INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED23
24-28 JULY 2023, BORDEAUX, FRANCE

A VALUE-DRIVEN DESIGN APPROACH FOR THE VIRTUAL VERIFICATION AND VALIDATION OF AUTONOMOUS VEHICLE SOLUTIONS

Bertoni, Marco (1);
Thorn, Stefan (2)

1: Blekinge Institute of Technology;
2: Volvo Autonomous Solutions

ABSTRACT

Autonomous vehicle solutions (AVS) are regarded as a major enabling technology to support the realization of 'total site solutions' in the construction equipment industry. Their full-scale deployment is hindered today by the need to test autonomous driving capabilities against the varying conditions an AVS is expected to be exposed to during its lifetime. Therefore, using virtual simulation environments is common to overcome the cost and time limitations of physical testing. A caveat in this virtual verification and validation (V&V) work is how to trade off the 'realism' of the V&V output (using high-fidelity models across many scenarios) against computational time. This research investigates expectations and needs for value-driven decision support in the virtual V&V process, proposing an approach and a tool to raise awareness among decision-makers about the value associated with using selected simulation models/components in the virtual verification and validation task for AVS. Verification activities performed on the initial prototype show that its main benefit lies in facilitating cross-domain negotiations and knowledge sharing when negotiating the desired features of the virtual simulation environment.

Keywords: Value-Driven Design, Case study, Systems Engineering (SE), Virtual Engineering (VE), Digital / Digitised engineering value chains

Contact:
Bertoni, Marco
Blekinge Institute of Technology
Sweden
marco.bertoni@bth.se

Cite this article: Bertoni, M., Thorn, S. (2023) 'A Value-Driven Design Approach for the Virtual Verification and Validation of Autonomous Vehicle Solutions', in *Proceedings of the International Conference on Engineering Design (ICED23),* Bordeaux, France, 24-28 July 2023. DOI:10.1017/pds.2023.44

1 INTRODUCTION

The ongoing digital servitization transformation is pushing solution providers to adopt innovative development approaches to manage the value co-creation process with customers (Struwe and Slepniov, 2023). As an example, autonomous vehicle solutions (AVS) are often regarded as a major enabling technology to support the realization of digital servitization business models in farming, forestry, construction, and mining (Leminen et al., 2022). In the latter two, AVS are often pinpointed as critical enablers for 'total site solutions' (Frank et al., 2019). Yet the full-scale implementation of such 'solutions' is limited today due to several concerns, no latter safety (Rezaei and Caulfield, 2021), which confine AVS to sealed-off production areas. One of the most critical aspects for the deployment of 'site solutions' is found in the limitations linked to vehicle-level testing. It has long been known (see: Koopman and Wagner; 2016) that it is infeasible to perform in-real-life physical try-outs thoroughly enough to ensure ultra-dependable system operation. There are simply too many arbitrary and/or dangerous situations that cannot be covered even by the most extensive physical testing, i.e., the latter is both impractical and too costly (Koopman and Wagner, 2016).
For this reason, virtual verification and verification (V&V) activities are becoming popular to inform decision makers about the resilience and scalability of AVS against the varying conditions they are expected to be exposed to during their lifetime. A caveat in this work is how to trade-off the ability of these virtual models to provide a realistic output against computational time (Schlager et al., 2020). Increasing the resolution and the level of detail of the virtual V&V models is not automatically a recipe for success, because this is also expected to increase computational time, which leads to a reduction in the number of scenarios that can be tested in a given timeframe (Schlager et al., 2020). Hence a question remains about where the optimum trade-off between 'fidelity' and 'scenario coverage' shall be found when designing a virtual V&V strategy and configuring its toolchain.
The aim of this research is to explore the use of Value-Driven Design (VDD) and, more in detail, qualitative value models, to guide early-stage decisions about how to design a virtual V&V strategy (including its simulation components) for AVS. The research question can be described as: 'How can qualitative value models be applied to communicate the value of a virtual V&V strategy to the decision makers?'. This objective is to present the results from a study conducted in collaboration with a Swedish provider of autonomous transport solutions for off-road applications. After presenting the findings related to the expectations and needs for value-driven decision support in the virtual V&V process, the paper describes a decision-making tool named the 'Value Visualiser for Virtual Vehicle Verification and Validation' (V6), which was developed in co-production mode with the company partner. The application of the V6 is exemplified in a case study related to the development of an autonomous dumper for offroad applications. The feedback from preliminary verification activities and pointers to future research are presented in last section of the manuscript.

