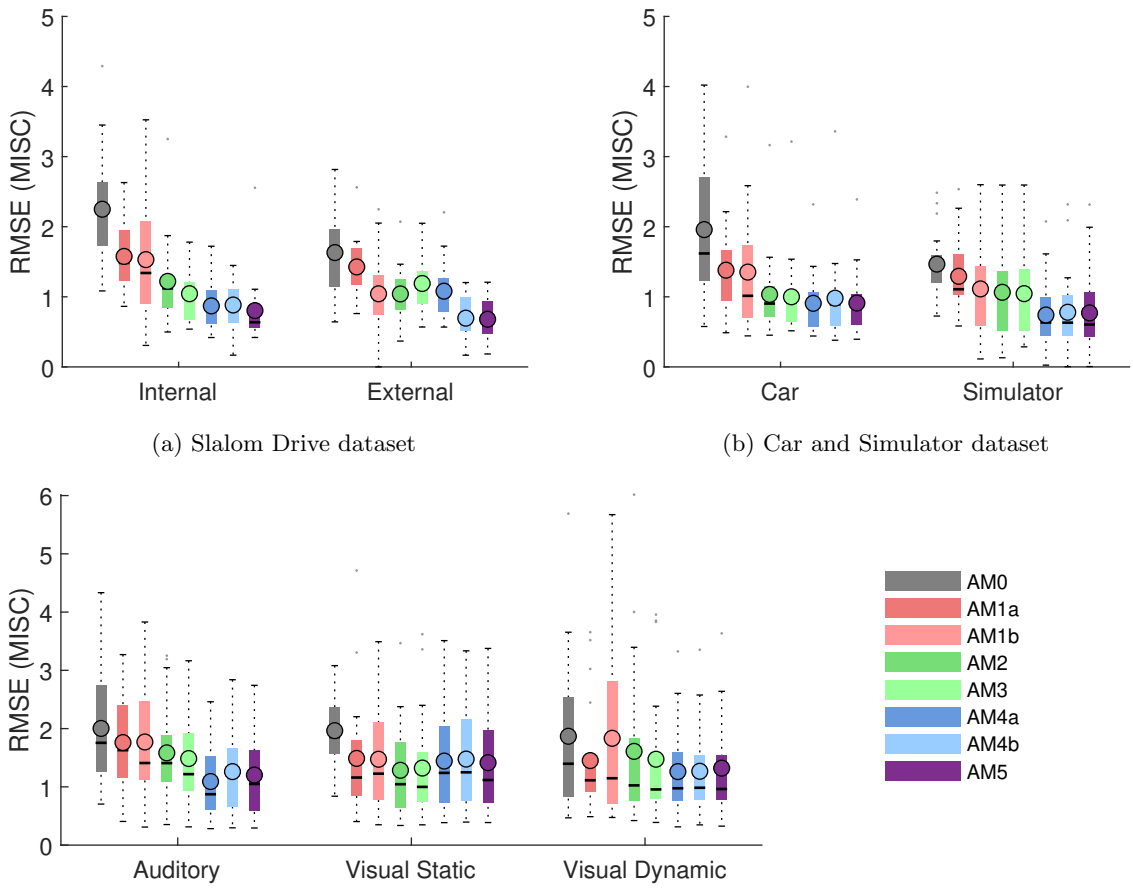


Figure 7: Sicknes with NDRT dataset. Motion sickness responses (MISC) in experiments by Metzulat et al. (2024) in black, fitted AM2 model predictions (MISC) in green, fitted AM5 model predictions (MISC) in dashed violet, fitted AM0 model predictions (MISC) in dotted grey, and MSI predictions from Hill function in orange for 4 participants (participant label shown on the left) for the case in the auditory (left column), visual static (middle column) and the visual dynamics (right column) task.

3.1.2 Parameter selection

In the previous section, it was observed that the model with two parameters (AM2) is able to model the individual responses. To show this quantitatively, we evaluated the need for each of the parameters. We did this by reducing the number of parameters by using relations found during the previous fitting and also what was observed and reported in various studies. We used 5, 4, 3, 2, 1, and 0 parameters to evaluate the need for the parameters, see table 3. From figure 8a it is clear that reducing the number of parameters below 2 leads to a 36% increase in RMSE (from 1.2 to 1.6 in internal vision case and from 1.0 to 1.4 in external vision case when comparing AM2 to AM1a model) in the Slalom drive dataset. Hence, any model with 2 or more parameters is sufficiently accurate to capture the motion sickness development. We also compared the individual models with a group-averaged version of the accumulation model (AM0), where the parameters are the same for all participants in the dataset. It is observed that AM0 has on average 1.7 times more RMSE (for example, from 2.25 to 1.22 in the internal and from 1.63 to 1.04 in the external vision case of the slalom drive dataset) as compared to the AM2 model. Thus, parameters estimated for each individual are more accurate than group-averaged parameters.

