



Karl-Johan Larsson

Autoliv Research,
Vårgårda 447 83, Sweden;
Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Gothenburg 412 96, Sweden
e-mail: karl-johan.larsson@autoliv.com

Jonas Östh

Volvo Cars,
Gothenburg 405 31, Sweden;
Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Gothenburg 412 96, Sweden
e-mail: jonas.osth@volvocars.com

Johan Iraeus¹

Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Gothenburg 412 96, Sweden
e-mail: johan.iraeus@chalmers.se

Bengt Pipkorn

Autoliv Research,
Vårgårda SE-44783, Sweden;
Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Gothenburg 412 96, Sweden
e-mail: bengt.pipkorn@autoliv.com

A First Step Toward a Family of Morphed Human Body Models Enabling Prediction of Population Injury Outcomes

The injury risk in a vehicle crash can depend on occupant specific factors. Virtual crash testing using finite element human body models (HBMs) to represent occupant variability can enable the development of vehicles with improved safety for all occupants. In this study, it was investigated how many HBMs of different sizes that are needed to represent a population crash outcome through a metamodel. Rib fracture risk was used as an example occupant injury outcome. Morphed HBMs representing variability in sex, height, and weight within defined population ranges were used to calculate population variability in rib fracture risk in a frontal and a side crash. Two regression methods, regularized linear regression with second-order terms and Gaussian process regression (GPR), were used to metamodel rib fracture risk due to occupant variability. By studying metamodel predictive performance as a function of training data, it was found that constructing GPR metamodels using 25 individuals of each sex appears sufficient to model the population rib fracture risk outcome in a general crash scenario. Further, by utilizing the known outcomes in the two crashes, an optimization method selected individuals representative for population outcomes across both crash scenarios. The optimization results showed that 5–7 individuals of each sex were sufficient to create predictive GPR metamodels. The optimization method can be extended for more crashes and vehicles, which can be used to identify a family of HBMs that are generally representative of population injury outcomes in future work. [DOI: 10.1115/1.4064033]

Introduction

Vehicle occupant epidemiology indicates greater fatality and injury risks for female, obese, and elderly occupants [1–5]. While it has been indicated that the increased risk of several injuries for females is due to differences in crash and vehicle types [6], better understanding of the effect of occupant variability in the design of safety systems can enable development of protection systems that further reduce the number of occupant fatalities and injuries.

Vehicle occupant injury risk has traditionally been evaluated using anthropomorphic test devices (ATDs) in physical crash tests and virtual simulations of crash tests. Crash test simulations are common during the vehicle development phase as it reduces the need for physical prototypes and speeds up the development process. Physical crash testing is mandated in regulations and applied by consumer information organizations, such as new car assessment programs, providing publicly available safety ratings of vehicles.

The population of adult vehicle occupants in crash testing is represented by three standard ATD sizes, representing a fifth percentile female, a 50th and a 95th percentile male. These percentiles refer to male and female height and weight distributions

of the 1970s U.S. population [7]. ATDs are mechanical representations of humans and use sensor measurements to estimate body region injury risks.

To provide more detailed insights into underlying injury mechanisms, finite element human body models (HBMs) are used as complements to ATDs in crash simulations. Detailed HBMs used for occupant safety simulation, such as the total human model for safety (THUMS) [8], GHBM [9], and SAFER HBM (SHBM) [10] represent the human anatomy at a higher level of detail than the ATDs. For instance, in the HBMs all ribs in the ribcage are modeled in detail while in the ATDs the ribcage is represented with a few ribs. Therefore, the HBMs have the potential to evaluate injury at the tissue level by predicting local physical measurements such as stress, strain, and pressure, that are related to the actual physical injury mechanism. As an example, SHBM (v.9) was validated for predicting strain in the rib cortical bone and a strain-based rib fracture risk [11,12]. Following the ATD size convention, the THUMS and GHBM models exist in three versions representing individuals of the same height and weight as the standard ATD sizes, while the SHBM is only available in one size, representing a 50th percentile male. To enable representation of a variation of anthropometries with HBMs, mesh morphing methods, that modify the geometry described by the HBM finite element mesh, have seen increased use in recent years and have been implemented in frameworks such as PIPER [13] and parametric HBM morphing

¹Corresponding author.

Manuscript received May 31, 2023; final manuscript received October 3, 2023; published online January 29, 2024. Assoc. Editor: Matthew B. Panzer.

[14,15]. Studies including validation of morphed HBMs (MHBMs) have demonstrated a potential to represent anthropometric occupant variability in impact conditions [16–18].

Considering the statistics of injury risk differences across the population and the potential in using HBMs to represent occupant variability, the consumer information organization Euro NCAP has communicated their intention to include virtual crash testing with HBMs to represent occupant diversity in vehicle safety evaluations [19]. However, how many HBMs, and of which anthropometries has not yet been specified. A potential approach, investigated in this study, can be to select a number of anthropometries which are representative of population injury risk outcomes due to the variability existing in the general population. In this approach, a prerequisite is knowledge of the injury outcome across the population.

As crash simulations with detailed finite element HBMs are computationally expensive (hours to days of multicore calculations per simulation depending on the specific scenario and computing resources), evaluating the crash outcome for a population of occupants through direct computation can become prohibitively time consuming. To aid in analysis, metamodeling based on computer simulation results (also referred to as response surface methodology, surrogate modeling, emulators) has been proposed as a viable methodology for crash simulations with MHBMs [20–22]. In this context, a metamodel is a mathematical model constructed to reproduce some input–output relationship of the crash simulations but is substantially faster to evaluate than the crash simulations it is built upon. Constructing a metamodel requires parametrization of the crash simulation model to create input parameters that modify the details of the simulation, running a set of simulations at different combinations of inputs (i.e., input points) to generate known input–output relationships (training data), and selecting and fitting a metamodel to the training data.

As an example, Ref. [20] used 61 combinations of six parameters to generate training data for driver injury risks in frontal impacts. The driver sled model represented a driver environment and was equipped with a steering wheel airbag and a seatbelt. Parameters described variation in impact speed, seatbelt load limit force, seat position, seat back angle, and height and body mass index (BMI) of the occupant. Height and BMI variations were achieved by morphing the 50th percentile male THUMS. Second-order polynomial regressions were used as metamodels for different MBHM injury metrics, as they had the lowest error for predicting the training data (training error) compared to other evaluated functional forms [20].

However, the training error can be optimistic in measuring the performance of a metamodel for two reasons. First, the metamodel is often constructed through algorithms minimizing the training error, and perfectly fitting the training data is achievable for flexible models. But the purpose of a metamodel is to predict the output of the crash simulation in new points. The error for prediction in new points, not part of the training data (test error), is a more appropriate metric for judging metamodel quality [23]. Second, crash simulation outputs contain numerical noise. While they are deterministic—running the same crash simulation twice on the same computer will produce the same results [24]—small changes in input parameters can spuriously lead to unexpectedly large changes in output, due to finite precision numerical calculations. In the presence of noise, it is possible to construct metamodels predicting the noise, rather than underlying systematic trends in the output [25]. This leads to models with small training errors that does not predict the response well in new input points (large test errors), which is a phenomenon known as overfitting.

To estimate the test error, a portion of the data can be held out before generating the metamodel, at the cost of a reduced training set. A technique to enable use of all data in the metamodel construction, while obtaining an estimate of the test error is k -fold cross-validation (CV), where the training data are divided into k groups. One of the k groups is taken as the test data, and the metamodel is constructed using the remaining $k - 1$ groups, and a

test error is calculated for the held-out set. This is repeated k times, such that all groups have been used as a test set, and the CV test error estimate is the average over the k iterations [23].

Reference [21] used the ten-fold CV error to compare different regression methods as metamodels for a combined MHBM injury metric in a frontal impact scenario. Special attention was given to the effect of tuning metamodel hyperparameters (i.e., parameters affecting the metamodeling method itself) on the CV error. A male GHBMC model representing either a 50th percentile male, or as morphed to an obese male represented variation in BMI. Fifteen additional parameters represented various safety system configurations. Evaluated methods included linear regression with second-order polynomial terms with and without regularization, regression forests, support vector regression, and neural network regression. Among these methods, the linear regression with second-order terms regularized with least absolute shrinkage and selection operator (LASSO) [26] had the lowest CV error over a fixed set of 450 training points.

A drawback with using a fixed training set size for metamodel construction is that it is not known if the train and test errors have converged, or if additional training points could substantially improve metamodel fit. Due to the inherent numerical noise in crash simulation outputs, which can have different magnitudes for different crash scenarios, depending on the specific models involved and computational settings [24], there exists no general error thresholds for an acceptable metamodel fit. For constructing metamodels of lateral head excursion and rib fracture risk from a morphed 50th percentile male GHBMC, Ref. [22] used a batch-sequential approach to metamodel construction. The crash simulation included variation in height, weight, and waist circumference of the MHBMC, and variation in seatbelt load limiting force and impact configuration of the vehicle. The neural network regression based metamodels predicted the outputs in 45 new test points. The outputs were then obtained in these points, enabling calculation of test errors for the metamodels constructed from the previous batches. The sequential calculation of metamodel errors allows to study if metamodel error magnitudes converge for some training set size, which can be used in the judgment if sufficient training data have been generated.

In the broader context of metamodeling based on computer simulation output, Gaussian process regression (GPR), also known as kriging, has been widely used [27,28]. Previous applications include biomechanical finite element simulations [29,30] and HBM simulations [31,32]. Its applicability for metamodeling of MHBMC outputs has not yet been evaluated, but it appears promising as it is a nonparametric method (i.e., it does not assume a functional form for the output) with relatively few hyperparameters that have likelihood formulations, which simplifies the model fitting process.

While metamodeling has been demonstrated as a viable approach to enable prediction of crash outcomes for MHBMs representing anthropometric variability, only male anthropometric variability has been explored. Further, as a general limitation, metamodels can only make predictions involving the parameters existing in the training data. In a vehicle development phase, where new prototype designs are invented based on evaluation outcomes of current designs, metamodeling has limited utility as unknown future features cannot be included in the parametrization. In this context, knowledge of a limited number of HBM anthropometries that can be used to efficiently generate a metamodel representative of population variability in crash outcomes enables more rapid evaluations.