2 V&V ACTIVITIES FOR AVS IN A VIRTUAL ENVIRONMENT

Nowadays, V&V activities for AVS are largely performed virtually in proprietary and open-source simulation environments, such as AirSim, CARLA or LGSVL Simulator. These systems are based on gaming engines, such as Unreal Engine® and Unity 3D®, which provide a flexible platform to parallelize the testing activities under a variety of possible traffic, lighting, and environmental conditions (Riedmaier, 2020). Here, virtual V&V activities can be decomposed into two methodologies: (1) integrated system, where the overall simulation toolchain is tuned to replicate a distinct manoeuvre, and (2) sub models-based, where each ingredient of the simulation pipeline is individually validated with respect to its physical counterpart (Donà and Ciuffo, 2022), as shown in Figure 1.
Virtual models for sensors (see: Schlager et al., 2020) and vehicles (see: Schramm et al., 2016) belong to 3 major categories; Low-, Medium- or High-fidelity models (also named as 'Black-', 'Grey-', or 'White-box' models). Low-fidelity vehicle models use a point-mass representation, while the sensor models are only able to detect objects that are inside the field of view, not occluded by any other object. Medium-fidelity vehicle models exploit single- or double-track chassis models with linear or non-linear tires, while the sensor models can detect the shape and texture of an object together with other environmental effects. High-fidelity vehicle models consider suspensions, drivetrains, wheels, and tires along with electronic controllers, while sensor models try to replicate the physical phenomena regulating the interaction between the sensors and the external environment in simulation.

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

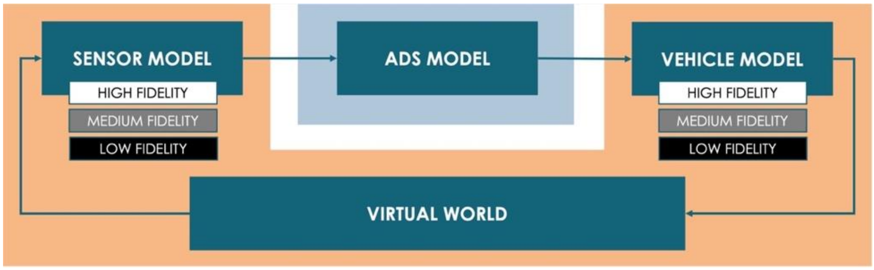

Figure 1. Simulation environments' setup for V&V of AV systems (Donà and Ciuffo 2022).

Table 1 exemplifies the pros and cons for Low-, Medium-., and High-fidelity sensor models, with emphasis on the computational power needed vs. the level of realism obtained by the simulation.

Table 1. Pros and cons of different fidelity levels (adapted form Schlager et al., 2020).

Model Fidelity →	**Low (black-box)**	**Medium (grey-box)**	**High (white-box)**
Operating principles	Geometrical aspects	Physical aspects, detection probabilities	Rendering (rasterization, ray tracing, etc.)
Input	Object lists	Object lists	3D scene (mesh)
Output	Object lists	Object lists or raw data	Raw data
Pros	Low computational power needed	Trade-off between computational power and realistic output.	Most realistic output
Cons	High abstraction level, no realistic output	Lots of training data may be required	High computational power needed

2.1 Major limitations of current virtual V&V processes

A simulation model, no matter how detailed, will always approximate the real-world phenomena and will only serve the need of the specific application it aims at replicating. For this reason, the literature stresses the context-dependent nature of the validity analysis, postulating that the absolute validity of a virtual testing toolchain is generally not achievable. As shown by Donà and Ciuffo (2022) an ultimate 'validation criterion' is not available. One a practical level, the validation procedure must be first tailored to the specific application domain, and one shall be aware that obtaining validation-grade data might be very challenging, extremely costly, or even impossible (e.g., for a system not yet existing), thus limiting the applicability of the analysis. Another gap identified in literature concerns the selection of scenarios used for AV validation. Most are staged in residential areas and motorways, neglecting off-road applications. Furthermore, their ability to handle 'exceptions' from the classic 'sunny day' (e.g., to include rain, snow, fog, ice, dust and more) is limited. Worse still is that the combinations of environmental factors, road characteristics and driving conditions that can occur are simply too many to enumerate in a classical written requirements specification. Perhaps not all combinations need to be covered if results are likely to be innocuous, but the requirements should be clear about what is within the scope of system design, as well as what is not (Koopman and Wagner, 2016). Furthermore, the literature rarely addresses the value of including new virtual model components (e.g., to simulate rain or fog) in the V&V simulation and does not elaborate on how to balance (computational) cost and value when designing a virtual V&V strategy. Little is said about how decision makers can be supported when weighting the pro and cons of choosing a computationally intensive alternative for the virtual simulation (e.g., using high-resolution simulation components) vs. a more approximated one.

3 RESEARCH APPROACH

The overall research effort can be framed in the Design Research Methodology (DRM) proposed by Blessing and Chakrabarti (2009) and is based on a single-case design (Yin, 2011). The research question was defined in collaboration with a company that develops and commercialises industrial autonomous transport solutions for offroad applications. This setting has provided a unique viewpoint from which to investigate the issue related to the value of virtual V&V modelling components. The way the research was conducted likens Participatory Action Research (PAR) (Argyris and Schön, 1989), with researchers and practitioners mutually involved in the research design and development.