Further, to demonstrate that the two parameters are the optimal number of parameters, the previous test was repeated with car and simulator dataset by Talsma et al. (2023). As shown in Figure 8b, there is a 25% increase in RMSE values (from 1.03 to 1.38 in the case of motion sickness in the car and from 1.06 to 1.29 in the simulator) when comparing AM2 to AM1a. When comparing the individual models with a group-averaged version of the accumulation model, it is observed that the group-averaged model, AM0, has 1.64 times more RMSE (from 1.96 to 1.03 in the internal and from 1.47 to 1.06 in the external vision case) as compared to the AM2 model. This increase is close to what was seen in the Slalom Drive dataset where a 1.7 times increase was observed. Thus, parameters estimated for each individual are more accurate than group-averaged parameters. This trend was also seen in the sickness with NDRT dataset by Metzulat et al. (2024), where the AM2 model offers the best balance between performance and the efficiency.



(a) Slalom Drive dataset

(b) Car and Simulator dataset

(c) Sickness with NDRT dataset

Figure 8: Root Mean Squared Error (RMSE) between the predicted MISC from the models and the actual MISC for the Slalom Drive dataset and the Car and Simulator Dataset. Also shown are the mean (circled), median (horizontal solid line), and interquartile range (in colored rectangle)

It is clear from the results in figure 8 that this model framework, is able to capture two different conditions with a single set of 2 parameters. And the accuracy is better (by around 30% in RMSE) than using a group averaged model. This is applicable for various vision conditions, as well as different motions from real cars to simulators.

To show this statistically we used various model selection criteria, such as the Akaike information criterion (AIC) by Akaike (1998) and Bayesian information criterion (BIC) by Schwarz (1978) which take into account the tradeoff between the goodness of fit and the simplicity of the model. In other words, these criterions balance the risk of overfitting and underfitting. Simpler model also means that the simulations will be computationally fast. The model criteria we used are the AIC and BIC which are defined as,

$$AIC = 2k + n\ln(RSS/n)$$

$$BIC = k\ln(n) + n\ln(RSS/n)$$

where, k is the number of parameters, n is the sample size and RSS is the Residual Sum of Square. The model with the lowest score is ideal. The RSS of each model is calculated by summing the RSS for the fits of each participant and each condition. In the model fittings, RMSE was used as a cost function. RMSE is proportional to RSS. RMSE is the square root of the average RSS per observation. Thus, the cost function used in the model fitting, RMSE, directly relates to RSS and the model fitting algorithm would also minimize the model criteria such as the AIC and BIC.

However, we found that the penalty on the number of parameters, $2k$ or $k\ln(n)$, was not enough, which led to the criterion selecting the model with the least error. To solve this problem, we followed Drop et al. (2018) where a similar situation was observed. Here, they increased the gain on the term of k to increase the penalty. We did the same and the modified BIC is defined as follows,

$$mBIC = c\ k\ln(n) + n\ln(RSS/n)$$

where, c is model complexity (number of parameters) penalty parameter which is to be tuned to avoid false-positives while maintaining sensitivity to small yet important contributions. To do this, we calculated mBIC values with c varying from 1 to 5. The variation of the mBIC with c is plotted in figure 9.
Out of all the available values for c (ranging from 1 to 5), we select $c = 2$ for model assessment, as this choice strikes a balance between model complexity and fit. This value allows for a moderate penalty for additional parameters, enabling models to demonstrate predictive power without excessive complexity, hence reducing the risk of overfitting.
Furthermore, we must avoid selecting AM0, which has zero estimated parameters and acts merely as an averaged model. This choice would lead to oversimplification, neglecting crucial dynamics within the data. Additionally, AM5, with its five estimated parameters, poses a high risk of overfitting despite potentially fitting the training data exceptionally well. These two conditions are close to being breached in c values of 3 and 1 respectively, where the models AM5 and AM0 have smaller values of the mBIC metric ad become more favorable than 50% of the other models, which is not desirable.
By favoring AM2 at $c = 2$, we achieve an optimal balance of accuracy and simplicity. This model effectively captures the underlying trends while ensuring robustness in predictions, making it the most suitable choice based on statistical criteria and model performance.

