Adaptive Models for Safe Maintenance Planning of Cyber-physical Systems

Manuel Müller*, Andreas Löcklin*, Nasser Jazdi*, Michael Weyrich*

**Institute of Industrial Automation and Software Engineering, University of Stuttgart, Stuttgart, Germany (e-mail: manuel.mueller@ias.uni-stuttgart.de).*

Abstract: The progress of digitalization and Internet of Things enables more and more complex, networked and powerful Cyber-physical Systems (CPS) operating in uncertain environments. This complexity and uncertainty, however, makes it unfeasible to model every aspect in advance. This causes the models to leave their scope and reach their capability limits. Specifically, in safe maintenance planning for highly-automated trucks, this fact causes waste of valuable resource, since maintenance models are often more rule-of-thumb (e.g. operation hours) than precise. In order to counteract this issue, we propose extending the digital twin concept by artificial intelligence such that the models get *dynamic* and *adaptive*. Having described the general approach and its architecture, we showcase and evaluate the approach in a highly-automated truck scenario.

Keywords: adaptive models, automated system, safety, maintenance planning.

1. INTRODUCTION

The progress of digitalization and Internet of Things enables more and more complex and powerful Cyber-physical Systems (CPS). This technology push awakens great hopes especially with regard to increasing efficiency and new fields of application. In particular, people envisages the use of such CPS in uncertain and dynamic environments. However, according to the survey of (Luo et al., 2019), there are still large unsolved questions. This is especially true in the area of safety and the late product life cycle phases focusing on maintenance. Classic CPS execute complex but fixed procedure in static environment. They often occur in the domain of production, but modern trucks, which we take as an example, count as CPS, too. Classic safety and maintenance planning concepts focus such CPS. The safety process usually starts with planning phase and ends with certification. Afterwards the operating phase begins where the maintenance plan created as part of the safety analysis takes effect. Maintenance often take place at specified intervals based on static models. Under the assumption of static conditions and with the appropriate experience, this approach works very well. However, for the safe maintenance planning of the future, this approach leads to a dilemma. Maintenance costs time and money. Especially safety-critical assemblies must be replaced in advance. Therefore sometimes assemblies are replaced that, in retrospect, would still have had a considerable service life. Thus, valuable resources are wasted. This is where the demand for increased efficiency arises. However, since considering a system in an uncertain environment, the safety engineer must assume the worst-case scenario of use and estimate wear and tear very conservatively. Especially due to the uncertainties, large discrepancies between the a priori planned maintenance and a posteriori determined lifetime of the components can be expected.

Our approach to the solution of this problem is the consideration of *dynamic* and *adaptive models.* These models automatically evaluate the process data of the CPS, thus taking into account the context and the concrete type of use. Consequently, they enable a more precise (worst-case) estimation. We build on the concept of the digital twin, which provides a detailed reproduction of system components. However, according to our survey (Löcklin et al.), current approaches hardly exploit the potential of the digital twin. Furthermore, digital twins do not automati-cally adapt to previously unknown dependencies between models and emerging influencing factors. If, for example, a deteriorated shock absorber influences the wear of a wheel, the digital twin usually does not model this effect and the connection between wheel and shock absorber if it has not been modelled before. However, modelling all details in advance is impracticable (West and Blackburn, 2018). For this reason, we propose to extend the digital twin concept by aspects of artificial intelligence (AI) according to the Intelligent Digital Twin's architecture (Ashtari Talkhestani et al., 2019) to provide adaptive models. Detected anomalies (monitoring) are divided into three groups, namely (1) little deviations from prediction (2) fundamental deviations from prediction, and (3) short, but intense deviations with random characteristic. Based on this differentiation, we present a holistic architecture handling all three cases individually, analysing them with rule extracting algorithms and directed graph reconstruction, and thus uncover yet unhandled phenomena. We investigate this issue in the further course of this paper by means of the maintenance planning for a highly automated truck. For this reason, maintenance work can only be carried out at dedicated maintenance locations (Allal et al.), so both the estimation of the service life of wear parts and estimated journey time to next location becomes relevant from a safety point of view. However, the service life of many wearing parts such as brake pads depends much more on their actual use than on static metrics like operating hours. Driving through congested urban areas with frequent stop-and-go phases is much more stressful, high fuel consumption and erosion than smooth driving over free highways. The totality of the individual influencing factors is strictly dependent on the use case and therefore

unknown at design time. At operating time, however, these influencing factors can be read out of the operating data, the digital twin can adjudge life cycle models accordingly and thus provide more precise and reliable predictions.
The simplified logistic scenario is detailed described in Section 2.1. Reference to related work follows (section 2.2). Subsequently, we explain our concept (Section 3) and present simulated results of the scenario (Section 4). Finally, a conclusion follows (Section 6).