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

Field data were collected through (1) regular bi-weekly virtual gathering, (2) the review of internal company documents, and (3) co-located workshops at the company facilities, to ensure triangulation. These data were used in the research to create preliminary demonstrators of a qualitative value model, that were discussed with a cross-disciplinary group of experts (about 20 people, participating physically and on-line) having knowledge in engineering design, vehicle dynamics, sensor modelling, virtual prototyping, and autonomy. The lessons learned gathered through the demonstration of emerging modelling concepts have allowed the researchers to close the look-think-act learning circles typical of PAR. The validation of the impact of the Prescriptive Study results (the V6 tool) corresponds to the 'Support Evaluation' phase of the DRM. Due to the lead time of new autonomous vehicles development projects this paper does not encompass the 'Descriptive Study II' stage that concerns the 'Application Evaluation' and the 'Success Evaluation' phases.

4 DESCRIPTIVE STUDY FINDINGS

The literature analysis and the case study data have both shown the need to work in a more systematic way when defining proxies for 'good' vs. 'bad' decisions in the design of a virtual V&V strategy. The practitioners raised two main issues in this regard: how to define a hierarchy of metrics/dimensions for ‘good decisions’ (considering all the involved parties), and how to visualise the trade-off among conflicting dimensions in a way to understand how much of a trade-off (i.e., between fidelity and computational cost) can be tolerated by the decision makers. The notion of Value Creation Strategy (VCS), as proposed by Monceaux and Kossmann (2012) and Isaksson et al. (2013), was extensively discussed in this regard. On the one end, the virtual simulation environment shall be configured so to predict with precision the behaviour of an AVS. For instance, the computer-based environment shall make possible to virtually test if the AVS is able to 'read' a scenario correctly, so to prevent accidents, as well as if it is able to continue its operations in front of non-critical disturbances (e.g., a bird standing on the track, a pile of soft snow), to maintain the desired level of productivity. Increased capabilities for testing such behaviours come with additional requirements, not only in terms of computational time and cost, but also with regards to the availability of human expertise, the accessibility to computational resources, and more. These aspects must all be factorised in when defining a V&V strategy and play an important role in the timely delivery of site solutions. Furthermore, more intangible aspects of value were discussed during the study. These include the opportunity of growing new knowledge on a particular software or model add-on, as well as the possibility to spin-off virtual V&V activities into new businesses, and more. The study further identified the need to be able to manipulate the VCS through an interactive user interface, so to simulate alternative scenarios in real time, to facilitate both communication and decision making in the cross-functional team.

4.1 Defining the value function for virtual V&V studies

Fidelity was indicated early on by the industrial practitioners as the main dimension for 'value' in virtual V&V activities. However, the widespread opinion among the industrial experts was that comparing virtual V&V studies using 'fidelity' as a metric is far from being a trivial task, mainly because a conclusive unit of measure for the concept is missing. Furthermore, even though several approaches in literature claim to provide a scalar number for fidelity, it remains still uncertain how to interpret such a number. For instance, experts and process owners have debated in the study what actions could be triggered by receiving the information that a simulation has received '79% fidelity' score. The latter neither clearly indicates the way forward when architecting a virtual V&V study, nor suggests how much 'fidelity' can be traded-off with cost, effort, and computational time. While the practitioners agree that such a trade-off is a moving target - following the evolution of computer technologies and the increase in the capacity of CPUs and graphical cards - it is still important to quantify the level of fidelity that is just 'good enough' for the purpose of a study. As pointed out by one of the respondents:

"If we want perfect fidelity, we'll probably need to run the (virtual) scenarios at a speed that is slower than real life. In fact, we want to run the simulation way faster than that."

Another issue being discussed is how to balance the number of scenarios being studied and their level of precision and completeness. A unique revelatory aspect connected with the case study is that virtual V&V activities shall be performed both for existing off-road sites (for which a detailed description of

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

exists) and for future sites (for which detailed information is not simply available). This is seen to significantly increase the number of scenarios to investigate and raises the level of complexity of the virtual V&V activity to the next level.

4.2 Assessing the value of virtual model components

Early on in the study, the practitioners identified 'weather' as a major virtual modelling capability of interest for off-road applications. For this reason, the study was directed towards exploring of how assess the value of rain, snow, fog, smoke, and dirt model add-ons for Unreal Engine® and Unity 3D® engines. All these add-ons introduce 'noise' in the simulation, which must be dealt with by the virtual sensor model (see Figure 1) to ensure the correctness of the behaviour prediction (Hospach et al., 2016), such as recognizing obstacles and deciding if and how to reroute the AVS. A main reason for prioritizing this 'category' of models is that, although many simulators come with weather models, currently used image data sets are recorded largely under good weather conditions (Cordts et al., 2016). Even though von Bernuth et al. (2019) show the degradation of object detection quality with growing intensity of the weather conditions, physically correct influences on the optical sensor are often lacking.
The descriptive study went on to investigate how much does ‘value’ increase with increased capabilities/technical requirements of these model components. The demonstrators being developed during the study played with the idea that it should be possible to benchmark the value of alternative weather-related add-ons to identify the one that, given the specific requirements for the virtual V&V work, provides the highest value for the application. A caveat in this work is that simulating thousands of rain drops or even trillions of fog water particles and the corresponding light rays - as well as their interactions - is, while possible, terribly time consuming. The descriptive study showed that design decision support is needed to help the cross-functional team in identifying an optimum when it comes to the accuracy of a virtual model component vs. the computation continuum - and possibly against a longer list of value creation factors (as captured in the VCS).