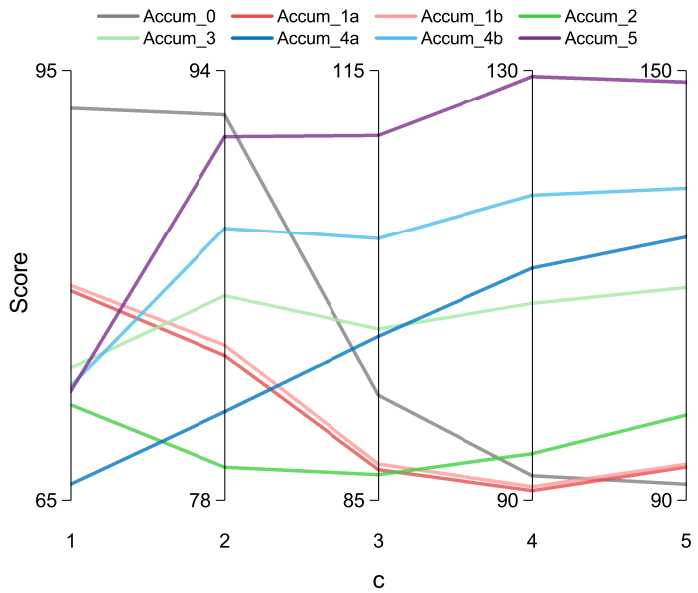


Figure 9: mBIC scores with varying value of c for various models

Table 4 shows all models with the number of parameters and metrics of their fits - RMS, AIC, BIC and mBIC (with $c = 2$). It is clear from the table that though AM4a has the least RMSE, AIC and BIC. However, our new metric, mBIC, which puts more penalty on the number of parameters in the model, chooses AM2 as the optimal model.

Table 4: Model selection criterion

Model	Participants (n)	Parameters (k)	RMSE	AIC	BIC	mBIC ($c = 2$)
AM0	44	0	2.0192	92.334	92.334	92.334
AM1a	44	1	1.6737	77.82	79.604	83.388
AM1b	44	1	1.681	78.203	79.987	83.771
AM2	44	2	1.4652	68.116	71.685	**79.253**
AM3	44	3	1.4451	68.897	74.249	85.602
AM4a	44	4	**1.2628**	**59.032**	**66.169**	81.306
AM4b	44	4	1.3644	65.84	72.977	88.114
AM5	44	5	1.3014	63.68	72.601	91.522

3.2 Part 2: Probabilistic modelling

The parameter sets of 44 participants of the Car and Simulator and motion sickness with NDRT datasets is shown in figure 10 for the AM2 model. Vehicle motion is used as input in both of these datasets. The Slalom drive dataset is omitted here for consistency in parameters due to slalom drive using head motion instead of vehicle motion as is the case in the other two. In the figure, it can be seen that there are three distinct sets of parameters, which are found my using the k-means clustering algorithm:

- High susceptibility are those with high K_1 and low T_1, shown as red dots.
- Low susceptibility have the opposite, low K_1 and high T_1, shown as green dots.
- Medium susceptibility have low K_1 and low T_1, shown as blue dots.

With this knowledge, representative parameters can be sampled to test motion profiles on different motion sickness susceptibilities. Additionally, percentiles can be defined to see which percentile of subjects do or do not get motion sick.

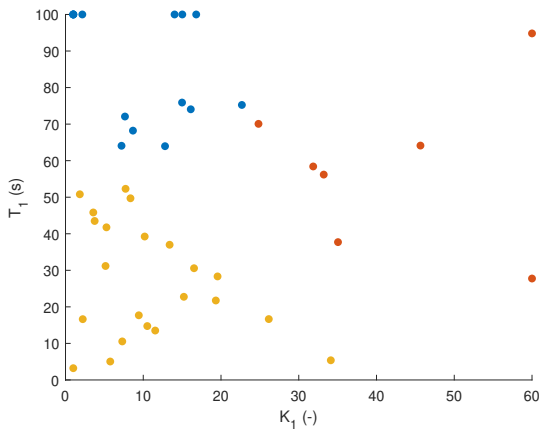
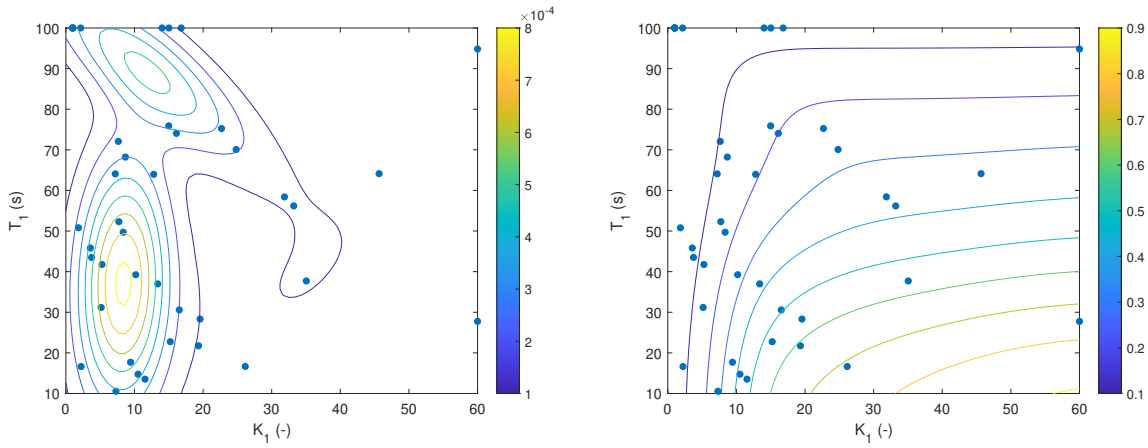


Figure 10: Parameter distribution (estimated gain (K_1) and time constant (T_1)) for the AM2 model.

To do this, we have fitted a three variable probability distribution with the parameter sets (of 44 individuals) which can be seen below. The mean value of the parameters are 15.2 for K_1 and 52.3 for T_1.