Therefore, the aim of this study was to investigate how many, and which, MHBMC anthropometries that are needed to represent a population crash outcome. As an example outcome, occupant rib fracture risk was used in this study, as rib fractures remain a prevalent occupant injury in real-world crashes [5] that crash testing with ATDs have limited capability to predict [33–35]. Thus, virtual evaluation of rib fracture risk with MHBMs has the potential to enable development of protection systems that contribute to rib fracture injury risk reduction for the population of vehicle occupants.

Methods

The methods are divided into three main steps. A brief overview is given here, while detailed methods are provided in the below methods subsections.

First, metamodel training data representing rib fracture risk outcomes in front and near-side impacts for a population were generated. The population was defined to include 90% of U.S. males and females, in terms of height and weight ranges. Male and female samples were selected, and for each sample, a corresponding MHBM was created by morphing the SHBM to the corresponding sex, height, and weight. For each MHBM, the risk to sustain two or more fractured ribs (NFR2+) was computed in two different crash scenarios, representing a frontal impact and a side impact using generic occupant compartment models.

Second, using the frontal and side impact NFR2+ results, two different regression methods were evaluated for their capability of metamodeling the population NFR2+ risk in each impact scenario, by studying convergence of each metamodel training and test errors for increasing training set sizes. Here, training error refers to the error between metamodel prediction of NFR2+ and the corresponding MHBM NFR2+ outputs in the training data. Test error refers to the errors between metamodel NFR2+ predictions and corresponding MHBM NFR2+ outputs in a held-out test set. Metamodels for both sexes combined, as well as separate models for each sex were evaluated. The metamodeling technique that had the overall lowest errors for predicting NFR2+ in the held-out test set was selected for the next step.

Third, with the selected metamodeling technique, it was investigated how many, and which, MHBMs within the training sample that should be used to represent the population NFR2+. In this context, to represent the population meant that the metamodel constructed using only the training data represented by these MHBMs, had the lowest error for prediction of NFR2+ across the entire population, averaged over both impact scenarios.

Generating the Population Rib Fracture Risk Training Data. The population to be represented by the MHBMs was defined to include 90% of adult males and females in the U.S., in terms of height and weight, based on the 2013–2016 National Health and Nutrition Examination Survey (NHANES) data, Fig. 1. Height and weight ranges were defined by 90% probability regions, calculated using the NHANES weighting factors [36]. Briefly, a 90% confidence region of a bivariate normal distribution constructed

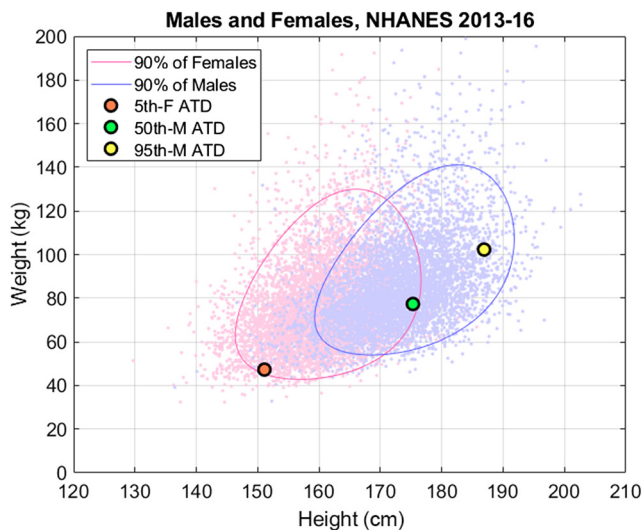


Fig. 1 Height and weight ranges including 90% of U.S. males (blue) and females based on the 2013–2016 NHANES. Height and weight specifications of fifth percentile female, 50th, and 95th percentile male ATDs [7] indicated for comparison. (Color version online.)

using the correlation between height and weight, centered at the mean height and weight, was computed for each sex using a custom MATLAB script (R2017b, The MathWorks, Inc., Natick, MA). The logarithm of weight was used due to right-skewed weight distributions [36]. The three standard ATD sizes are included in Fig. 1 for reference.

A sequential space filling strategy with room for 500 samples of height and weight within each sex was used to sample MHBM sizes, as it was unknown how many samples were needed for converged NFR2+ metamodel test errors. As the sample regions were elliptical, the starting point was an existing space-filling design of 500 points within a two-dimensional unit disk [37]. Here, the two-dimensional coordinates within the disk corresponded to normalized height and $\log(\text{weight})$. These 500 samples were then subsampled by a sequential strategy that started from the sample closest to the disk center, and then sequentially chose the next sample as the one furthest away from the already selected samples. The samples were then mapped to the corresponding region defined for each sex in Fig. 1. This enabled a sequence of height and weight samples that, for every sample size, were spread out within the regions for each sex. Based on previous MHBM metamodeling studies [21,22], an initial sample size of 400 MHBMs (200 male, 200 female) was used, and the sequential sampling strategy allowed to sequentially add additional samples, up to 300 more of each sex.

Morphing to Generate Morphed Human Body Models. The SHBM (v.10) [10] was the baseline HBM used for morphing. The parametric HBM morphing method [15] was used for morphing of the SHBM to the sex, height, and weight targets selected in the sampling. This morphing method uses statistical human shape models of body surface, ribcage, pelvis, femur, and tibia, all parametrized with respect to sex, age, height, and weight, to describe the target geometry for the corresponding HBM parts. The use of human shape statistical models ensured that the resulting MHBM geometries represented human shape trends with respect to sex, height, and weight, both in external body shape and for skeletal structures. The age parameter was fixed at 45 years for all MHBMs in this study. As sex does not influence rib cortical bone material properties [38], male and female MHBMs used the same rib material properties. Similarly, the rib cortical bone thickness was not adjusted for sex, as male and female thicknesses are not significantly

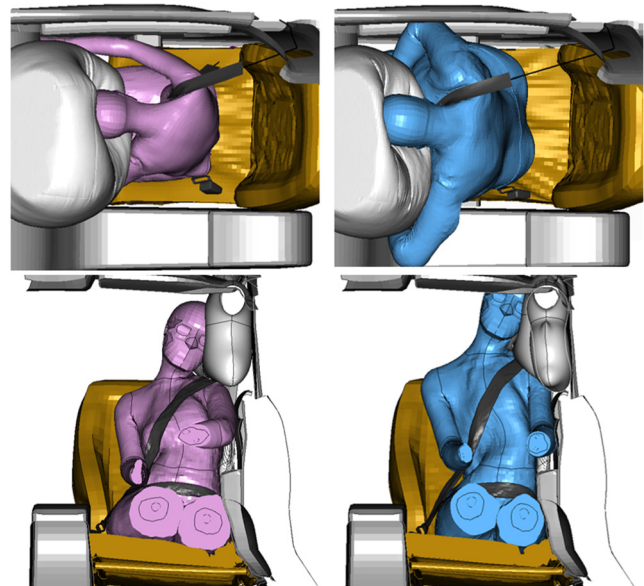


Fig. 2 Examples of frontal and near-side crash simulation with MHBMs. Top row: frontal impact at 108 ms. Left: female 165 cm, 58 kg. Right: male 181 cm, 130 kg. Bottom row: near-side impact at 54 ms, parts hidden for visibility. Left: female 162 cm, 74 kg. Right: male 188 cm, 79 kg.

different below 55 years of age [39]. Morphed versions of SHBM (v.9) were previously validated using postmortem human subject tests [18] and updated validation results of morphed SHBM (v.10) in frontal and side impacts are presented in (See [Supplemental Materials](#) on the ASME Digital Collection, Appendix A).

Occupant Crash Simulations for NFR2+ Risk. A frontal and a near-side lateral impact crash simulation were carried out with each MHBM seated as a front row passenger in generic vehicle interior models (Fig. 2). The interior models were based on the models used by Pipkorn et al. [12] and Iraeus and Lindquist [40]. Descriptions of the interior model updates and validations are provided in (See [Supplemental Materials](#), Appendix B).

Assuming a 220 mm seat track length, the seat was located 170 mm rearward of the frontmost position along the seat track, corresponding to the average of real-world front row passenger seat locations [41]. MHBM hip (H-point) location relative to the seat was adjusted according to sex, height, and weight [42]. The seatbelt model was routed along the shortest path using the belt fitting algorithm in the Primer preprocessor (v.17.0, Oasys Ltd., London, UK). For the frontal impact, the delta-velocity was 48 km/h, and the seat belt retractor load limiting force was 4 kN. For the near-side impact, the lateral delta-velocity was 24 km/h and peak intrusion of the B-pillar was 184 mm. In both impacts, the crash conditions were selected such that a male MHBM corresponding to the 50th male ATD height and weight predicted 50% NFR2+ risk.

The NFR2+ risk from each MHBM in each crash was calculated using a probabilistic age- and strain-based rib fracture risk method [43,44]. Briefly, the maximum value of first principal strain throughout the crash simulation was extracted from each rib cortical bone mesh, resulting in 24 strain values (one from each rib). Next, a fracture risk was calculated for each rib using a strain-based and age-adjusted risk function [44]. Finally, NFR2+ throughout the ribcage was calculated using a generalized binomial model and the 24 fracture risks from each rib [43]. Using the probabilistic method, rib fracture risk was calculated in a postprocessing step, after the crash simulations had concluded. Thus, fracturing of the ribs was not explicitly modeled (by, e.g., element deletion) in the MHBMs. Age was fixed at 45 years also in the risk calculations. All simulations were performed using LS-Dyna (16 cores, v.9.3.1 mpp, Livermore Software Technology, Livermore, CA).

The animation output of each simulation was inspected for signs of errors, such as contact intersections or numerical instabilities (shooting nodes). All NFR2+ results were carefully inspected, and if any NFR2+ result deviated from the result obtained from MHBMs of similar height and weight and sex, all rib strain time histories used in the NFR2+ calculation were carefully examined to rule out numerical issues.