5 RESULTS: THE V6 QUALITATIVE VALUE MODEL

The study brought to the definition of a qualitative value model named the 'Value Visualizer for Virtual Vehicle Verification and Validation' (V6). The process steps shown in Figure 2 are designed to inform decision makers about the value contribution associated to the utilization of virtual simulation model components in V&V tasks. The V6 performs three main functions:

- correlate the characteristics of one or more off-road scenarios for autonomous driving to the VCS for virtual verification and validation;
- provide a scalar score to represent the value of virtual model components against the selected VCS, together with a maturity score to indicate how much the value score can be trusted;
- support the codification and storage of the knowledge that describe the relationship between the characteristics of the model add-ons and the value drivers of interest.

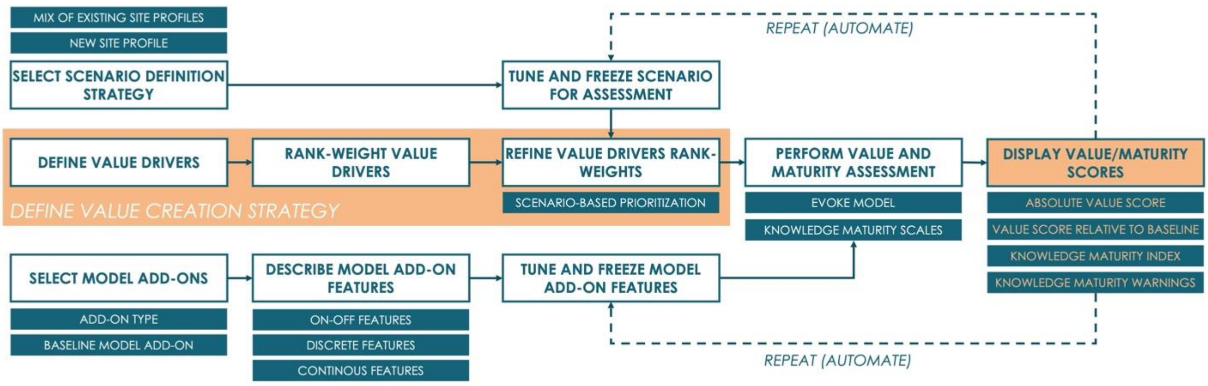

Figure 2. Process steps of the V6 approach.

The V6 demonstrator described in the section below is composed of a method, a tool and usage guidelines, and targets the evaluation of virtual model components used to simulate various weather conditions in a confined working area (e.g., a quarry). In its current version, it can evaluate 2 classes of model add-ons (a snow and a rain model) and includes 63 value drivers across 8 value dimensions.

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

5.1 Eliciting the VCS

The V6 aims to tap into the tacit knowledge of the cross-functional team to codify all factors of interest on which the value analysis shall be performed. Eliciting the VCS is designed as a 4-step process, which is conducted in cross-functional workshops or through a larger crowdsourcing activity.

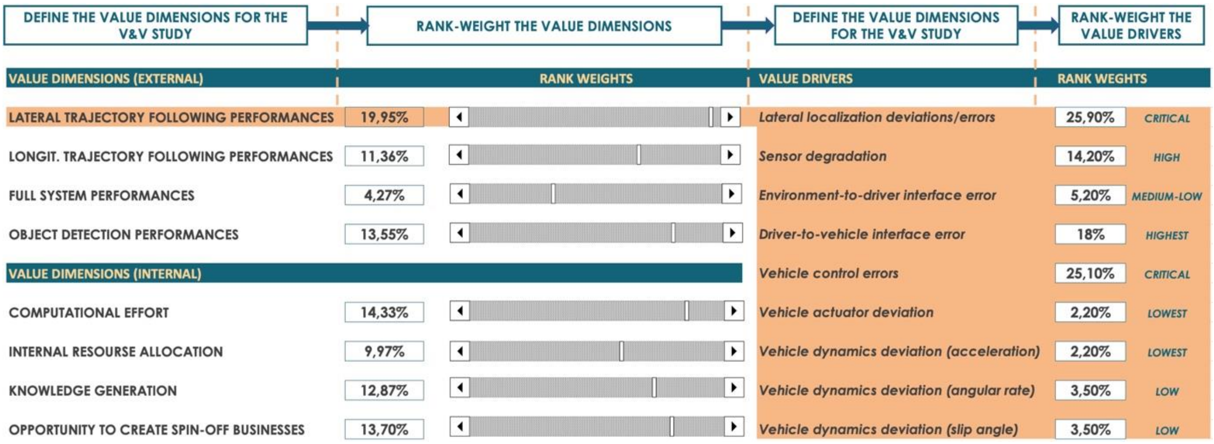

Figure 3. Extract from the VCS for the electrical autonomous dumper.