(a) Probability density function

(b) Cumulative density function

Figure 11: Three variable probability distribution of the parameter distribution (estimated gain (K_1) and time constant (T_1)) for the AM2 model.

With this PDF, we can sample parameters randomly or based on the probability to get sick (from 0 to 1). Here, figure 12 shows an example of randomly sampled 40 parameter sets overlayed on the actual parameters estimated from the other datasets. The joint probability distribution can be defined as:

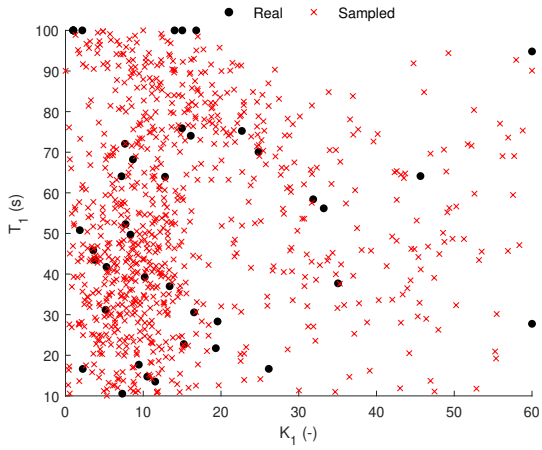
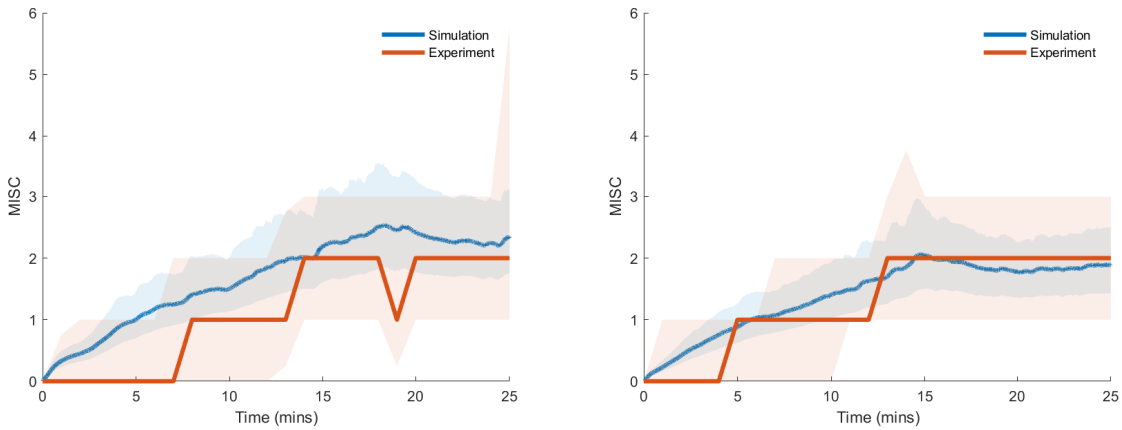


Figure 12: Sampled parameter sets (red cross) from the probability density function of the parameter distribution (estimated gain (K_1) and time constant (T_1)) for the AM2 model (black dots).

To show a use case of this PDF, we simulated these 40 sampled parameter sets on a completely independent dataset with 47 participants. This dataset is collected with passengers in manual driving in a semi-urban environment, and driving the same vehicle in automated mode on a test track with the same participants. The median MISC with the 25th and 75th percentile is shown below in figure 13. This is overlayed with the median, 25th and 75th percentile from the 40 sampled parameter sets. It is observed that the two are highly similar but the simulation somewhat underestimates the sickness variance. This can be seen especially at the start where the experiment data shows some people jumping to MISC of 1 very quickly which is not seen in the model predictions. This is particularly observed at the beginning, where the experimental data shows that some individuals quickly jump to a MISC of 1. This is not predicted by our model.



(a) On Road

(b) On Track

Figure 13: Experimentally reported MISC and Predictions of MISC from sampled parameter sets from the probability density function of the parameter distribution (estimated gain (K_1) and time constant (T_1)) for the AM2 model.

This proves that we can use this probability distribution to predict expected motion sickness levels for untested experiments and scenarios, including those using various motion planning or motion cueing algorithms. Also, this can also be used as a proxy to Motion sickness dose value (MSDV), which is a ISO standard (International Organization For Standardization, 1997) and uses weighted root mean squared acceleration as a measure of motion sickness severity.

4 Discussion

rework the discussion section

add about probabilistic model

This paper introduces and validates a novel combination of models to create a framework to predict an individual's motion sickness level in vehicles and simulators. This model framework combines a group-average 'conflict generation' model with an individual 'conflict accumulation' model to capture individual susceptibility differences. This is important to capture and predict differences in motion sickness susceptibility. Furthermore, by using a 'conflict generation' model that includes visual inputs, various vision conditions (such as external, internal, and only vision) can be simulated as well. This is crucial in simulators where motion sickness occurs due to a strong influence of visual cues. We hypothesized that using a 'conflict accumulation' model with individualized parameters will result in greater accuracy compared to a model that uses group-averaged parameters. Hence, in this paper, we assessed the feasibility and accuracy of this new model approach for motion sickness predictions.

