Accelerating HBM Simulations Through Reduced Order Modelling: Impact of Model Choice and Training Data Size.

Jovana Aleksic, Marius Rees, Philipp Wernicke, Steffen Peldschus

I. INTRODUCTION

Vehicle safety development increasingly relies on simulation to inform design decisions and ensure occupant protection in crash scenarios. Human Body Models (HBMs) capture detailed anatomy and dynamics but are computationally expensive, which limits their use in large parametric studies across anthropometries and restraint configurations. Model order reduction (MOR) addresses this limitation by approximating high-fidelity simulations in a low-dimensional representation. This enables fast predictions while preserving relevant kinematics. Previous work of the authors has demonstrated the feasibility of MOR for HBM applications [1], but the influence of model choice and training data size on prediction accuracy remained unclear. This study investigates linear and nonlinear reduction techniques, including autoencoder (AE) based methods, in combination with different surrogate modelling techniques.

II. METHODS

A simplified HBM setup based on THUMS v5 was used, focusing on pelvis and belt interaction to capture the dominant kinematics relevant for occupant restraint behavior. The model was positioned on a rigid flat seat subjected to a prescribed horizontal acceleration pulse, while the pelvis was restrained by a simplified seatbelt system with a double-sided spring-based pretensioner (Fig. 1). Six input parameters describing loading, anthropometry, and restraint configuration were varied using Sobol sampling (Table 1). Belt anchorage position defines the horizontal distance from where the belt is anchored to the seat. Simulations were carried out in LS-DYNA R9.3 (MPP) with a duration of 100 ms.

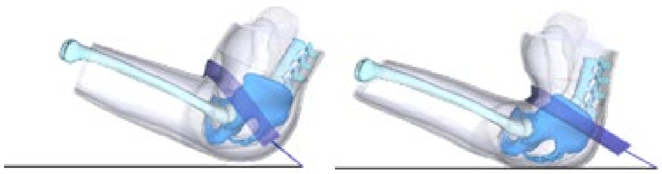

Fig. 1: Simplified THUMS v5 HBM setup at 0ms(left) and 60ms(right)

Table 1: Input parameters and corresponding ranges

Parameter	Range
Acceleration pulse factor	[15 g – 30 g]
Torso mass	[10 kg – 35 kg]
Rotation of HBM upper legs	[-4 ° – 20 °]
Friction between belt and HBM	[0 – 1]
Stiffness of soft tissue parts	[0.5 – 1.5]
Belt anchorage position	[-30 mm – 100 mm]

The full dataset comprised 300 simulations, from which subsets of 50, 100, 200, and 300 samples were used for training to analyze the influence of training data size. The resulting time-dependent nodal displacement fields were split into training and test sets. MOR techniques were applied to the training data to obtain a low-dimensional representation. Linear reduction was performed using Proper Orthogonal Decomposition (POD), while nonlinear representations were learned using AE-based methods [2], variational AE (VAE) [3], and conditional variants (CAE, CVAE) [3,4]. In contrast to POD, AE-based methods learn nonlinear mappings, with variational approaches introducing a probabilistic latent space and conditional variants incorporating input parameters. Surrogate models were trained to map input parameters to the reduced representation using Gaussian process regression (GPR) and radial basis function (RBF) interpolation.

During evaluation, latent predictions were reconstructed to the full size and compared to reference simulations from the test set to assess prediction accuracy. Although the full model includes surrounding soft tissue and upper legs, the analysis was restricted to pelvis and belt nodes (2,700 nodes), as these capture the dominant kinematics relevant for restraint interaction, such as submarining.

III. INITIAL FINDINGS

Prediction errors of the x-displacement of pelvic bone and belt nodes, quantified using the mean absolute error (MAE), defined as the average absolute deviation from the reference displacements, decreased consistently with increasing training dataset size across all methods (Fig. 3). The largest improvement was observed between

J. Aleksic is a MSc student at TUM and at BMW AG, M. Rees (e-mail: marius.rees@bmw.de) and P. Wernicke are safety engineers at BMW AG, S. Peldschus is a professor at Ludwig-Maximilians University Munich.