2. SCENARIO AND RELATED WORK

In this section first the scenario and the underlying problem is described (section 2.1). Afterwards the reference to related works is established (section 2.2).

2.1 Scenario of safe maintenance planning for trucks

In the scenario, we assume an autonomously driving truck that delivers parts to various customers. On its way, the truck stops at maintenance bases where wear parts can be replaced and other maintenance work can be carried out. Therefore, the autonomous truck requires autonomous maintenance planning. The truck mainly drives on similar routes. The sections of the routes contain physical properties such as the gradient or the quality of the road surface. There are also other properties such as traffic density. These properties are statically assigned to a route section and are always similar within a statistical fluctuation range. Occasionally, however, the road sections traveled change, for example due to road closures, customer fluctuation, priority change etc., or the properties of the road sections change, for example due to road renovation. In addition to the rather slowly changing properties, events such as traffic jams or rain etc. can occur. Characteristic for events is that they occur stochastically distributed over different road sections, that they have a short but significant influence on process data and that they affect the target value, the wear. The autonomous truck reacts to these influences automatically. For example, it reduces its speed when it rains. It is obvious that it is not possible to manually build a separate model for each route section and each event, which models the environment and system reaction. Instead, the models themselves must adapt to the new conditions without having to rely on external help. As already mentioned, the goal of the modelling is to adequately estimate the wear of the components and to choose the maintenance time with the appropriate reliability. Note that the maintenance planning of an autonomously driving truck is not only an availability problem but also a safety problem. Due to a lack of on-board personnel, failures on the track may result in long intervention times. Traction failures caused by material fatigue therefore may endanger other road users, especially at railroad crossings or in rescue routes etc. That is why maintenance planning must fulfil safety requirements. For this purpose, on the one hand, the approach predicts the arrival time at the destination and on the other hand, it individually evaluates the wear and tear based on the route travelled.
The main idea in this scenario, namely to perceive the parameters of the surrounding system, analyse them, draw conclusions from them and then adjust the models and their relations, is cross-domain. This idea can be transferred to all areas of automation technology, e. g. manufacturing etc.

2.2 Related work

The shortcomings of traditional safety concepts for autonomous systems especially when it comes to safeguard dynamically made decisions is still an open issue. Today's safety analysis tools are often not suitable for real-time application. To still meet the need for safety in unknown environments, (Bajcsy et al., 2019) propose a real-time safety analysis based on Hamilton-Jacobi accessibility to calculate the Backward Reachable Set in real time. The basic idea of this approach is to build a specific model, i.e. an environment map, according to a fixed scheme at runtime and to identify a secure environment in this map using the Backward Reachable Set. A different approach, but in principle with the same goal, namely the control of uncertainties (Chen et al., 2019). They extend existing imitation learning methods with better performance and a safe set based safety controller. Speed and distance-dependent ellipses define the safety areas that are not touched by safe set-based safety controllers. A third approach (Magdici and Althoff), also based on safe sets, provides an emergency trajectory in case an unexpected event occurs. All three approaches have in common that they use variable models for a very specific sub-area, navigation, to ensure the safety of the vehicle during operation. Another piece of the puzzle is the safety monitoring of models. For example, (Machin et al., 2018) contribute to this topic. Their Safety MOnitoring Framework (SMOF) automatically generates safety regulations based on the concept of safety margins. The approach builds on a hazard analysis and formal verification techniques to synchronize the regulations. However, the approach assumes trusted information sources. (Müller et al., 2019) contribute to the trustworthy merging of different information sources based on subjective logic. (Di Franco and Bezzo, 2020) present an interesting avionics approach to human interpretability. They propose to apply decision trees at runtime to monitor the system behaviour in an interpretable way and thus avoid collisions of their quadcopter. The algorithms classify the trajectories as safe or unsafe depending on the training set trajectories generated. If a planned trajectory becomes too similar to an unsafe trajectory, the system considers this trajectory unsafe, too and triggers re-planning. In this case, Di Franco and Bezzo propose to use a different decision tree to evaluate the trajectory, which is most similar to the current one and manoeuvre the drone out of the risk situation. Since the decision trees are rule-based in principle, they argue that, unlike modern black-box models of artificial neural networks, they offer interpretability.

