Towards Interpretable Thoracic Injury Prediction in Frontal Crashes Using Kolmogorov-Arnold Networks to Assess the Impact of Load-Limiting Seatbelt Force and Delta-V

Manuel Valdano, Alejandro Polo-Molina, Corina Klug, Jesús R. Jiménez-Octavio, Bengt Pipkorn, Francisco J. López-Valdés

I. INTRODUCTION

Virtual testing enables the evaluation of occupant safety, in a broader range of real-world scenarios than standardised load cases, by including variations in occupant sizes, vehicle types, and impact conditions. Finite element human body models (HBMs) are computationally expensive, limiting the number of feasible simulations, which can be addressed by using surrogate models, mathematical representations also known as metamodels, that approximate the relationship between input and outcomes. This minimises the costs to obtain it from full simulations, enabling to upscale simulation results. Commonly used techniques include least absolute shrinkage and selection operator (LASSO), neural networks, support vector machines, and gaussian process regression (GPR).

Traditional Machine Learning offers accurate forecasting, with a lack of interpretability, which can make it difficult to understand complex variable relationships. Kolmogorov-Arnold Networks (KANs) [1] offer a more transparent and interpretable framework for modeling outcomes through learnable univariate functions. The metamodel output can be described by combining these functions in an explicit analytical expression, enabling humans to interpret the relationships between the inputs. In this study, a KAN-based metamodel is used to analyse the joint effects of load-limiting seatbelts and changes in velocity (delta-v) on thoracic injury risk, providing insights into these factors' individual and combined contributions. The KAN model's accuracy is benchmarked against the currently most frequently used GPR.

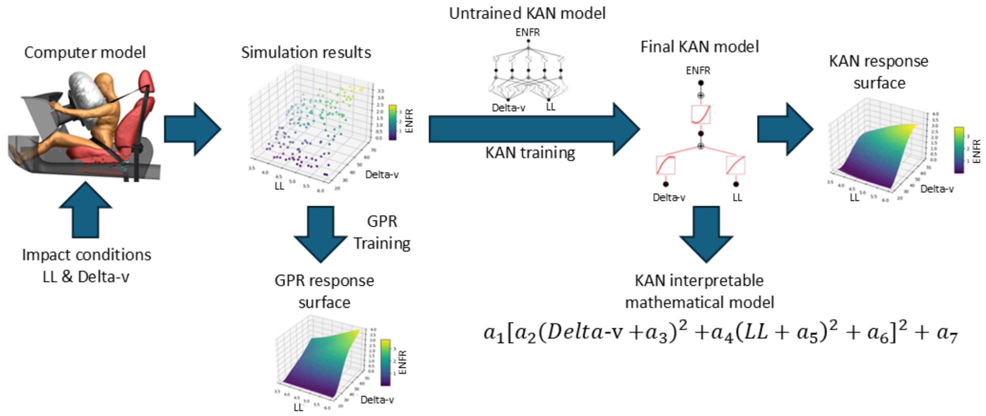

Fig. 1. Flowchart illustrating the method employed throughout the study and the resulting equation.

II. METHODS

Frontal crash simulations, combining a generic vehicle interior (GVI) model and an HBM, were performed to predict thoracic injuries for different delta-v and forces applied with the load-limiting device. A modified version of the GVI model from [2] was used by replacing the seat with an open-source front seat model [3] and the retractor with a simplified generic model [4]. The occupant VIVA+ model was utilised for the HBM positioned as driver, with its two anthropometries: the 50th percentile male (50M) and female (50F). The model included a three-point seatbelt. The retractor included a 2 kN pre-tensioner and a load-limiting device, with forces ranging from 3.5 to 6 kN. The airbag model remained unchanged from [2]. Crashes were simulated with delta-v values from 20 to 70 km/h, and the pulse shape was estimated using [6]. A total of 300 simulations were conducted to predict thoracic injuries by estimating the number of fractured ribs (ENFR), described in [7], for various combinations of load-limiting forces and delta-v values. ENFR was calculated as a continuous variable starting from zero.