50 and 100 training simulations, followed by smaller gains with further increasing dataset size. This indicates diminishing returns, suggesting that a moderate number of full-order FE simulations is sufficient to capture the dominant system behavior. Figure 4 illustrates the spatial distribution of prediction errors for a representative simulation case, which indicates generally low prediction errors with localized deviations.

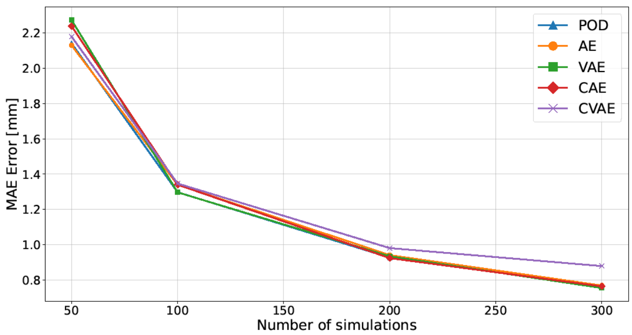

Fig. 3: Mean absolute error (MAE) of predicted x-displacements for pelvic bone and belt nodes as a function of training dataset size.

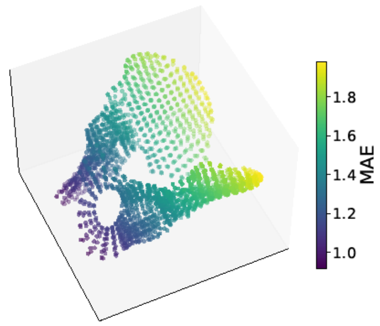

Fig. 4: Spatial distribution of MAE for predicted x-displacements of pelvic bone nodes for a representative test simulation (PCA + RBF, trained with 200 simulations).

Across all dataset sizes, only minor differences were observed between the linear and nonlinear reduction methods. Principal Component Analysis (PCA) achieved accuracy comparable to the autoencoder-based approaches, and all methods showed a similar trend with increasing training data. In addition, RBF interpolation consistently showed competitive or improved performance compared to GPR, while requiring lower computational effort. The kernel PCA (kPCA) approach was not further considered, as it required extensive hyperparameter tuning with limited improvement in prediction accuracy. Overall, the results indicate that both model choice and training data size influence prediction quality, suggesting that combinations of simple reduced order modelling and surrogate modelling techniques can already provide robust approximations for the investigated HBM configuration.

IV. DISCUSSION

The results indicate that accurate predictions can be achieved with moderate training data sizes and simple modelling approaches for the investigated setup. The comparable performance of linear and nonlinear reduction methods suggests that the system behavior can be represented effectively in a low-dimensional linear subspace. This is likely due to the limited nonlinearity of the simplified configuration and restricted parameter space. As a result, more complex nonlinear methods, such as AE and its variants, do not provide a clear advantage in this specific case. This finding cannot be generalized to all HBM applications. The present study is based on a simplified model setup and focuses on a limited set of output quantities, namely displacement fields of pelvis and belt nodes. More complex scenarios, such as full-body HBMs with strongly nonlinear soft tissue behavior, larger environmental configurations involving interactions with airbags or advanced restraint systems, as well as alternative injury-related outputs, may require nonlinear reduction techniques to capture the underlying dynamics more accurately.

Future work will therefore extend the proposed framework to more complex model configurations and broader output definitions relevant for injury assessment. In particular, the transferability of the findings to full-scale HBMs and more strongly nonlinear response regimes will be investigated. Initial investigations will consider the load cases defined in the ENCAP HBM-qualification protocol CP550 [5]. Overall, the results provide guidance for efficient surrogate model development, showing that accurate approximations can be achieved with moderate training data sizes and simple modelling approaches.

V. REFERENCES

[1] Go L et al., Comput Methods, Biomed Eng, 2023.
[2] Michelucci U, arXiv, 2022.
[3] Boussam S et al., IACR Trans CHE, 2025.
[4] Liu et al., EMNLP, 2022.
[5] Euro NCAP Technical Bulleting CP 550.