Figure 3 exemplifies the construction of the VCS for an autonomous electric dumper, which is a complete solution designed to operate autonomously in a quarry or a mine. The first step foresees the team making clear what are the value 'macro-categories' to be considered in the virtual V&V tasks. These 'value dimensions' cover aspects related to the functionality of the simulation (e.g., the ability to detect objects), non-functional aspects (e.g., from knowledge generation to resource allocation) and cost/effort-related aspects (e.g., computational time). Each dimension is rank-weighted by the team using ad-hoc methods - from multivoting to Multi Attribute Decision Making (MADM) techniques such as the Analytic Hierarchy Process. The team can also define more specific 'value drivers' for each dimension to increase the granularity of the assessment, which are in turn rank-weighted using MADM or other methods.

5.2 Describing the scenarios and the virtual model add-ons

Figure 4 (left) shows how a virtual snow model component is described in the V6, and how this can be enabled/disabled, as well as tuned in its specifications when conducting the assessment.

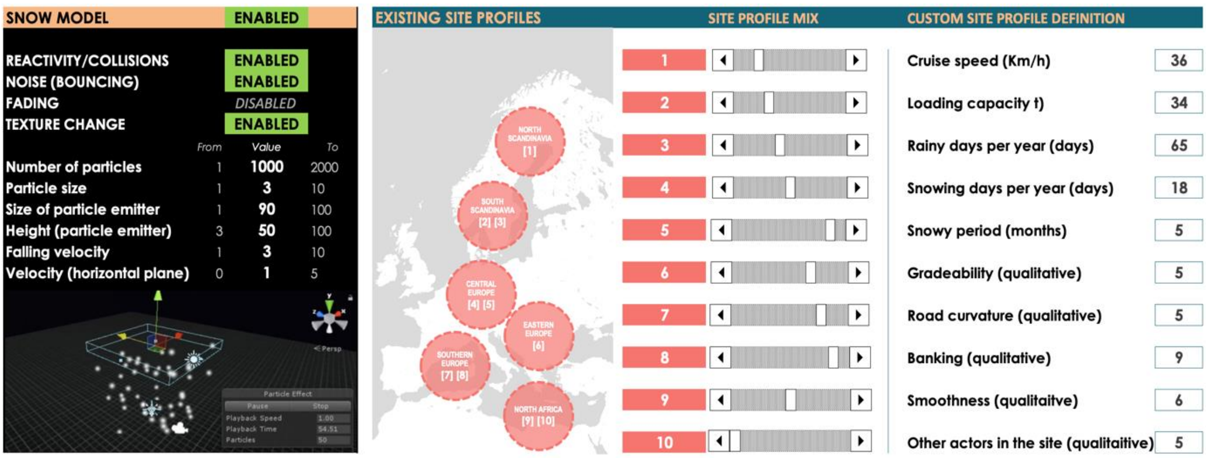

Figure 4. Extract of the model add-on (left) and scenario (right) definition modules.

Different configurations for the same component can be generated by fine-tuning the list of features that describe how this is implemented in the simulation environment. These are parametrised by using continuous, discrete, or binary (on/off) variables. Collision physics, fading and texture change are examples of binary features that can be enabled/disabled by the decision makers, while particle size and falling velocity can be fine-tuned to match a specific value. In practice, a Design-of-Experiment

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

(DoE) approach is often applied to reduce the number of combinations being investigated. Surrogate modelling techniques can then be deployed to support the identification of (what could be considered) the optimum point in the design of the component.

The descriptive study has also shown that a virtual model component holds an intrinsic and an extrinsic value. A component can be intrinsically valuable for a VCS (e.g., a 'reactive snow' feature is highly relevant to virtually verify longitudinal trajectory deviations), yet its implementation in a virtual validation toolchain makes little sense if the target market is located in a tropical region where snowy conditions are not an issue. Similarly, a fog model might be of little significance if the targeted application domain is rarely affected by the phenomenon. For this reason, the right-end side of Figure 5 shows how decision makers can input the characteristics of the scenario in the value analysis, by defining a site profile to fine-tune the weights of value dimensions and drivers. Such a profile can be inputted manually, using both qualitative and quantitative information across a set of parameters for a site, or can be automatically generated from existing site data. Site profiles descriptors are agreed upon by the cross-functional team dealing with V&V activities and typically include road, weather, and other environmental conditions of interest, modelled from existing datasets.

5.3 Assessing the value of the model components

The description of an add-on weather model component is inputted in a Multi Criteria Decision Making (MCDM) matrix adapted from the EVOKE model proposed by Bertoni et al. (2018), and originally by Eres et al (2014). The matrix is implemented in MS Excel® and computes a value score associated to a given configuration together with a knowledge maturity index, which indicates how much the results of the assessment can be trusted. The decision makers must first indicate how much an add-on model feature is expected to impact a given value driver, typically by using a high, medium, and weak or null correlation, expressed through a coefficient (0.9, 0.3, 0.1, and 0 as in Figure 5).