It is clear from the results (see figures 5 and 6) that the Oman model as a 'conflict accumulation' model with individualized parameters is able to better model the motion sickness responses of individuals as compared to using the group-averaged models such as the AM0. In addition to this, it also captures the recovery phase of the experiment (see figures 5 and 6) and, theoretically, the hypersensitivity in the following second motion exposure. This recovery phase takes a few minutes and is not captured by the Hill function, which predicts MSI, due to its time constant of 12 min Kamiji et al. (2007); Liu et al. (2022). These results are in line with the work by Irmak et al. (2020) where individualized fits with the Oman model reduced the prediction error by a factor of 2 in a slalom drive with a frequency of 0.2 Hz and lateral accelerations with a peak amplitude of 0.4 g with eyes closed. Our work extends this, by using a six degrees of freedom 'conflict generation' model to generate the conflict, in realistic driving conditions with varying vision conditions. This way a single set of parameters, estimated using data from all required conditions, can characterize an individual across various motion and vision conditions.

This approach is not limited to the SVC model which we used as the 'conflict generation' model. Any other 'conflict generation' model can be used such as any of the version of SVC by Kamiji et al. (2007); Liu et al. (2022); Wada et al. (2020) or the Multi-Sensory Observer Model by Clark et al. (2019).

We also looked into reducing the number of parameters in the model. We found that using the relations and values mentioned in previous studies (such as in Oman (1990); Irmak et al. (2020, 2022)), the number of individual parameters can be limited to 2. This way we can be sure that the parameter set estimated is unique for each individual and is not an over-fit on the available sparse (in conditions) dataset. Another by-product is that the computation time required for the estimation of parameters is significantly reduced (by a factor of 4, from 48 to 11 seconds for 40 min of simulation). However, even with 5 parameters the model can be used for real-time applications.

Further we created a distribution of the parameter sets which can be used to sample an any number of parameters (figure 12 and simulate the motion profile to get the expected motion sickness levels. This has been demonstrated in figure 13 where the simulated sickness predictions closely match the experimentally observed median sickness levels. The difference in spread of sickness is likely due to the selection of participants with extremely high or low sickness susceptibility which is sometimes not correctly reported in self-reported Motion Sickness Susceptibility Questionnaire. Additionally, at the start of the experiments, some of the participants quickly reported a MSIC of 1. This is not captured by our models, but instead, due to the inherent nature of the models, there is a smooth increase in MISC over time. If this is instead discretized as is reported by the participants, the sickness predictions match even better and the spread of MISC is completely inside the experimentally reported MISC (see appendix C). This proves that we can use this probability distribution to predict expected motion sickness levels for untested experiments and scenarios, including those using various motion planning or motion cueing algorithms. Also, this can also be used as a proxy to Motion sickness dose value (MSDV), which is a ISO standard (International Organization For Standardization, 1997) and uses weighted root mean squared acceleration as a measure of motion sickness severity. This has been known to not accurately capture the dynamic nature of motion sickness especially the quick recovery from sickness which is not captured by MSDV due to its monotonous nature and lack of any leakage term.

Using these models, the motion profiles in simulator experiments can be tuned to reduce dropout of participants in simulator experiments due to motion sickness. Not only is the model useful in designing experiments but also can be employed in motion cueing algorithms in driving simulators, as well as in automated vehicles. Using a model-based control method, this model framework can be included in the plant model to forecast motion sickness levels, as in Jain et al. (2023). This way, motion sickness levels can be controlled by taking into account each individual's susceptibility. Vehicle motion in automated vehicles and platform motion and tilt coordination in simulators can be optimized to minimize their effect in eliciting motion sickness. Thus, this model is feasible to use both post-experiment and in real-time during the experiment.

We now tuned individual parameters of the accumulation model while keeping the many parameters of the sensory conflict model constant. A main limitation of the model is that it fails to accurately fit the subject's responses who get highly motion sick during internal vision and do not get sick during the external vision case (for example P14 in figure 5). This sharp shift in motion sickness dynamics cannot be captured by our model framework. To do this, at least one additional parameter needs to be estimated, which could be the vision gain in the 'conflict generation' model or other perception parameters. People not only differ in motion sickness

susceptibility, but also in their perception of motion. This has been shown in Irmak et al. (2023), where a correlation coefficient of 0.74 was observed between individuals' overall sickness sensitivity and their subjective vertical time constant. Thus each human will have a different contribution of visual and vestibular signals for their state estimation. By adjusting the parameters in the 'conflict generation' model, the contribution of vision to the estimates can be tuned.
Another limitation is that the model gain, K_1, needs to be adapted for the two datasets, and presumably depending on the location of the IMU (head or seat). This can be seen in the difference in the range of estimated gain (K_1) values in the two datasets (see **????**). Also, the Hill function accumulation model predicted a reasonable MSI magnitude in the Slalom Drive but highly underestimated sickness in the Car and Simulator dataset. Ideally, we would like to always use recorded head motion, which is more representative of the motion experienced by the vestibular system. If this is not available we could use biomechanical human/seat models or linear transfer functions to convert 6D vehicle motion to 6D head motion. This way it can be ensured that the estimated parameters can be used for any case.