Population Metamodeling. The metamodeling of NFR2+ outputs as a function of sampled input parameters is a regression problem. In a regression problem, n pairs of inputs and output (x_i, y_i) are known (training data), where y_i is the known output from the corresponding input vector (bold notation) $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{ij}]$, representing the i th choice of j input parameters. Here, NFR2+ was the output and sex, height, and weight were the input parameters. A solution to the regression problem is an estimated regression function, \hat{f} , based on the training data, capable of predicting new output values $\hat{y} = \hat{f}(\mathbf{x})$ given new input data \mathbf{x} . The root-mean-square error (RMSE), Eq. (1), was used to measure overall metamodel fit to training data, test data, and across the population sample.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

As a complementary measure of fit, R^2 was used, see Eq. (2), where \bar{y} is the mean of the output sample. R^2 measures the proportion of

variance in the output that is explained by a model. $R^2 = 1$ indicates that all output variance is explained by the model. A model predicting \bar{y} in all points will have $R^2 = 0$, while models with worse predictions can obtain $R^2 < 0$.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

Comparison of Regression Models. Both LASSO regularized linear regression with second-order polynomial terms (LASSO) [26] and Gaussian process regression with separable Gaussian kernel (GPR) (see, e.g., Ref. [28]) were used to create metamodels for NFR2+ risk in both impacts. All metamodeling was performed using R v.4.2.2 [45], including packages “glmnet” v.1.0 [46] and “laGP” v.1.5-8 [47]. Test and training set splits were done using the “caret” package v.6.0-93 [48]. All inputs were normalized to unit range before training. For hyperparameter tuning, 100 LASSO regularization factors ranging from 0.0001 to 1 were evaluated, and the one resulting in the smallest five-fold CV RMSE for the training data were selected. The range and number of LASSO regularization factor levels were determined after initial evaluations. GPR hyperparameters included a kernel length scale for each input parameter and a nugget (corresponding to noise not explained by inputs) and were computed through simultaneous likelihood optimization based on the training data, thus avoiding the need for user inputs to the fitting [28,47].

Train and test RMSE were calculated for training data subsets of increasing sizes to study the evolution of LASSO and GPR metamodel fits to NFR2+ outcomes as functions of training set size. All available data were split into a test and a training set. The test set size was 25% of all data and was only used to calculate test error. Each metamodel was fitted 50 times for randomly sampled (without replacement) subsets of the training data. This was repeated for increasing training subset sizes, until the full (75%) training set was used in the training. Average and standard deviation train and test RMSE were plotted against training set size. The results were used to determine if sufficient training data to represent population outcomes of NFR2+ had been created and to judge which of LASSO and GPR was more accurate for metamodeling of population NFR2+.

Frontal and near-side impact NFR2+ metamodels were created for males, females, and for males and females combined.

Selecting Individuals Representative of Population NFR2+ Risk. To identify sample designs of individuals representative of the population NFR2+, an optimization was performed. In this optimization, those individuals that resulted in the overall best metamodel prediction for the population (the total set of all MHBM outputs in each crash) were selected. The optimization was repeated for an increasing number of individuals, starting from three females and three males. To avoid suboptimization for the specific NFR2+ outcomes in a single accident scenario, the optimization target was to minimize an averaged normalized RMSE (ANRMSE), Eqs. (3) and (4), for metamodel NFR2+ predictions in both front and side impact simultaneously

$$\text{NRMSE}_i = \frac{\text{RMSE}_{i,\text{All}}}{y_i^{\text{max}} - y_i^{\text{min}}} \quad (3)$$

$$\text{ANRMSE} = 0.5 * \text{NRMSE}_{\text{Front}} + 0.5 * \text{NRMSE}_{\text{Side}} \quad (4)$$

where $\text{RMSE}_{i,\text{All}}$ was the metamodel RMSE calculated for all available points in the population, and y_i^{max} and y_i^{min} were the largest and smallest outputs among the currently evaluated candidate training points. Normalization was performed as the range of obtained NFR2+ outputs could be different in the two impacts.

For every subpopulation size, a corresponding number of individuals were selected, one metamodel for each impact was constructed using the training data represented by those individuals, and ANRMSE was calculated. The optimization objective was then to select the individuals that minimize ANRMSE for the population. Selecting the best k individuals from a total sample size of n is known as an “ n -choose- k ” problem which tends to grow to large numbers of possible candidate solutions. E.g., there are 19,900 ways to choose two out of 200, 1.3×10^6 ways to choose three, 65×10^6 and 2.5×10^9 ways to choose four and five, respectively.

Therefore, minimization of ANRMSE was performed using genetic algorithms (GAs). R package “GA” v.3.2.3 [49], with population, mutation, and crossover functions tailored for “genes” containing unique integer indices. Here, each index in a gene corresponded to a specific sample of sex, height, and weight in the MHBM training data. For every optimization, a sequence of three GAs was used. This was done to obtain random restarts and to tradeoff exploration and exploitation by reducing mutation probability in each following GA. Each GA was initially seeded with 1000 random candidate genes. The termination criteria for each GA were 1200 iterations or 300 consecutive iterations without improving ANRMSE. The single best candidate solution from the previously terminated GA was kept in the initialization of the following GA. The GA implementation was verified using brute-force searching to identify the best possible solutions for choosing two and three out of 200. For these subpopulation sizes, the GA implementation successfully identified the best solutions within the first few iterations of the first GA in the sequence.

After the optimization, metamodels for frontal and side impact NFR2+ were constructed using the individuals identified for each subpopulation size, and RMSE and R^2 for those metamodels were calculated using all available samples, corresponding to RMSE and R^2 for the population. As reference for optimization performance, the corresponding measures were calculated from metamodels using

the same number of individuals but selected by the sequential space filling strategy instead.

Results

NFR2+ risk was calculated for 200 male and 200 female MHBMs in both front and near-side impact (in total 800 simulations). All simulations reached normal termination.

In the frontal impact NFR2+ ranged from 32% to 100% for the males and from 1% to 100% for the females. In the side impact NFR2+ ranged from 1% to 70% for males and from 0% to 54% for females, Fig. 3.

To better visualize differences between male and female MHBM NFR2+ results, the same NFR2+ results are additionally shown in perspective plots, Fig. 4.

For females in frontal impact, NFR2+ tended to initially increase with increasing weight, then decrease to lower values for the heavier females. This weight-effect on NFR2+ was also influenced by height. For males in frontal impact, NFR2+ also increased with weight, in a height-modulated manner, but remained saturated at 100% for the heavier males, Fig. 3.

For the females with high weight and low resulting NFR2+, the seatbelt slipped up toward the neck resulting in less seatbelt loading to the ribcage, Fig. 5. The seatbelt did not slip up toward the neck for male MHBMs of higher weights. For males and females of similar height and weight, frontal impact NFR2+ tended to be higher for females, Figs. 3 and 5.

In the side impact, male NFR2+ changed mainly with weight, initially increasing, and then decreasing with increasing weight (Figs. 3 and 6). For most female height and weight combinations, NFR2+ was low. In a limited region, approximately bounded by 160–175 cm in height and 70–100 kg in weight, NFR2+ was influenced by height and weight. For females above 160 cm, increasing weight resulted in an initially increasing NFR2+, which decreased again for weights above approximately 90 kg, Fig. 3. For males and females of similar height and similar, lower, weights, the

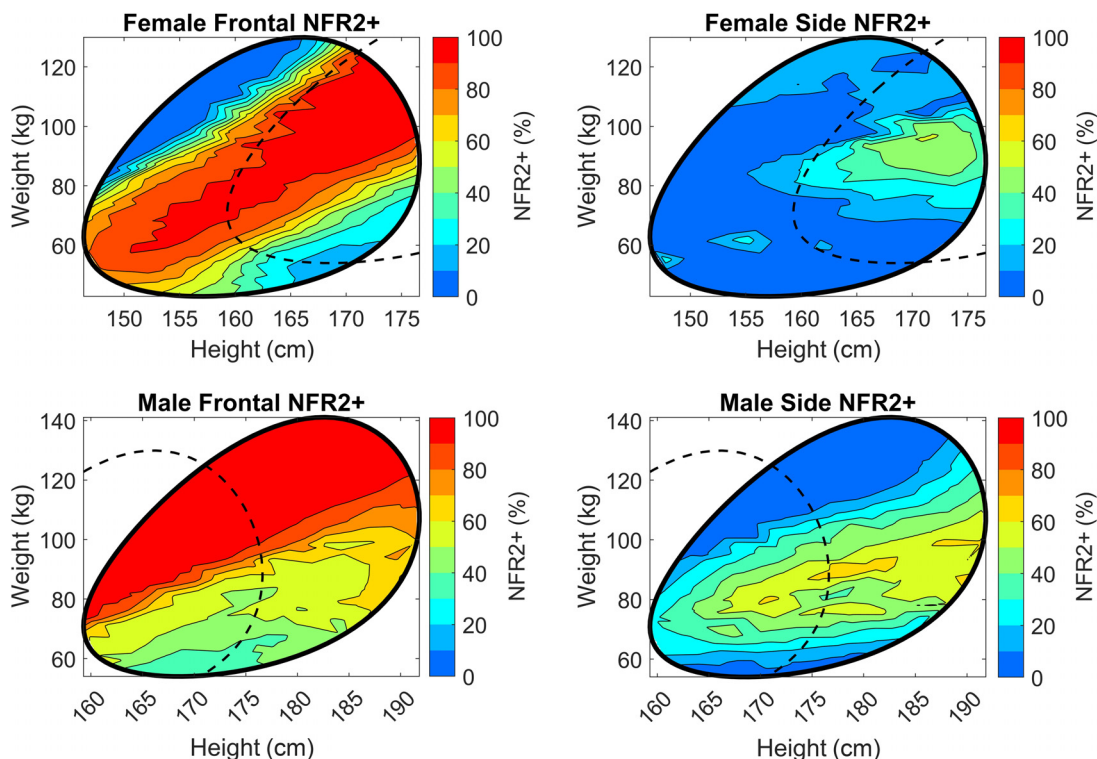


Fig. 3 Contour plots of NFR2+ versus height and weight. The color corresponds to NFR2+ risk in percentage. Dashed curves indicate male and female overlapping height and weight regions. Top row: females. Bottom row: males. Left column: frontal impact. Right column: near-side impact. (Color version online.)