To compute a scalar score for each proposed component configuration, they must also indicate the type and shape of value function that links the model feature with the value driver of interest. The V6 uses four different functions to calculate a performance value score for a proposed design: maximising (Max), minimizing (Min), optimizing (Opt), and avoiding (Avo). The first one describes a situation when value is increased by increasing the number associated to virtual model component feature (i.e., the maximum level of fulfilment is reached when the value of the parameter gets close to infinity). In the case of a minimization functions, the logic is opposite (i.e., the maximum level of fulfilment is reached when the value of the parameter approximates zero). In the case of an optimization function, the maximum value is achieved when a model component parameter reaches a target value set by the cross-functional team. An avoidance function follows the opposite logic than the optimization one.

		REACTIVITY/COLLISIONS	ENABLED			
LATERAL TRAJECTORY FOLLOWING PERFORMANCES	*Normalised weights*	Correlation	Function type	Function shape	Maturity index	Value score
Vehicle actuator deviation (torque)	8,4%	0,30	MAX	STEP	2	100,00%
Vehicle actuator deviation (brake force)	18,7%	0,90	MAX	STEP	5	100,00%
Vehicle actuator deviation (steering angle)	1,9%	0,90	MAX	STEP	3	100,00%
Vehicle dynamics deviation (lateral acceleration)	22,5%	0,90	MAX	STEP	4	100,00%

		Particle size		3,00		
OBJECT DETECTION PERFORMANCES	*Normalised weights*	Correlation	Function type	Function shape	Maturity index	Value score
Sensor degradation	16,4%					0,00%
Field of view	2,5%	0,10	MIN	DEFAULT	3	71,94%
Occlusions	22,4%	0,90	MIN	CONCAVE	2	66,14%
Visibility	41,0%	0,30	MIN	LINEAR	5	77,78%

Figure 5. Extract from the EVOKE model linking value drivers and add-on model features.

In the V6, the value function can assume 4 different shapes: linear, concave, convex and step, to better capture the non-linear dependencies between value drivers and model component features. Step functions, for instance, are employed to model binary variables, such as a model component feature being activated or not (e.g., the Reactivity/collision feature presented in Figure 4). A Knowledge Maturity Index (KMI) is computed together with a value score to indicate how much the cross-functional

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

team might trust the results of the value assessment activity. Following the approach proposed by Johansson et al. (2017), the team is required to assess the maturity of the knowledge used at each intersection in the matrix to set correlation coefficients, value functions and their shapes. This is done by using a qualitative scale - from 1 (minimum) to 5 (maximum) - computed over three dimensions: input, method (tool), and expertise (experience). If the KMI is below a critical value set by the team, the model issues one or more 'warnings' that are presented to the decision makers together with the total KMI index for a configuration, which is obtained by aggregating the scores from each individual assessment. Both the value score and the KMI are logged at each iteration and used to support the selection of the best design (or to pinpoint areas where knowledge is not sufficient to decide on the most value-adding design). The results form the value assessment activity are visualised to the design team in a way to emphasise how a new model configuration is related to a given baseline design.

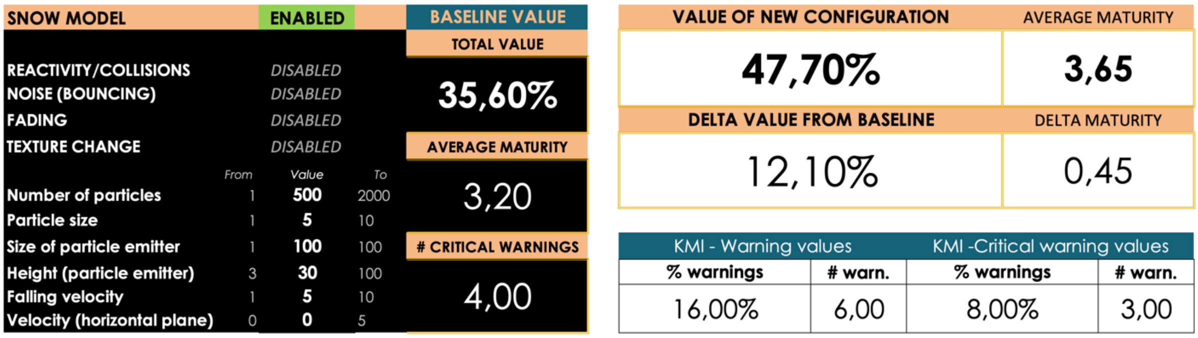

Figure 6. Results visualization panel for the V6 tool.

The panel on the left-end side of Figure 6 displays the features of a baseline design together with its calculated value and maturity information. The panel on the right-end side displays the results for a new configuration, pinpointing the delta with the baseline design as well as providing in depth information about the KMI warnings. It is further possible to visualise the value contribution across each driver by clicking on the interface.