Thus, there are approaches that already propose adaptive models to ensure safety during operation. However, these approaches mainly refer to motion planning, not to maintenance planning of safety-critical components. For this reason, specific models are also focused on, namely mainly card and motion models that expand according to specific processes. There are also approaches that monitor the safety-related behaviour and make new decisions based on the situation. In addition, there are approaches like by (Müller et al., 2019) which evaluate the information’s value of a data source and reclassify it during operation. There are also efforts to make this monitoring interpretable for humans. However, to the best of our knowledge, there are no holistic approaches,

combining adaptive models, model monitoring and interpretability for lifetime estimation of safety-critical parts under wear. Therefore, the authors contribute the concept descripted in Section 3 to close this gap.

3. CONCEPT

The architecture for dynamic and adaptive models to achieve a reliable lifetime estimation of safety-critical parts is shown in Fig.1. It follows the MAPE-K-scheme (monitor, analyze, process, execute, and knowledge) introduced by autonomic computing (Kephart and Chess, 2003). Monitoring, analysis and information processing exploit machine learning algorithms in order to identify shortcomings in the models and automatically fix the issues. Execute module follows static procedures to handle the generated knowledge gain. The Knowledge is managed according to the concept of the digital twin (Ashtari Talkhestani et al., 2019). For the normal operation, physical models calculate the lifetime estimation of the wear parts and the operation hours until next maintenance. Context model choose the appropriate physical models and their parameterization while the state model activates the needed context models. In contrast, event models serve for handling specific exceptions. They are introduced in detail later on. The information model links the specific models to the process steps, models to models, and (context) models to historical data. This creates a structure of loosely networked models that can be reorganized dynamically. The intelligent algorithms from the MAPE steps, on the other hand, use this plasticity to dynamically adapt the models to the system and its environment and to integrate new findings.

The monitoring has the responsibility to detect abnormal behavior of the system. Therefore, different algorithms from simple thresholds to sophisticated ones like by (Lindemann et al., 2019) surveil the process data. The selected algorithms depend on the input values. We discuss the concrete choice for this work in Section 4. From the Anomaly Detection Module the system differentiates three conditions:

(1) Little deviations from prediction
(2) Fundamental deviations from prediction, and
(3) Short, but intense deviations with random characteristic.

In order to differentiate the three states, the spread monitoring algorithms position the process data in the light of the historical curves and automatically derive the decision boundaries. Little deviations from estimation (1) refers to the normal case, where the module detects no anomaly. In this case, in the analysis step, stability analysis module increases the stability metric of the models involved in the prediction. Processing state continues training the physical parameters of the physical models to further improve the prediction quality. In contrast to the first case, (2) fundamental deviations show off in anomaly detection. They occur when either the system assigned a wrong context to a given situation or the system faces a persistent environmental change. Referring to our scenario, the latter could be the effect of rearranging the routes due to customer fluctuation or diversions. To react to those situations our approach instantiates a new context model and a new physical model. This event comes with high safety margin since the system has little information and thus estimates conservatively. To avoid this effort if possible, in the analysis step context analysis algorithms search the other context models. Simulating the prediction of the other context models, in the simplest case, Context analysis module diagnoses the confusion of the contexts. In the following processing step, Correlate States module inserts a new link from state model to the newly discovered context model. Otherwise, the closest context models and therefore physical models serve as template for the new model instances. In this case, the Directed graph reconstruction algorithms of the processing step come into play. The core idea of the directed graph reconstruction algorithms is to extract a directed graph from the process data where the nodes of the graph represent intersection points with other processes (i.e. driving other routes with their individual manifestation) and the edges model the characteristic curves of the process data. The output of the directed graph reconstruction algorithms, obviously a graph, principally serves as the system's state model. However, process step duplicates are not yet handled. That is why the state correlation algorithms are subsequently involved identifying and merging duplicates and thus reducing the graph.