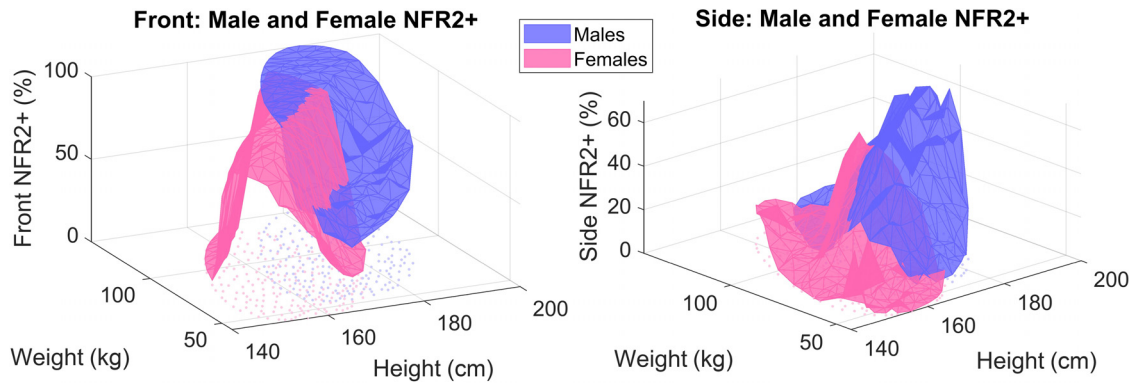


Fig. 4 Perspective plots of NFR2+ versus height and weight. Color corresponds to sex with pink for females and blue for males. Left: frontal impact. Right: near-side impact. (Color version online.)

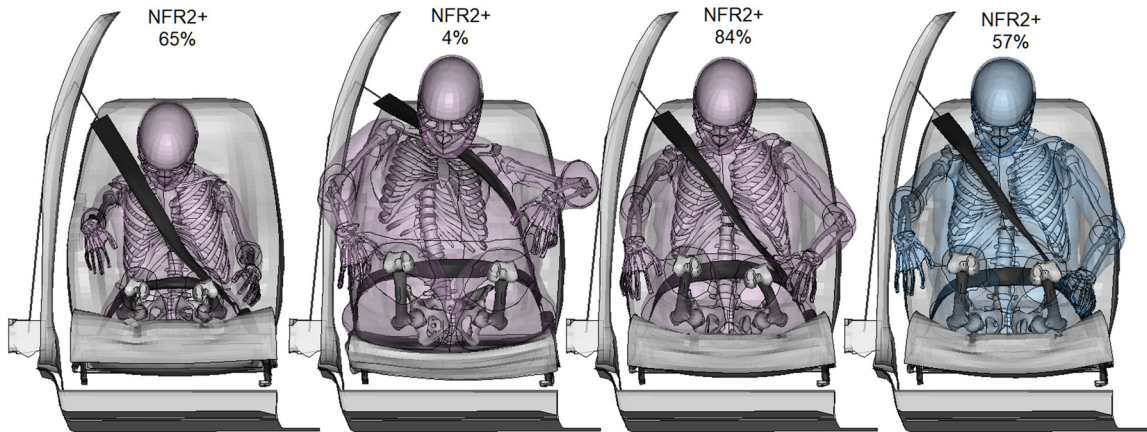


Fig. 5 MHBMs in frontal impact at 92 ms (approx. time of maximum rib strains), parts hidden for visibility. Left to right: female 154 cm and 46 kg, female 161 cm and 116 kg, female 170 cm and 82 kg, male 171 cm and 82 kg.

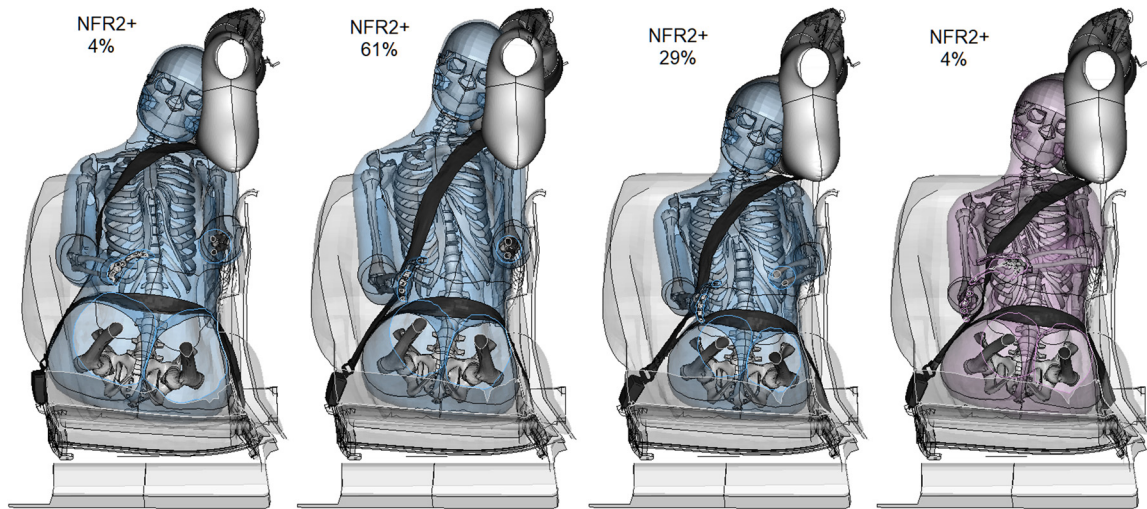


Fig. 6 MHBMs in near-side impact at 49 ms (approx. time for maximum rib strains), parts hidden for visibility. Left to right: male 176 cm and 127 kg, male 190 cm and 101 kg, male 159 cm and 73 kg, female 159 cm and 74 kg.

male NFR2+ was higher. For higher weights and similar heights, female risk was higher, Fig. 4.

Comparison of Regression Models. The reduction in test set error (error between metamodel prediction of NFR2+ and corresponding test set MHBM NFR2+ outputs) for both LASSO and GPR with increasing training set sizes was evaluated. Test error (RMSE) decreased at a high rate with increasing training set sizes up until about

50 (for both males and females) or 25 (single sex) random training points, Fig. 7. Beyond 25 training points for each sex, the relative gain in test error reduction with increased training set sizes was diminished. The exception was GPR for female frontal impact NFR2+, which showed a steady reduction in test errors with increasing training set sizes all the way up to maximum set size, Fig. 7.

Overall, separate models for each sex resulted in lower test set RMSE and higher R^2 for different training set sizes, Fig. 7 and

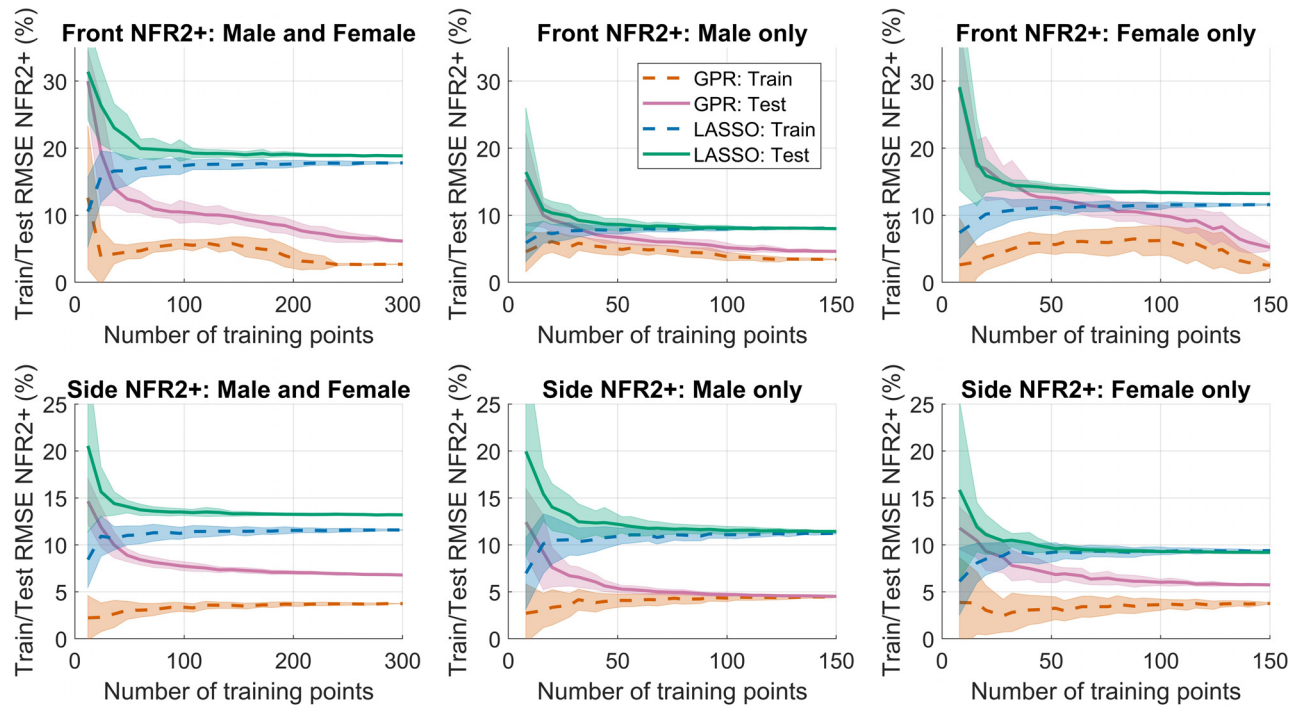


Fig. 7 Average training (dashed curves) and test (solid curves) RMSE for GPR and LASSO predictions of MHBM NFR2+ outputs in frontal (top row) and side (bottom row) impact. Shaded regions represent ± 1 standard deviation. Left to right: models using male and female data, only male data, and only female data.

Table 1. GPR showed smaller test set RMSE and higher R^2 than LASSO for most training set sizes.

Overall, GPR metamodel test set RMSE was judged to be converged for 150 MHBMs of each sex, indicating that 200 males and 200 females were a sufficient sample size of MHBMs for representing population NFR2+. The exception was females in frontal impacts, but here GPR test set R^2 was anyway as high as 0.99 for 150 training points, indicating little room for improvement by adding more training samples.