5 Conclusion

rework the conclusion section

add about probabilistic model

A new model framework was developed to fit and predict individual motion sickness levels in vehicles and simulators. This model framework contained two parts: 'conflict generation' and 'conflict accumulation'. The model utilizes acceleration and angular rotational data as inputs and adjusts the parameters of the 'conflict accumulation' model to fit each individual's motion sickness level (reported using MIsery SCale (MISC)) response. This framework is able to fit individual responses with just two estimated parameters (a gain K_1 and time constant T_1) with an average RMSE of 1.1 MISC. By combining the two models, better fits, on average by a factor of 1.7 (for AM2 model), of an individual's motion sickness levels are achieved as compared to using only the group-averaged models. We fit both conditions (External and Internal vision conditions for the Slalom Drive dataset and driving in a car and a simulator in the Car and simulator dataset) together to estimate one set of parameters. The model is able to simulate both conditions using a single set of estimated parameters for each participant. This framework has the potential to show the impact of vision on an individual level, offering a more personalized approach compared to previous models. This is important especially in driving simulators, as the effect of delays in visual cues, can be quantified with this model. This model framework also shows robustness by accurately modeling different datasets with completely different motion and vision conditions. These individual models can be used in further studies for predicting motion sickness levels of the same individual in a different motion scenario, thus reducing the dependence on extensive experiments on humans. Finally, this model framework can also aid in faster testing of new motion cueing algorithms for driving simulators due to less sickening driving simulator experiments and may reduce the dropout of participants in driving simulator evaluations.

Acknowledgements

We thank Tugrul Irmak (Delft University of Technology, The Netherlands), Tessa Talsma (Skoda Auto, Czechia) and Myriam Metzulat (Wurzburg Institute for Traffic Sciences (WIVW GmbH)) for sharing the experimental data. The contribution of Varun Kotian was financially supported by Toyota Motor Europe.

References

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. *Biogeochemistry*, 1998:199–213.

Baumann, G., Jurisch, M., Holzapfel, C., Buck, C., and Reuss, H.-C. (2021). Driving simulator studies for kinetosis-reducing control of active chassis systems in autonomous vehicles. In *Driving Simulation Conference Proceedings*, pages 51–58, Munich, Germany. Driving Simulation Association.

Bos, J., Kuiper, O. X., and Schmidt, E. A. (2020). Motion Predictability and Sickness. In *Driving Simulation Conference Proceedings*, pages 179–182, Antibes, France. Driving Simulation Association.

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

Bos, J. E. and Bles, W. (1998). Modelling motion sickness and subjective vertical mismatch detailed for vertical motions. *Brain Research Bulletin*, 47(5):537–542.

Clark, T. K., Newman, M. C., Karmali, F., Oman, C. M., and Merfeld, D. M. (2019). Mathematical models for dynamic, multisensory spatial orientation perception. *Progress in Brain Research*, 248:65–90.

Cleij, D., Venrooij, J., Pretto, P., Pool, D. M., Mulder, M., and Bülthoff, H. H. (2018). Continuous subjective rating of perceived motion incongruence during driving simulation. *IEEE Transactions on Human-Machine Systems*, 48(1):17–29.

De Winkel, K. N., Tuğrul Irmak, Â., Kotian, V., Daan, Â., Pool, M., and Happee, Â. R. (2022). Relating individual motion sickness levels to subjective discomfort ratings. *Experimental Brain Research 2022*, 1:1–10.

Drop, F. M., Pool, D. M., Van Paassen, M. M., Mulder, M., and Bülthoff, H. H. (2018). Objective Model Selection for Identifying the Human Feedforward Response in Manual Control. *IEEE transactions on cybernetics*, 48(1):2–15.

Golding, J. and Markey, H. M. (1996). Effect of frequency of horizontal linear oscillation on motion sickness and somatogravic illusion. *Aviation Space & Environmental Medicine*, 67(2):121–126.

Griffin, M. J. and Mills, K. L. (2002). Effect of frequency and direction of horizontal oscillation on motion sickness. *Aviation, space, and environmental medicine*, 73(6):537–543.

Hogerbrug, M., Venrooij, J., Pool, D. M., and Mulder, M. (2020). Simulator Sickness Ratings Reduce with Simulator Motion when Driven Through Urban Environments. In *Driving Simulation Conference Proceedings*, pages 175–178, Antibes, France. Driving Simulation Association.

Howarth, H. V. C. and Griffin, M. J. (2003). Effect of roll oscillation frequency on motion sickness. *Aviation, space, and environmental medicine*, 74(4):326–331.