Finally, the third case namely short but intense deviations with random characteristic (3) is considered. In case of trucks, this could be spontaneous traffic jam. The core problem is often a missing information. You cannot measure this directly, it

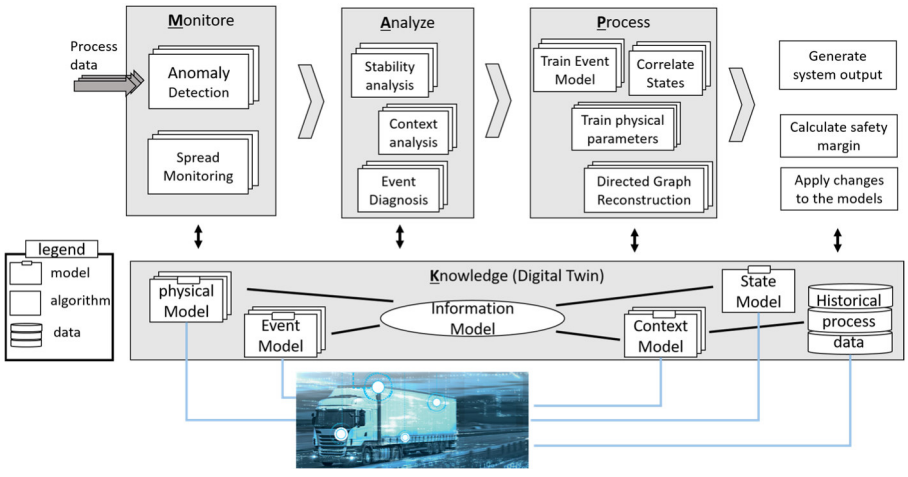

Fig. 1: Following the MAPE-K scheme, this architecture realizes the model extension for safe lifetime estimation.

happens suddenly, somehow randomly (e.g. concerning street segment) and it shows off in an anomaly (e.g. vehicle speed). Nevertheless, there are significant correlations to certain contexts (e.g. rush hour). We refer to this as event. We define an event, in the common sense "something that happens, especially when it is unusual or important" (Collins English Dictionary) in this context as the triple E = {P, R, I}, where P is the event's characteristic pattern (input), R is a rule defining when the event occurs, and I is the influence of the event on certain values (outcome). Note that the precision of the event's parameters increase with the number of observed events. The event diagnosis module of the analysis step uncovers those events checking for sudden, short but high anomalies occurring on stable models. However, events show off patterns, which algorithms can uncover. Therefore, in the processing step, a two-step process trains the event models. First step is to uncover rules. A combination of RUDE (Lud and Widmer, 2002) and Sequential Covering (SC) (Halpern, 1977) provide human-interpretable rules describing the anomaly. Second step is determine the effect on the target value. Artificial neural networks can do this.

The execution step does not infer artificial intelligence but executes the actual actions, calculate the safety reserves and apply the changes to the models according to the previous training step. The execution step may be realized in a semi-autonomous manner, where human operator has to confirm the model extensions or in fully autonomous way where the suggested actions are directly executed. A special focus lies on the calculation of the safety margin. Here, the information obtained from the available and generated models come into play. The spread monitoring not only provides the basis for estimation of the stochastic noise distribution of the physical models or the spread in the characteristics of an events and the event's influences. Combined with the analysis step, the monitoring also provides information of the probability of events occurring assuming a given process. Moreover, the graph reconstruction checkpoints testing the predictions. This allows to calculate worse case and realistic estimations of the metrics of interest (i.e. remaining lifetime of a component), where at the same time, since RUDE+SC and Directed Graph Reconstruction algorithms provides interpretable rules, human operator can understand what stands behind a certain event.

The introduced architecture expresses our very general approach for extending the Digital Twin by AI thus achieving dynamic and adaptive models. The intelligent algorithms differentiate three cases and improve on the models according to them. In this way, the approach uncovers new contextual situations or new events, assigns human-interpretable rules to them and learns their effect. In order to evaluate this concept we developed a simulation to our showcase (Section 2.1). This simulation is described in the following section (Section 4).

4. SIMULATION TO EVALUATE THE APPROACH

The simulation consists on an environment simulation (ES) and the Digital Twin under Test (DTuT) of the truck. The ES is realized in a Unity3D simulation (Fig. 2). The ES provides a map with four maintenance bases and two road junctions. In the map, the ES models physical aspects such as gradients of

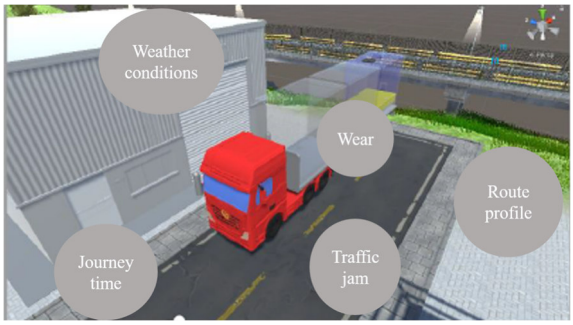

Fig. 2. Visualization of the simulation.