Selecting Individuals to Represent Population NFR2+ Risk.

Only GPR was used when identifying representative individuals through minimizing ANRMSE, due to the overall lower test errors for GPR compared with LASSO. The optimization was performed for subpopulation sizes of 3–28 training points for each sex. It was stopped due to diminished RMSE and R^2 improvements, Fig. 8. Height and weight of selected males and females for subpopulation sizes up to ten are provided in (See [Supplemental Materials](#) on the ASME Digital Collection, Appendix C).

For both impacts and sexes, GPR models based on optimized training points obtained smaller errors and higher R^2 than corresponding GPR models trained using the space filling sequence of samples, Fig. 8. For models built using the training points selected by optimization, the most substantial reductions in RMSE and gains in R^2 for predicting the population outcome were seen up to about

5–7 points. For more than ten points, gain in RMSE was lower (Fig. 8).

In Fig. 9, the samples selected by the ANRMSE optimization for male and female subpopulation sizes of seven are shown, together with the first seven sample coordinates from the sequential space filling strategy. For this sample size, the optimization selected males spread out across the lower half of the weight range. For females, the selected individuals appear more evenly spread over the region.

The NFR2+ predictions from GPR metamodels based on outputs from seven individuals of each sex that were selected in the optimization, can be seen in Fig. 10. The population RMSE and R^2 for these metamodel predictions are shown for $N = 7$ in Fig. 8. While the metamodels constructed using seven selected training points deviate locally from the population outputs, they predict the overall trends in the crash simulation outputs.

Discussion

The aim of this study was to investigate how many, and which, MHBM anthropometries that are needed to represent a population injury outcome, here NFR2+ in front and side crashes. The considered population included 90% of U.S. males and females in terms of height and weight and NFR2+ across this population was represented by metamodels trained on MHBM front and side crash simulation results.

Main Findings. The MHBM NFR2+ rib fracture risk, which was 50% for a 50th percentile male ATD-sized MHBM, varied nonlinearly with respect to height and weight within each sex and ranged between 1% and 100% in frontal impact and 0–70% in side impact when male and female population height and weight variability were represented in MHBM crash simulations, Fig. 3. Targeting a 50% NFR2+ risk for a 50th percentile male ensured that NFR2+ was sensitive to changes in sex, height, and weight as this result was in the middle of the NFR2+ range. In both higher and lower severity crashes, NFR2+ could be close to 100% or 0%, respectively, for many more body sizes. Therefore, the obtained

Table 1 RMSE and R^2 for predicting MBHM NFR2+ results in the 25% test set, for models using the full (75%) training set

Inputs	Frontal impact				Side impact			
	LASSO		GPR		LASSO		GPR	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
Male and female	18.9	0.62	6.1	0.98	13.2	0.62	6.8	0.94
Male	8.0	0.88	4.6	0.97	11.4	0.70	4.5	0.95
Female	13.2	0.87	5.1	0.99	9.2	0.47	5.7	0.89

NFR2+ results represent somewhat of a worst-case scenario, in terms of sensitivity to sex, height, and weight variations.

The effect of height and weight on NFR2+ results were different for male and female MHBMs even for similar heights and weights, Figs. 3 and 4. Thus, there was no additional benefits to metamodel predictive accuracy (test error RMSE), for either LASSO or GPR, by combining the sexes in a single metamodel, Fig. 7. This implies that female NFR2+ is not representative of male NFR2+, and evaluation for both sexes is necessary for predicting the population NFR2+.

In frontal impact, increasing weight was associated with increasing MHBM NFR2+, reaching saturation at 100% depending on height. Increasing body mass increases the kinetic energy that needs to be absorbed by the safety systems and vehicle interior parts to stop the relative motion of the occupant within the limited space of the vehicle interior, resulting in greater loads to the torso. For female MHBMs, NFR2+ initially increased, and then decreased with increasing weight, due to the seatbelt slipping toward the armpit and neck for females of higher weights. This applied loading to the neck, and even though the NFR2+ risk decreased it does not indicate improved safety for the occupant, but rather a failure of the shoulder belt to restrain the thorax. This phenomenon was not observed among the heavier males, indicating that this was due to female specific body shape features. As an example, the presence of breasts for the females influenced the initial position of the seatbelt.

The MHBM NFR2+ predictions were less sensitive to height and weight variations in the near-side impact than in the frontal impact. While NFR2+ initially increased with weight for males of all heights, and females above 160 cm, after adding additional weight it decreased again. Increasing weight increases the inertia of the body, resulting in larger side impact forces. However, it also increases the soft tissue thickness, which appears to provide protection for the ribs, ultimately resulting in a lower rib fracture risk for the MHBMs of higher weights. If this is a MHBM modeling artifact, or an effect present for occupants in real-world crashes is not known. Increasing BMI (kg/m^2), which corresponds to increasing weight for a fixed height, have both been identified to increase thoracic injury risk [4], and to have a nonsignificant effect [50] in logistic regression models

of real-world near-side crash injury risk. It has been demonstrated that BMI can have a nonlinear effect on real-world rib fracture outcomes [51], which means that statistical models assuming linear BMI effects for rib fracture and thoracic injury risk can underestimate the significance of BMI as a predictor, which can contribute to conflicting findings in accident data analysis. The results from the current MHBM simulation study showed nonlinear effects on NFR2+ by increasing weight for a fixed height, which can support nonlinear modeling of height, weight, and BMI effects on risk in epidemiological studies of vehicle crash injuries.

For the majority of training set sizes, GPR models obtained lower train and test RMSE, indicating a better fit to the MHBM NFR2+ outputs, compared to the LASSO models. Separate GPR-based metamodels for each sex tended to converge to a test NFR2+ RMSE of about 5%-units when using 150 MHBM training examples of each sex, with little improvement in test error for training set sizes beyond 100. The exception was the highly nonlinear female MHBM frontal impact NFR2+ response, for which the GPR metamodel test RMSE showed continuous improvement with increasing training set sizes and reached a test set $R^2 = 0.99$ for 150 training samples. Previous studies metamodeling HBM rib fracture risk reported 8% average test error [22] (three or more fractures), and test set $R^2 = 0.78$ [32] (one or more fractures). The obtained differences between training and test set RMSE for both LASSO and GPR confirmed that metamodel training set errors can be overly optimistic for predictive accuracy, especially for smaller training set sizes, Fig. 7. This highlights the importance of estimating metamodel test error through a test set or k -fold CV to estimate the predictive capabilities of the metamodel.

Combining the male and female MHBM data and including sex as an input parameter for the metamodels did not improve the capability of metamodels to predict NFR2+, even though this effectively doubled the training set size. This is in line with findings in Ref. [32], in which metamodels did not improve when combining training sets otherwise separated by factor inputs. In this study, possible reasons are that the male and female height and weight regions only partially overlap, Fig. 1, and that the MHBM NFR2+

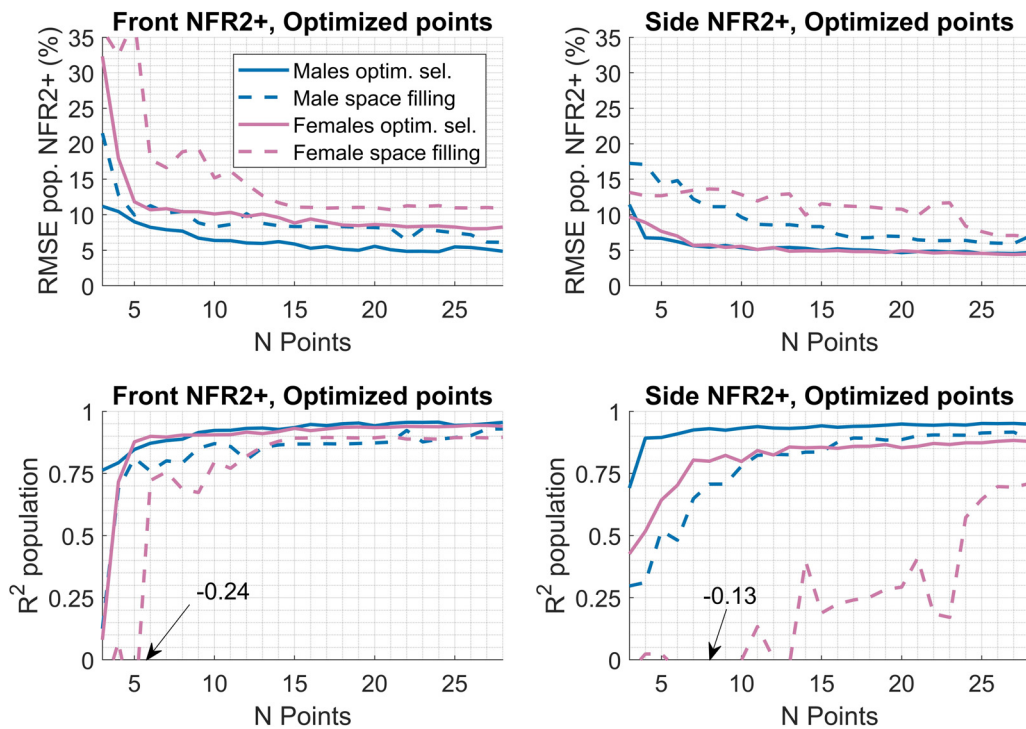


Fig. 8 GPR metamodels fitted using optimized training points and GPR metamodels fitted using the sequence of points from the sequential space filling design. Top row: RMSE for predicting MHBM NFR2+ outputs. Bottom row: R^2 . Left column: frontal impact NFR2+. Right column: near-side impact NFR2+.