International Organization For Standardization (1997). ISO 2631-1:1997 - Mechanical vibration and shock - Evaluation of human exposure to whole-body vibration - Part 1: General requirements.

Irmak, T., De Winkel, K. N., Pool, D. M., Bülthoff, H. H., and Happee, R. (2021). Individual motion perception parameters and motion sickness frequency sensitivity in fore-aft motion. *Experimental Brain Research*.

Irmak, T., Kotian, V., Happee, R., de Winkel, K. N., and Pool, D. M. (2022). Amplitude and Temporal Dynamics of Motion Sickness. *Frontiers in systems neuroscience*, 16.

Irmak, T., Pool, D., de Winkel, K., and Happee, R. (2023). Validating models of sensory conflict and perception for motion sickness prediction. *Biological Cybernetics*.

Irmak, T., Pool, D. M., and Happee, R. (2020). Objective and subjective responses to motion sickness: the group and the individual. *Experimental Brain Research*.

Jain, V., Kumar, S. S., Papaioannou, G., Happee, R., and Shyrokau, B. (2023). Optimal Trajectory Planning for Mitigated Motion Sickness: Simulator Study Assessment. *IEEE Transactions on Intelligent Transportation Systems*.

Kamiji, N., Kurata, Y., Wada, T., and Doi, S. (2007). Modeling and validation of carsickness mechanism. In *Proceedings of the SICE Annual Conference*, pages 1138–1143. IEEE.

Kolff, M., Schwienbacher, M., Venrooij, J., Pool, D. M., and Mulder, M. (2022). Motion Cueing Quality Comparison of Driving Simulators using Oracle Motion Cueing. In *Driving Simulation Conference Proceedings*, Strasbourg.

Kotian, V., Irmak, T., Pool, D., and Happee, R. (2024). The role of vision in sensory integration models for predicting motion perception and sickness. *Experimental Brain Research*, 242(3):685–725.

Liu, H., Inoue, S., and Wada, T. (2022). Motion Sickness Modeling with Visual Vertical Estimation and Its Application to Autonomous Personal Mobility Vehicles. In *2022 IEEE Intelligent Vehicles Symposium (IV)*, pages 1415–1422.

McCauley, M., Royal, J., Wylie, C., O'Hanlon, J., and Mackie, R. (1976). Motion Sickness Incidence: Exploratory Studies of Habituation, Pitch and Roll, and the Refinement of a Mathematical Model. Technical report, Canyon Research Group Inc Goleta Ca Human Factors Research Div.

Metzulat, M., Metz, B., Landau, A., Neukum, A., and Kunde, W. (2024). Does the visual input matter? Influence of non-driving related tasks on car sickness in an open road setting. *Transportation Research Part F: Traffic Psychology and Behaviour*, 104:234–248.

Oman, C. M. (1990). Motion sickness: a synthesis and evaluation of the sensory conflict theory. *Canadian Journal of Physiology and Pharmacology*, 68(2):294–303.

Reuten, A. J. C., Bos, J., and Smeets, J. B. J. (2020). The metrics for measuring motion sickness. In *Driving Simulation Conference Proceedings*, pages 183–186, Antibes, France. Driving Simulation Association.

Schwarz, G. (1978). Estimating the Dimension of a Model. *https://doi.org/10.1214/aos/1176344136*, 6(2):461–464.

Talsma, T. M., Hassanain, O., Happee, R., and de Winkel, K. N. (2023). Validation of a moving base driving simulator for motion sickness research. *Applied Ergonomics*, 106:103897.

Wada, T., Kawano, J., Okafuji, Y., Takamatsu, A., and Makita, M. (2020). A Computational Model of Motion Sickness Considering Visual and Vestibular Information. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1758–1763. IEEE.

A Appendix A: Motion stimuli in each of the datsets

add imu data plots

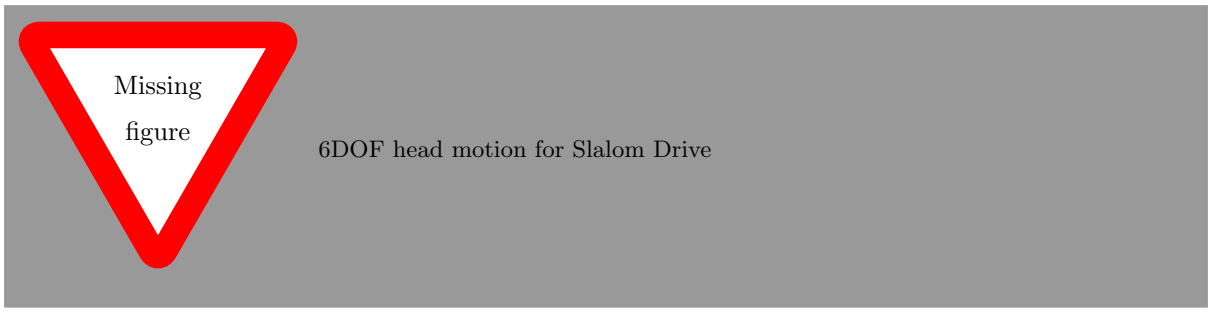


Figure 14: Slalom Drive dataset (Irmak et al., 2020) 3D accelerations and 3D rotations