6 DISCUSSION

The goal with the development of the V6 was to create a 'demonstrator', which has a primary use as communication tool, to provide evidence of product benefits (see Bobbe et al. 2023). The V6 is mostly aimed at communicating the hypothesis that value models are beneficial to shape the virtual V&V strategy for AVS, rather than at evaluating such hypothesis. Being a demonstrator and not a full-fledged prototype, the V6 wants to engage practitioners and experts in providing cues on how to interact with qualitative value models, so to better understand (and lower the barrier related to) how to introduce a value thinking in the V&V task. The design principle of 'try' (see: Bobbe et al. 2023) was at the core of every iteration of the tool, which was first and foremostly designed with the intention of letting the potential users to try the solution and experience it first-hand.

The verification activities performed on the V6 show that the tool is deemed suitable by industrial practitioners to raise awareness about the desirability, viability, and feasibility of adopting alternative model components in the virtual V&V toolchain. The interaction with the practitioners showed that the main advantage of value-driven design approach is found in the possibility to support a set-based engineering approach early in the V&V strategy design process, helping the team in filtering out those strategies that are considered to be of low 'value' for both customers and solution providers. A qualitative value-based approach was found to be mainly beneficial for the decision makers when elaborating on (1) what model components shall be included in the V&V work, depending on the specific product and business case, and (2) the resolution of such components, in terms for trade-off between realism vs. computational time and effort (and more, depending on what the provider considers as an internal value dimension of interest). The V6 was found to help individuals in learning about dependencies (and specify differences) across the organizational boundaries, which is to catalyse the tacit knowledge related to the virtual V&V task and to communicating the rationale of V&V decisions to those roles that do not have in depth experience with the technicalities of the virtual testing platform.

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

On the other end, the work pinpointed issues when it comes to the utilization of qualitative value models in the V&V process. An important lesson learned from the work is that it is not intuitive for several roles and expertise in the cross-functional team to understand the contribution of models such as the V6 in the virtual V&V toolchain. Simply stated, high-fidelity, white-box models are perceived by many as 'best' no matter what, and one shall always strive for obtaining the most realistic output from the simulation work. Even though the descriptive study shows that challenges do exist when striving from absolute realism, more work needs to be done to adapt qualitative value models to the needs of the simulation experts and to integrate the tool with the other items in their engineering toolbox. More specifically, when looking at the V6 demonstrator, the ability of modularizing the approach, being applied to different sites and being scaled up through the assessment of a large number of model components, has room for improvement. At the same time, the value scored obtained from the model shall be better related to the concept of 'fidelity', which was found to be a particularly popular 'boundary object' in the cross-functional in the team.

7 CONCLUSIONS

Virtual V&V activities for AVS are becoming increasingly common, yet it remains unclear how the increasing level of granularity and detail of a simulation shall be compared against computational power and cost. The findings from the work show that the knowledge base on which V&V decisions are taken (e.g., with regards to what model components shall be included in the virtual validation toolchain) needs to be expanded, so to raise awareness about a broader set of factors critical for the successful deployment of servitized solutions to the customers. Since the optimum is not found at the end of the scale (i.e., maximum realism) but rather as a trade-off between such dimensions, qualitative value models become of interest to identify the most value adding virtual V&V strategy. The paper shows that the development of such a strategy shall be treated as a value-driven design problem, balancing precision and effort in the virtual modelling work. The study further demonstrates the applicability of a value modelling approach to support the design team in identifying the characteristics of the modelling environment that are believed to deliver the best value for the system decision makers. The sweet spot for the application of qualitative models is found in the earliest stages of the virtual verification and validation process, when quick what-if analysis are performed to filter out those strategies that are believed to be too computationally intensive, or, on the other end, not enough 'realistic' for the scope of the V&V activity. Noticeably, the notion of Knowledge Maturity (KM) was also found to be very relevant to provide a trustworthiness measure of the value modelling results, highlighting those areas needing more knowledge about the relationship between virtual model components and V&V strategy goals.

Future work will aim at further developing the proposed demonstrator and at expanding its application outside the offroad sector. They will also aim at scaling up and expanding its coverage, applying it not only in the frame of the evaluation of virtual simulation model add-ons, but also to assess the value of other aspects of the autonomous transport system (such as connectivity protocols and more) to orient early-stage design decisions towards more value adding solutions. Experimentation activities in selected design episodes - typically through protocol analysis - are also in focus to further verify how the tool is used (and its main benefits and drawbacks) in different contexts.

ACKNOWLDGEMENTS

The research work presented in this paper was performed within the 'Validering av virtuella modeller för simulering av autonoma fordonssystem' (VaViM) project funded by the Swedish Innovation Agency (VINNOVA) through the Strategic innovation programme 'Drive Sweden'.