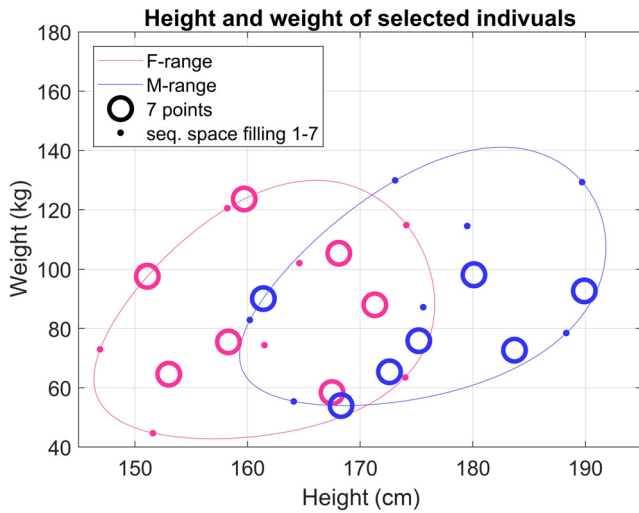


Fig. 9 Male and female individuals identified as minimizing ANRMSE for a population size of seven (circles), and the first seven points from the sequential space filling design (dots)

result trends were different between males and females, also in the overlapping region, Figs. 3 and 4. This ultimately reduced the benefit of sex as an input in a combined model. That is, including female MHBMs NFR2+ trends due to height and weight did not improve the model fit for male MHBMs NFR2+ trends and vice versa.

For randomly selected training samples, the most substantial reduction in test errors were seen for up to about 25 samples of each sex, Fig. 7. For predicting the entire population output, the GPR metamodels constructed using the sequential design of samples all obtained $R^2 > 0.8$ for 13 samples, except for the GPR model of female NFR2+ in the side impact, that reached an R^2 of 0.70 for 26 samples, Fig. 8. The reason for the low R^2 values in this case was that female MHBMs side impact NFR2+ was relatively low and unaffected by height and weight in many cases, except for in a small region where NFR2+ increased, Fig. 3. Therefore, metamodels that did not have sufficient samples in this region consistently predicted too low NFR2+ in this region.

Taking these results into account, there is an indication that future metamodeling efforts of population NFR2+, and other nonlinear crash outcomes, in other vehicles or crash conditions, can expect reasonably accurate GPR metamodel predictions of population outcomes from a training set consisting of at least 25 males and 25 females, evenly spread across their respective height and weight domains. This can be taken as a lower bound, general, answer to the question of how many MHBMs are needed to represent a population outcome in a given crash scenario. This recommendation is in agreement with a general metamodeling rule-of-thumb of sampling ten times the number of input dimensions [52–54].

Further reducing the number of MHBMs requires leveraging known information, which was done when selecting which male and female individuals to be used to construct population representative GPR metamodels through optimizing ANRMSE, Eq. (4), for the entire population across both impacts. Here, it was found that constructing metamodels using as few as five to seven MHBMs of each sex could produce metamodels capable of representing the overall trends in MHBMs population NFR2+ in both crashes, due to height and weight variations, Figs. 8 and 10. It is important to point out that the selected individuals were selected to be representative of the specific male or female NFR2+ outcomes, as only these two crash outcomes were included in the optimization criteria (ANRMSE, Eq. (4)).

That the resulting selection of individuals is related to these specific results can be seen by comparing the location of the selected individuals in Fig. 9, to the NFR2+ results in Fig. 3 as well as to the resulting metamodel predictions for front and side NFR2+ in Fig. 10. For the selected females, the height and weight sizes are located such that the resulting metamodels both can capture the rise and fall of female NFR2+ in the frontal impact as well as predict the increase of NFR2+ in the limited region where height and weight influenced female MHBMs NFR2+ in the side impact. For the males, no individuals in the higher end of the weight range were selected for a subpopulation size of seven, Fig. 9. Therefore, metamodels constructed from these selected males need to extrapolate beyond their training data to predict NFR2+ in the higher weight range. For these specific male MHBMs NFR2+ results, the metamodel extrapolated trends coincide with the MHBMs NFR2+ result trends, Fig. 10, which is why these metamodels performed well in terms of the optimization objective. This means that if an additional male

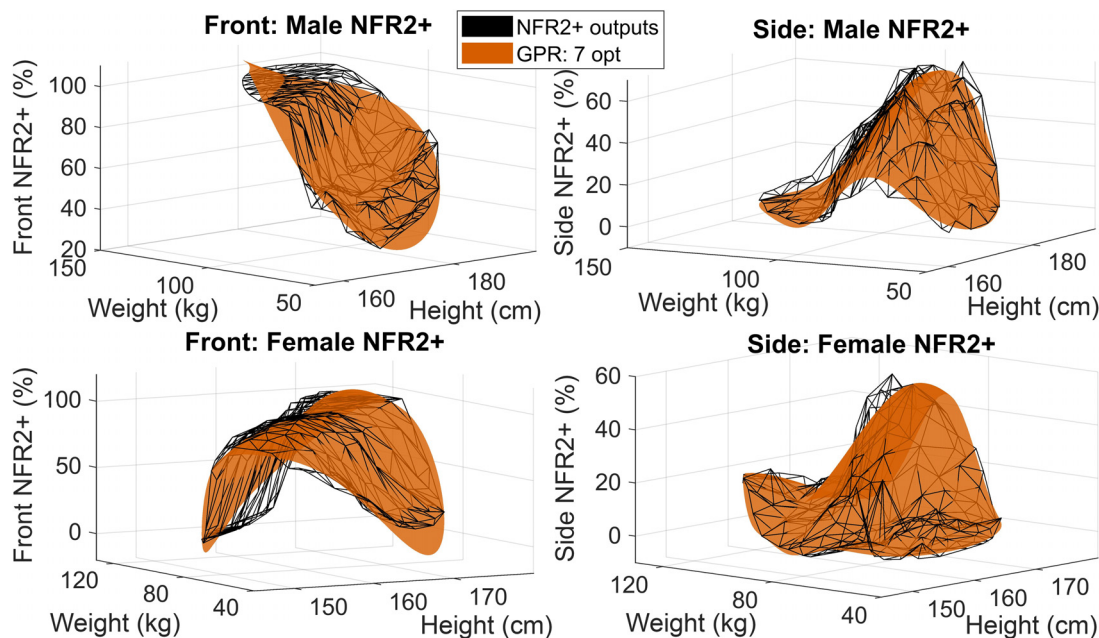


Fig. 10 Perspective plots of GPR models obtained from seven individuals selected by optimization (orange surfaces) together with corresponding crash simulation NFR2+ outputs (wireframe). Top row: males. Bottom row: females. Left column: frontal impact. Right column: side impact. (Color version online.)

crash outcome was added to the optimization objective, where the metamodel extrapolation for higher weights did not correspond to the outcome, it is likely that a different set of individuals would be selected. That is, the selection is not robust when based on only two crash outcomes for each sex. To improve the robustness of the selection, it is necessary to add additional crash outcomes to the optimization objective (further discussed under limitations and future work below).

While the metamodel errors and identified individuals are only valid for the specific crash conditions and NFR2+ results in this study, the optimization results serve as a proof-of-concept for a method capable of selecting a given number of individuals that can be used to create population representative metamodels for more than one crash outcome. For the specific crash outcomes in this study, GPR models constructed using the seven selected individuals of each sex, shown in Fig. 9, Appendix C, all have $R^2 > 0.80$, which is only marginally improved upon by increasing the number of selected individuals, Fig. 8. This indicates a possibility to represent a nonlinear crash outcome for the population of occupants, through metamodels based on around five to ten individuals of each sex in future work.

Methods. The parametric HBM morphing method [15] was used for creating the MHBMs. This method aligns the morphed HBM parts to statistical human shape models of body surface, femur, pelvis, tibia, and the ribcage, all parametrized with respect to sex, age, height, and weight of the human subjects they are based upon. Thus, the resulting MHBMs follow geometrical shape trends existing in the population. The parametric ribcage model [55] enforces sex differences in overall ribcage geometry, including rib cross-sectional dimensions, for which females tend to have more slender ribs [56] which can be influential for rib fracture risks [57,58]. Further, the parametric body shape model represents overall sex differences in body surface geometry [59], such as more pronounced breasts and increased soft tissue volumes around the hip and thighs for females. Thus, both the ribcage model and the external body shape models have implications for how the MHBMs interact with the seatbelt, airbags, and vehicle interior parts, which contributes to the observed sex differences in MHBM NFR2+, even for male and female MHBMs of similar height and weight. Therefore, it is important to include HBMs morphed to females in population evaluations.

Least absolute shrinkage and selection operator with second-order polynomial terms was used as it has been identified as performing well in previous metamodeling with MHBMs [20,21]. Reference [20] did not use LASSO regularization, but added regression terms in a stepwise manner, which can be seen as a manual method of regularization. In this study, GPR obtained lower test RMSE for predicting MHBM NFR2+ for most sample sizes, and the relatively poor performance of LASSO can be attributed to that all LASSO models were prespecified as second-order polynomials, limiting the space of resulting metamodel surface shapes to second-order surfaces. Increasing the polynomial degree could potentially improve the LASSO test RMSEs for NFR2+.

A benefit of GPR, beyond its capability of generalizing well for few training points in this study, is that it is a nonparametric method (i.e., no functional form needs to be assumed), making it suitable for metamodeling of MHBM outputs, even in cases where no prior knowledge about the shape of the resulting output surface exists. An additional feature of GPR, however not utilized in this study, is that it also generates a confidence interval about its predictions, which can guide an analyst in selecting new points for improving the response surface fit. This predictive uncertainty has also been utilized in the design of sequential adaptive sampling strategies that can be used to select training points with a high likelihood to improve the overall fit of the metamodel, based on the information gained from previous training data [28,60,61]. Based on the results in this work, GPR is a suitable method to generate metamodels of population injury risk from MHBM simulations results.

The optimization performed to identify representative individuals was only possible to conduct because the population outcomes were already known in 200 points for each sex and crash scenario. As the optimization objective (ANRMSE, Eq. (4)) depends on the results already obtained for each sex in both impact scenarios, the solutions identified by this optimization are products of these specific MHBM NFR2+ results. To decrease this bias, the average predictive accuracy was used as optimization criteria, which means that the identified subpopulations will not be optimal for any one of the crash outcomes. Nevertheless, for subpopulation sizes of 3–28, the identified individuals performed better, in terms of metamodel accuracy, than a sequential space-filling design, indicating that each selected subpopulation is more representative for the two outcomes. To identify individuals that are representative across more crash scenarios and to make the selection of individuals more robust to particularities in a single crash outcome, the optimization objective, Eq. (4), can easily be expanded to include outputs from more crash scenarios, provided that the population outcome is known.