Predicted ENFRs were used to develop two KAN metamodels, one for each VIVA+ anthropometry,

M. Valdano (e-mail: mvaldano@comillas.edu) is Assistant Professor, A. Polo-Molina is a PhD student, and Jesús R. Jiménez-Octavio and F. J. López-Valdés are Associate Professors at MOBIOS Lab, IIT, Universidad Pontificia Comillas, Spain. A. C. Klug is Assistant Professor at the Vehicle Safety Institute at Graz University of Technology in Austria. B. Pipkorn is an Adjunct Professor at Chalmers University and Director of Simulation and Active Structures at Autoliv Research in Sweden.

which resulted in a mathematical expression relating ENFR to delta-v and load-limiting forces. The architecture of the KAN metamodel was pre-defined to obtain a model based on linear and quadratic interactions between the variables. A Gaussian Process Regression (GPR) metamodel was also used to benchmark the accuracy of the predictions of the KAN metamodel.

III. INITIAL FINDINGS

Using KAN and GPR regression methods, four metamodels were developed to estimate ENFR for 50M and 50F occupants. The number of ENFR observed in the simulations ranged from 0 to 3.5. The KAN-based metamodel yielded mean absolute errors of 0.131 (50M) and 0.147 (50F) ENFR. In contrast, the GPR metamodel achieved substantially lower prediction errors (0.045 for 50M and 0.070 for 50F), representing a two- to three-fold improvement in accuracy over the KAN metamodels when estimating the HBM response surfaces. KAN regression models, though less accurate than GPR ones, provide interpretable mathematical expressions, identifying the linear and quadratic interaction terms between load-limiting forces and delta-v values, as shown in "KAN interpretable mathematical model" in Figure 1. The analytical expression from the metamodel indicated that an increase in load-limiting forces was associated with a larger ENFR. However, this relationship was influenced by delta-v, as the final univariate function was quadratic. Specifically, at low delta-v values, the impact of load-limiting forces is minimal, but it increases as delta-v rises. No other parameters were modified from the GVI model between the simulations. This configuration of the GVI model resulted in an equal or lower estimated ENFR for the 50F occupant compared to the 50M occupant when modifying the load-limiting force and delta-v.

IV. **DISCUSSION**

Using a generic vehicle interior, KAN metamodels were developed to predict ENFR for two occupants in the driver position. Although these models exhibited higher estimation errors than the GPR metamodels, they offered the advantage of yielding explicit mathematical expressions to describe the influence of load-limiting forces and delta-v on thoracic injuries. The present study emphasised model interpretability by restricting the univariate function forms to linear and quadratic polynomials. However, more flexible mathematical representations could enhance accuracy. While such an approach might reduce prediction error, it would likely come at the cost of model interpretability and the ability to extract concise mathematical expressions. A comparable trade-off could also arise from increasing model depth (adding more layers), which may further reduce errors but obscure the interpretability of the model.

This study focused on the ENFR as the outcome variable to assess potential thoracic injuries in vehicle occupants. ENFR was chosen over more conventional probabilistic metrics, such as the risk of sustaining 2+ or 3+ rib fractures or AIS2+ or 3+ injuries, because these metrics can experience saturation effects at lower and upper extremes (0% and 100%). While it is possible to capture these saturation effects by using a more complex KAN model, this approach may result in a low interpretability of the model due to the use of additional layers or more intricate univariate functions. It is important to note that ENFR may not directly correlate with the injury metrics more commonly used in developing and evaluating restraint systems. While other metrics as the max. rib strain could be used, this requires a metamodel per rib and needs to be combined to have a useful injury metric. While this study primarily examined variations in load-limiting forces and delta-v and does not represent the whole vehicle fleet, the methodology is flexible and can be applied to analyse other real-world crash scenarios and their interactions.

The use of KAN metamodels needs further research to assess the effectiveness of this approach, especially considering the limited number of simulations available in the biomechanics field. This exploration may dive into various network configurations that could potentially decrease the number of simulations required while maintaining accuracy and interpretability.

V. REFERENCES

[1] Ziming L *et al.*, arXiv preprint: 2404.19756, 2024.

[2] Iraeus J *et al.*, Accid Anal Prev, 2016.

[3] Trummler L *et al.*, H2020 project VIRTUAL, 2020.

[4] Eriksson L, TUC web page, 2021.

[5] John J *et al.*, Front Bioeng Biotechnol, 2022.

[6] Iraeus J *et al.*, Int J Crash, 2015

[7] Iraeus J *et al.*, Int J Crash, 2021