REFERENCES

Argyris, C. and Schön, D.A. (1989), "Participatory action research and action science compared: A commentary", *American behavioral scientist*, Vol. 32, No. 5, pp. 612-623. https://doi.org/10.1177/0002764289032005008

Bertoni, M., Bertoni, A. and Isaksson, O. (2018), EVOKE: "A Value-Driven Concept Selection Method for Early System Design", *Journal of Systems Science and Systems Engineering*, Vol. 27, No. 1, pp. 46-77. https://doi.org/10.1007/s11518-016-5324-2

https://doi.org/10.1017/pds.2023.44 Published online by Cambridge University Press

Blessing, L.T. and Chakrabarti, A. (2009), *DRM: A Design Research Methodology*, Springer, London, UK. https://doi.org/10.1007/978-1-84882-587-1_2

Bobbe, T., Opeskin, L., Lüneburg, L. M., Wanta, H., Pohlmann, J., & Krzywinski, J. (2023). "Design for communication: how do demonstrators demonstrate technology?". Design Science, Vol. 9, e3.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R. and Schiele, B. (2016), "The cityscapes dataset for semantic urban scene understanding", *IEEE Conference on computer vision and pattern recognition*, pp. 3213-3223.

Donà, R. and Ciuffo, B. (2022), "Virtual testing of Automated Driving Systems. A survey on Validation Methods", *IEEE Access*. Vol. 10, pp. 24349-24367. https://doi.org/10.1109/ACCESS.2022.3153722

Eres, M.H., Bertoni, M., Kossmann, M. and Scanlan, J. (2014). "Mapping customer needs to engineering characteristics: an aerospace perspective for conceptual design", *Journal of Engineering Design*, Vol. 25, No. (1-3), pp. 64-87. https://doi.org/10.1080/09544828.2014.903387

Frank, M., Ruvald, R., Johansson, C., Larsson, T., & Larsson, A. (2019). Towards autonomous construction equipment-supporting on-site collaboration between automatons and humans. International Journal of Product Development, 23(4), 292-308.

Hospach, D., Mueller, S., Rosenstiel, W. and Bringmann, O. (2016), "Simulation of falling rain for robustness testing of video-based surround sensing systems", *IEEE Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 233-236.

Isaksson, O., Kossmann, M., Bertoni, M., Eres, H., Monceaux, A., Bertoni, A., Wiseall, S. and Zhang, X. (2013), "Value-Driven Design–A methodology to Link Expectations to Technical Requirements in the Extended Enterprise", *INCOSE International Symposium*, Vol. 23, No. 1, pp. 803-819. https://doi.org/10.1002/j.2334-5837.2013.tb03055.x

Johansson, C., Wall, J. and Panarotto, M. (2017), "Maturity of models in a multi-model decision support system", In DS 87-6 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 6: Design Information and Knowledge, Vancouver, Canada, pp. 237-246.

Koopman, P. and Wagner, M. (2016). "Challenges in autonomous vehicle testing and validation", *International Journal of Transportation Safety*, Vol. 4, No. 1, pp. 15-24. https://doi.org/10.4271-2016-01-0128

Leminen, S., Rajahonka, M., Wendelin, R., Westerlund, M. and Nyström, A.G. (2022), "Autonomous vehicle solutions and their digital servitization business models". *Technological Forecasting and Social Change*, Vol. 185, 122070. https://doi.org/10.1016/j.techfore.2022.122070

Monceaux, A. and Kossmann, M. (2012), "Towards a Value-Driven Design Methodology–Enhancing Traditional Requirements Management Within the Extended Enterprise", *INCOSE International Symposium*. Vol. 22, No. 1, pp. 910-925. https://doi.org/10.1002/j.2334-5837.2012.tb01379.x

Rezaei, A. and Caulfield, B. (2021), "Safety of autonomous vehicles: what are the insights from experienced industry professionals?". *Transportation research part F: traffic psychology and behaviour*. Vol. 81, pp. 472-489. https://doi.org/10.1016/j.trf.2021.07.005

Riedmaier, S., Danquah, B., Schick, B. and Diermeyer, F. (2021), "Unified framework and survey for model verification, validation and uncertainty quantification", *Archives of Computational Methods in Engineering*, 28(4), 2655-2688. https://doi.org/10.1007/s11831-020-09473-7

Schlager, B., Muckenhuber, S., Schmidt, S., Holzer, H., Rott, R., Maier, F. M., et al. (2020), "State-of-the-art sensor models for virtual testing of advanced driver assistance systems/autonomous driving functions", *SAE International Journal of Connected and Automated Vehicles*, Vol. 3, pp. 233-261. https://doi.org/10.4271/12-03-03-0018

Schramm, D., Hiller, M. and Bardini, R. (2016), *Vehicle Dynamics: Modeling and Simulation*, Springer Berlin, Heidelberg.

Struwe, S., & Slepniov, D. (2023). Unlocking digital servitization: A conceptualization of value co-creation capabilities. Journal of Business Research, 160, 113825.

von Bernuth, A., Volk, G. and Bringmann, O. (2019), Simulating photo-realistic snow and fog on existing images for enhanced CNN training and evaluation, *IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 41-46. https://doi.org/10.1109/ITSC.2019.8917367

Yin, R.K., *Case study research: Design and methods*, Sage publications Ltd, London, UK.