Limitations and Future Work. There are several limitations with this study. First, only two different impact scenarios were used to obtain the NFR2+ results. The occupant compartments used for each crash scenario represented average vehicle geometries. It is therefore possible that the trends in MHBM NFR2+ with changes in sex, height, and weight seen in each separate crash scenario can be similar in other vehicles. However, to confirm if the current result trends are general for all vehicles, it is necessary to redo the investigation using several different vehicle interiors. As the NFR2+ result trends were different in front and near-side impact, it is also necessary to investigate the population outcome in several different crash configurations and severities, before a definitive set of individuals, generally representative of a population NFR2+ can be confirmed. It is possible that a representative set of individuals can be different in different crash scenarios, especially if small subpopulations are sought. In future work, to decrease the sensitivity of which individuals are selected based on particularities of single crash outcomes (i.e., to increase the robustness of the selected individuals), the optimization to identify representative MHBMs needs to be performed for an objective including population outcomes in more vehicles and crash scenarios. To economically obtain population outcomes in new crash scenarios, the individuals identified as representative can be used as starting points in a study design, and the application of a GPR with a sequential adaptive sample selection method has the potential to enable an economical search of the input space that identifies a population representative metamodel. Alternatively, a space-filling design of 25 MHBMs of each sex can be used.

Further, the influence of age variations in the population on NFR2+ was not considered. Increasing age has been identified to significantly and substantially increase occupant rib fracture and thoracic injury risk in crashes [4,5]. The age effect can potentially be even more influential for rib fracture risk than height, weight, and sex [4], and needs consideration in a population perspective. Potentially contributing factors for the age trend are, for example, that rib cortical bone material properties such as Young's modulus and ultimate strain degrades and that rib cortical bone thickness is reduced with increasing age [38,39,44]. However, the focus of this study was the influence of height, weight, and sex. It is likely that the influence of the height and weight parameters is similar for occupants of different ages, although the overall magnitude of rib fracture risk will be different when different ages are modeled. Including age effects in MHBMs can contribute to also identifying ages of individuals representative for population NFR2+ in future work.

Additionally, the population of real-world occupants will differ in a multitude of dimensions beyond those described by statistical shape trends related to height, weight, and sex. A certain height can be reached by different proportions of leg and torso lengths, and occupants of similar weight can have different proportions of fat and muscle mass, influencing overall body shape and resulting position

of the seatbelt, to name a few. Beyond individual body shape variations, each occupant can select various seating postures and placement of the seatbelt across the torso which alters initial crash conditions and resulting outcomes [62]. Further investigation of population variability in more dimensions may reveal important aspects to consider when representing crash outcomes of a population.

Finally, only NFR2+ was considered in this work. However, designing occupant safety often requires balancing injury risks in different body regions. As an example, in frontal impacts the seatbelt load limiting force should be low to limit loading to the chest, but it should also be high enough to limit forward excursion of the chest and head, such that hard contacts with vehicle interior parts are avoided. Therefore, additional body region risk outputs, such as head injury risk, should also be added to the optimization objective in future work, to enable selection of individuals that are representative also for additional injury risk outputs.

Conclusions

Subpopulations representative for the overall population NFR2+ outcomes in both a front and a side impact scenario were identified. Seven selected individuals of each sex enabled constructing GPR metamodels representing population NFR2+ outcomes in both crashes.

The subpopulation selection was based on known outcomes. In other vehicles and crash scenarios, where the outcome is unknown, our results indicate that GPR metamodels based on 25 male and 25 females can represent the population outcome.

Funding Data

- The work performed in this study was funded by FFI-Strategic Vehicle Research and Innovation, Vinnova-Sweden's Innovation Agency, the Swedish Energy Agency, the Swedish vehicle industry; Funder ID: 10.13039/501100001858.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

- [1] Bose, D., Segui-Gomez, M., and Crandall, J. R., 2011, "Vulnerability of Female Drivers Involved in Motor Vehicle Crashes: An Analysis of U.S. Population at Risk," *Am. J. Public Health*, **101**(12), pp. 2368–2373.
- [2] Viano, D. C., Parenteau, C. S., and Edwards, M. L., 2008, "Crash Injury Risks for Obese Occupants Using a Matched-Pair Analysis," *Traffic Inj. Prev.*, **9**(1), pp. 59–64.
- [3] Abrams, M., and Bass, C. R., 2022, "Female vs. Male Relative Risk of Body System Injuries in Fatal and Non-Fatal Crashes," *Proceedings of the IRCOBI Conference*, Porto, Portugal, Sept. 14–16, pp. 33–44.
- [4] Carter, P. M., Flannagan, C. A. C., Reed, M. P., Cunningham, R. M., and Rupp, J. D., 2014, "Comparing the Effects of Age, BMI and Gender on Severe Injury (AIS 3+) in Motor-Vehicle Crashes," *Accid. Anal. Prev.*, **72**, pp. 146–160.
- [5] Forman, J., Poplin, G. S., Shaw, C. G., McMurry, T. L., Schmidt, K., Ash, J., and Sunnevang, C., 2019, "Automobile Injury Trends in the Contemporary Fleet: Belted Occupants in Frontal Collisions," *Traffic Inj. Prev.*, **20**(6), pp. 607–612.
- [6] Brumbelow, M. L., and Jermakian, J. S., 2022, "Injury Risks and Crashworthiness Benefits for Females and Males: Which Differences Are Physiological?," *Traffic Inj. Prev.*, **23**(1), pp. 11–16.
- [7] Schneider, L. W., Robbins, D. H., Pflüg, M. A., and Snyder, R. G., 1983, "Development of Anthropometrically Based Design Specifications for an Advanced Adult Anthropomorphic Dummy Family," University of Michigan Transportation Research Institute, Ann Arbor, MI, Report No. UMTRI-83-53-1.
- [8] Shigeta, K., Kitagawa, Y., and Yasuki, T., 2009, "Development of Next Generation Human FE-Model Capable of Organ Injury Prediction," Proceedings of the 21st Enhanced Safety Vehicle (ESV) Conference, Stuttgart, Germany, June 15–18, Paper No. 09-0111.
- [9] Gayzik, F. S., Moreno, D. P., Vavalle, N. A., Rhyne, A. C., and Stitzel, J. D., 2012, "Development of a Full Human Body Finite Element Model for Blunt Injury Prediction Utilizing a Multi-Modality Medical Imaging Protocol," *Proceedings of the 12th International LS-DYNA Users Conference*, Detroit, MI, June 3–5, pp. 1–14.
- [10] Pipkorn, B., Östh, J., Brynskog, E., Larsson, E., Rydqvist, L., Iraeus, J., Perez-Rapela, D., and Jakobsson, L., 2021, "Validation of the SAFER Human Body Model Kinematics in Far-Side Impacts," *Proceedings of the IRCOBI Conference*, Online Conference, Sept. 8–10, pp. 444–476.
- [11] Iraeus, J., and Pipkorn, B., 2019, "Development and Validation of a Generic Finite Element Ribcage to Be Used for Strain-Based Fracture Prediction," *Proceedings of the IRCOBI Conference*, Florence, Italy, Sept. 11–13, pp. 193–210.
- [12] Pipkorn, B., Iraeus, J., Björklund, M., Bunketorp, O., and Jakobsson, L., 2019, "Multi-Scale Validation of a Rib Fracture Prediction Method for Human Body Models," *Proceedings of the IRCOBI Conference*, Florence, Italy, Sept. 11–13, pp. 175–192.
- [13] Beillas, P., and Berthet, F., 2017, "An Investigation of Human Body Model Morphing for the Assessment of Abdomen Responses to Impact Against a Population of Test Subjects," *Traffic Inj. Prev.*, **18**(Suppl. 1), pp. S142–S147.
- [14] Hwang, E., Hu, J., Chen, C., Klein, K. F., Miller, C. S., Reed, M. P., Rupp, J. D., and Hallman, J. J., 2016, "Development, Evaluation, and Sensitivity Analysis of Parametric Finite Element Whole-Body Human Models in Side Impacts," *SAE Paper No. 2016-22-0014*.
- [15] Hwang, E., Hallman, J., Klein, K., Rupp, J., Reed, M., and Hu, J., 2016, "Rapid Development of Diverse Human Body Models for Crash Simulations Through Mesh Morphing," *SAE Paper No. 2016-01-1491*.
- [16] Zhang, K., Cao, L., Wang, Y., Hwang, E., Reed, M. P., Forman, J., and Hu, J., 2017, "Impact Response Comparison Between Parametric Human Models and Postmortem Human Subjects With a Wide Range of Obesity Levels," *Obesity*, **25**(10), pp. 1786–1794.
- [17] Hwang, E., Hu, J., and Reed, M. P., 2020, "Validating Diverse Human Body Models Against Side Impact Tests With Post-Mortem Human Subjects," *J. Biomech.*, **98**, p. 109444.
- [18] Larsson, K.-J., Pipkorn, B., Iraeus, J., Forman, J., and Hu, J., 2022, "Evaluation of a Diverse Population of Morphed Human Body Models for Prediction of Vehicle Occupant Crash Kinematics," *Comput. Methods Biomech. Biomed. Eng.*, **25**(10), pp. 1125–1155.
- [19] Euro NCAP, 2022, *Euro NCAP Vision 2030: A Safer Future for Mobility*, Euro NCAP, Leuven, Belgium.
- [20] Yang, S., Zhou, Q., Nie, B., Tang, J., Yamamae, Y., and Katsuhara, T., 2023, "A Human Body Model Based Method for Injury Risk Prediction Considering Occupant Stature and Posture," *Int. J. Crashworthiness*, **28**(4), pp. 498–510.
- [21] Joodaki, H., Gepner, B., and Kerrigan, J., 2021, "Leveraging Machine Learning for Predicting Human Body Model Response in Restraint Design Simulations," *Comput. Methods Biomech. Biomed. Eng.*, **24**(6), pp. 597–611.
- [22] Perez-Rapela, D., Forman, J. L., Huddleston, S. H., and Crandall, J. R., 2021, "Methodology for Vehicle Safety Development and Assessment Accounting for Occupant Response Variability to Human and Non-Human Factors," *Comput. Methods Biomech. Biomed. Eng.*, **24**(4), pp. 384–399.
- [23] James, G., Witten, D., Hastie, T., and Tibshirani, R., 2013, *An Introduction to Statistical Learning*, Springer, New York.
- [24] Östh, J., Pipkorn, B., Forsberg, J., and Iraeus, J., 2021, "Numerical Reproducibility of Human Body Model Crash Simulations," *Proceedings of the IRCOBI Conference*, Online Conference, Sept. 8–10, pp. 431–443.
- [25] Druecker, J.-P., 2011, "Optimization of Restraint Systems of a Vehicle Architecture Using Meta Models," Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles (ESV), Washington, DC, June 13–16, pp. 1–9.
- [26] Tibshirani, R., 1996, "Regression Shrinkage and Selection Via the Lasso," *J. R. Stat. Soc.: Ser. B (Methodol.)*, **58**(1), pp. 267–288.
- [27] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., 1989, "Design and Analysis of Computer Experiments," *Stat. Sci.*, **4**(4), pp. 409–435.
- [28] Gramacy, R. B., 2020, *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*, CRC Press, Taylor & Francis Group, Boca Raton, FL.
- [29] Di Achille, P., Harouni, A., Khamzin, S., Solovyova, O., Rice, J. J., and Gurev, V., 2018, "Gaussian Process Regressions for Inverse Problems and Parameter Searches in Models of Ventricular Mechanics," *Front. Physiol.*, **9**, p. 1002.
- [30] Noé, U., Lazarus, A., Gao, H., Davies, V., Macdonald, B., Mangion, K., Berry, C., Luo, X., and Husmeier, D., 2019, "Gaussian Process Emulation to Accelerate Parameter Estimation in a Mechanical Model of the Left Ventricle: A Critical Step Towards Clinical End-User Relevance," *J. R. Soc. Interface*, **16**(156), p. 20190114.
- [31] Nie, B., Xia, Y., Zhou, Q., Huang, J., Deng, B., and Neal, M., 2014, "A Response-Surface-Based Tool for Vehicle Front-End Design for Pedestrian Impact Protection Using Human Body Model," *IJVD*, **66**(4), pp. 347–362.
- [32] Schneider, B., Kofler, D., D'Addetta, G. A., Freienstein, H., Wolkenstein, M., and Klug, C., 2022, "Approach for Machine Learning Based Design of Experiments for Occupant Simulation," *Front. Future Transp.*, **3**, p. 913852.
- [33] Kent, R., Patrie, J., and Benson, N., 2003, "The Hybrid III Dummy as a Discriminator of Injurious and Non-Injurious Restraint Loading," *Annu. Proc./Assoc. Adv. Automot. Med.*, **47**, pp. 51–75.
- [34] Brumbelow, M. L., 2020, "Can Front Crash Rating Programs Using Hybrid III Predict Real-World Thoracic Injuries?," *Proceedings of the IRCOBI Conference*, Munich, Germany, pp. 679–692.
- [35] Brumbelow, M. L., Jermakian, J. S., and Arbelaez, R. A., 2022, "Predicting Real-World Thoracic Injury Using THOR and Hybrid III Crash Tests," *IRCOBI Conference Proceedings*, Porto, Portugal, Sept. 14–16, pp. 264–287.