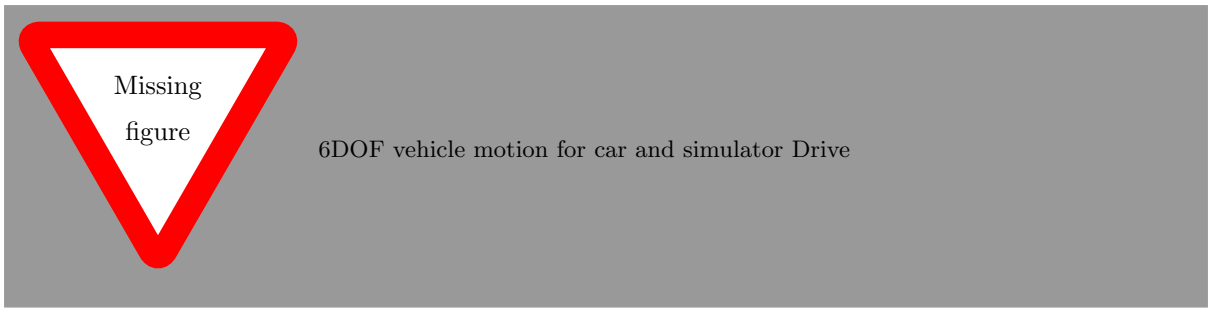


Figure 15: Car and Simulator dataset (Talsma et al., 2023) 3D accelerations and 3D rotations

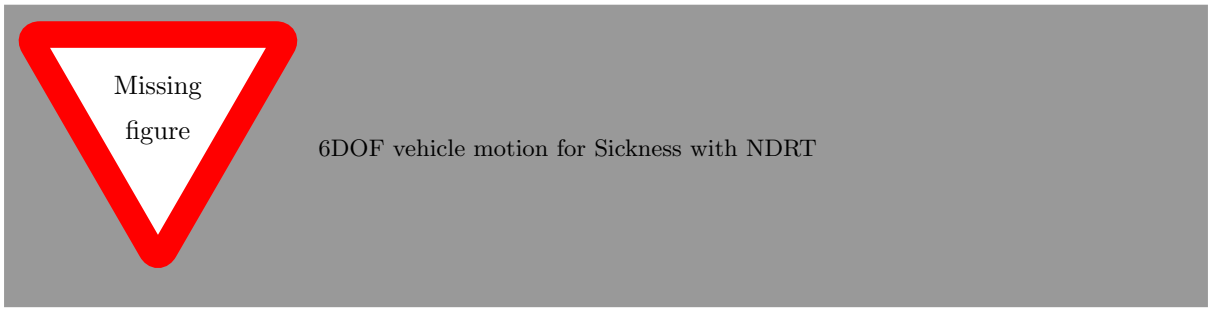
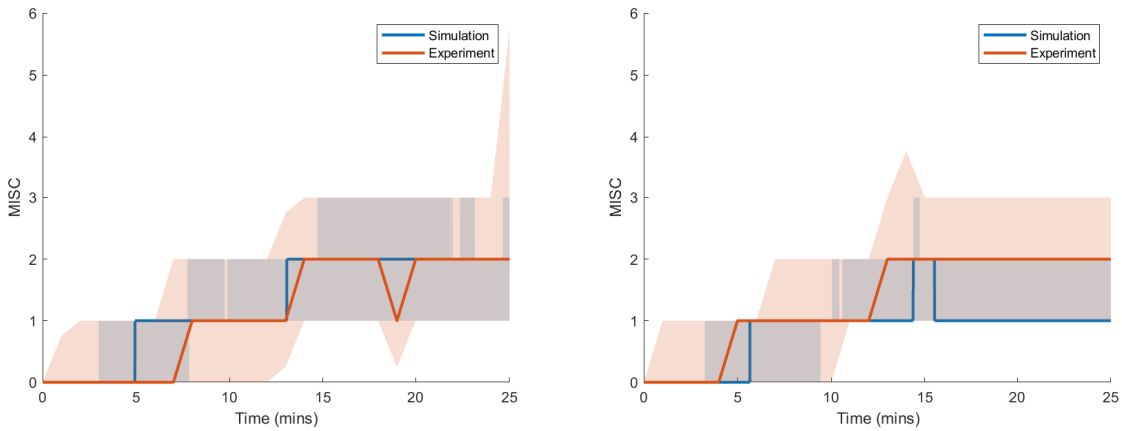


Figure 16: Sickness with NDRT dataset (Metzulat et al., 2024) 3D accelerations and 3D rotations

B Appendix B: Fitting results for all data

add plots of fitting of all participants

C Appendix C: Dicretized Predictions



(a) On Road

(b) On Track

Figure 17: Experimentally reported MISC and Predictions of MISC from sampled parameter sets from the probability density function of the parameter distribution (estimated gain (K_1) and time constant (T_1)) for the AM2 model.