- [36] Brolin, E., Högberg, D., and Hanson, L., 2020, "Skewed Boundary Confidence Ellipses for Anthropometric Data," *Proceedings of the 6th International Digital Human Modeling Symposium*, Skövde, Sweden, Aug. 31–Sept. 2, pp. 18–26.
- [37] Specht, E., 2021, "The Best Known Packings of Equal Circles in a Circle (Complete up to $N = 2600$)," accessed Jan. 24, 2023, www.packomania.com
- [38] Katzenberger, M. J., Albert, D. L., Agnew, A. M., and Kemper, A. R., 2020, "Effects of Sex, Age, and Two Loading Rates on the Tensile Material Properties of Human Rib Cortical Bone," *J. Mech. Behav. Biomed. Mater.*, **102**, p. 103410.
- [39] Holcombe, S. A., and Derstine, B. A., 2022, "Rib Cortical Bone Thickness Variation in Adults by Age and Sex," *J. Anat.*, **241**(6), pp. 1344–1356.
- [40] Iraeus, J., and Lindquist, M., 2016, "Development and Validation of a Generic Finite Element Vehicle Buck Model for the Analysis of Driver Rib Fractures in Real Life Nearside Oblique Frontal Crashes," *Accid. Anal. Prev.*, **95**, pp. 42–56.
- [41] Reed, M. P., Ebert, S. M., Jones, M. L. H., and Hallman, J. J., 2020, "Prevalence of Non-Nominal Seat Positions and Postures Among Front-Seat Passengers," *Traffic Inj. Prev.*, **21**(Suppl. 1), pp. S7–S12.
- [42] Park, J., Ebert, S. M., Reed, M. P., and Hallman, J. J., 2016, "A Statistical Model Including Age to Predict Passenger Postures in the Rear Seats of Automobiles," *Ergonomics*, **59**(6), pp. 796–805.
- [43] Forman, J., Kent, R. W., Mroz, K., Pipkorn, B., Bostrom, O., and Segui-Gomez, M., 2012, "Predicting Rib Fracture Risk With Whole-Body Finite Element Models: Development and Preliminary Evaluation of a Probabilistic Analytical Framework," *Ann. Adv. Automot. Med.*, **56**, pp. 109–124.
- [44] Larsson, K.-J., Blennow, A., Iraeus, J., Pipkorn, B., and Lubbe, N., 2021, "Rib Cortical Bone Fracture Risk as a Function of Age and Rib Strain: Updated Injury Prediction Using Finite Element Human Body Models," *Front. Bioeng. Biotechnol.*, **9**, p. 677768.
- [45] R Core Team, 2022, "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing, Vienna, Austria.
- [46] Friedman, J., Hastie, T., and Tibshirani, R., 2010, "Regularization Paths for Generalized Linear Models Via Coordinate Descent," *J. Stat. Software*, **33**(1), pp. 1–22.
- [47] Gramacy, R. B., 2016, "LaGP: Large-Scale Spatial Modeling Via Local Approximate Gaussian Processes in R," *J. Stat. Software*, **72**(1), pp. 1–46.
- [48] Kuhn, M., 2008, "Building Predictive Models in R Using the Caret Package," *J. Stat. Software*, **28**(5), pp. 1–26.
- [49] Scrucca, L., 2013, "GA: A Package for Genetic Algorithms in R," *J. Stat. Software*, **53**(4), pp. 1–37.
- [50] Pal, C., Tomosaburo, O., Vimalalithan, K., Jeyabharath, M., Muthukumar, M., Satheesh, N., and Narahari, S., 2014, "Effect of Weight, Height and BMI on Injury Outcome in Side Impact Crashes Without Airbag Deployment," *Accid. Anal. Prev.*, **72**, pp. 193–209.
- [51] Forman, J. L., and McMurry, T. L., 2018, "Nonlinear Models of Injury Risk and Implications in Intervention Targeting for Thoracic Injury Mitigation," *Traffic Inj. Prev.*, **19**(Suppl. 2), pp. S103–S108.
- [52] Chapman, W. L., Welch, W. J., Bowman, K. P., Sacks, J., and Walsh, J. E., 1994, "Arctic Sea Ice Variability: Model Sensitivities and a Multidecadal Simulation," *J. Geophys. Res.*, **99**(C1), pp. 919–935.
- [53] Jones, D. R., Schonlau, M., and Welch, W. J., 1998, "Efficient Global Optimization of Expensive Black-Box Functions," *J. Global Optim.*, **13**(4), pp. 455–492.
- [54] Loepky, J. L., Sacks, J., and Welch, W. J., 2009, "Choosing the Sample Size of a Computer Experiment: A Practical Guide," *Technometrics*, **51**(4), pp. 366–376.
- [55] Wang, Y., Cao, L., Bai, Z., Reed, M. P., Rupp, J. D., Hoff, C. N., and Hu, J., 2016, "A Parametric Ribcage Geometry Model Accounting for Variations Among the Adult Population," *J. Biomech.*, **49**(13), pp. 2791–2798.
- [56] Holcombe, S. A., Kang, Y.-S., Derstine, B. A., Wang, S. C., and Agnew, A. M., 2019, "Regional Maps of Rib Cortical Bone Thickness and Cross-Sectional Geometry," *J. Anat.*, **235**(5), pp. 883–891.
- [57] Agnew, A. M., Murach, M. M., Dominguez, V. M., Sreedhar, A., Misicka, E., Harden, A., Bolte, J. H., IV, Kang, Y., Stammen, J., and Moorhouse, K., 2018, "Sources of Variability in Structural Bending Response of Pediatric and Adult Human Ribs in Dynamic Frontal Impacts," *Stapp Car Crash J.*, **62**, pp. 119–192.
- [58] Larsson, K.-J., Iraeus, J., Holcombe, S., and Pipkorn, B., 2023, "Influences of Human Thorax Variability on Population Rib Fracture Risk Prediction Using Human Body Models," *Front. Bioeng. Biotechnol.*, **11**, p. 1154272.
- [59] Park, B.-K. D., Jones, M. L. H., Ebert, S., and Reed, M. P., 2022, "A Parametric Modeling of Adult Body Shape in a Supported Seated Posture Including Effects of Age," *Ergonomics*, **65**(6), pp. 795–803.
- [60] Svendsen, D. H., Martino, L., and Camps-Valls, G., 2020, "Active Emulation of Computer Codes With Gaussian Processes—Application to Remote Sensing," *Pattern Recognit.*, **100**, p. 107103.
- [61] Fuhg, J. N., Fau, A., and Nackenhorst, U., 2021, "State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging," *Arch. Comput. Methods Eng.*, **28**(4), pp. 2689–2747.
- [62] Leledakis, A., Östh, J., Iraeus, J., Davidsson, J., and Jakobsson, L., 2022, "The Influence of Occupant's Size, Shape and Seat Adjustment in Frontal and Side Impacts," *Proceedings of the IRCOBI Conference*, Porto, Portugal, Sept. 14–16, pp. 549–584.