road segments. Moreover, the ES provides the event *traffic jam* and the event *rain*. The user configures both events in their contextual and statistical properties. For example, *traffic jam* occurs on road segment 3, kilometer 125 to 175 in the time range from 17:30 h to 18:30 h in 60% of the simulated journeys. In this simulated environment, the DTuT gets virtual sensor values like velocity and temperature as well as distance to target and (simulated) time. The goal is to reconstruct all the simulated physical parameters, as well as events and their effect. According to common sense, the ES uses operating hours as basic wear metric. The challenge for the DTuT is to guarantee the maintenance intervals by estimating the travel time to the next maintenance base and, at the same time estimate the wear.

The simulation is subdivided in two phases. In the first phase, the DTuT derives a rough state model representing the road network. In this process, the algorithms neglect events. In the second phase, when the rough state model is converged, i.e. the change of the spread remains $< \varepsilon$, the event models are activated to further improve the prediction accuracy. For better readability, in the presented data, events are deactivated in the first phase as well. However, simulating events from the beginning does not show fundamental changes.

At the beginning, the truck only knows the maintenance bases and the paths connecting them. Properties of the streets and relevant road junctions are unknown. The truck drives to the maintenance bases via different routes, on which it supplies the customers. On its journey, the truck collects process data. The truck or more precisely the DTuT evaluates this process data and learns from it as descripted in Section 3. It first recognizes repeating route segments using the Directed Graph Reconstruction. In this way, the DTuT builds a road graph that combines road sections with the same properties using the directed graph reconstruction algorithm by (Berkolaiko et al., 2018). Fig. 3 visualizes this procedure on the example of a new

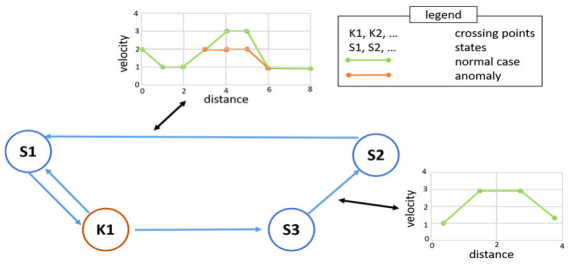

Fig. 3. Graph Reconstruction of the road segments.

discovered junction. The algorithm correlates curves of measured data from different journeys, recognizes recurring road sections and saves them as edges of the graph. Each edge represents a road segment. The context analysis algorithm assigns a context model to each of these edges. This context model links to several physical models. Physical models may apply for different contexts where the context models parameterize them. They are available in different degrees of maturity. In the simplest case the physical model breaks down to curve fitting if no further information is known at design time.

In our simulation, we take the temperature as an example. The temperature does not directly affect the operational time but it influences wear. Temperature models apply to several road segments. However, entering new climatic zone (e.g. truck drives from northern Europe to southern Europe) requires different parameterization. Fig. 4 visualizes the interplay of context model and physical model in case of a new instantiation. Since the process data deviate significantly more than the normal spread over the whole scope, a new model is instantiated. The model is just a floating average of the recorded data mapping temperature to time. However, there are different curves for the two different climatic zones. The context models maps the process data to the appropriate climatic zone and thus generate two temperature expectations depending on the road segment. Subsequently, during the next

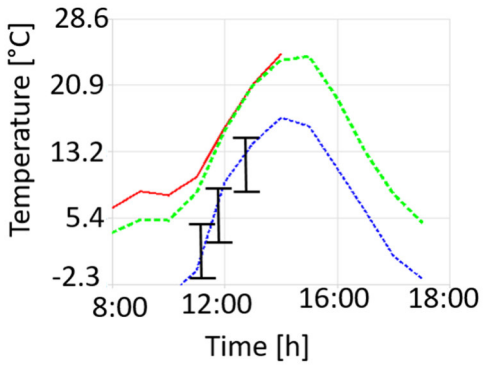

Fig. 4. Temperature curve indicating new context: The deviation from the measurement (red) to the known model (blue) significantly exceeds normal spread (black), so a new model is instantiated and trained (green).

runs, the temperature new instantiated model adapts to the new environmental conditions optimizing quadratic error.

Another physical model is the gradient model. This model is bound to a specific road segment, since every road segment is unique in this property. Figure 5 depicts the gradient model.

As visualized for *road segment 3* in Fig. 5 a) the trained physical model accurately represents the specified street properties. Now, the before event *traffic jam* comes into play. Fig. 5 b) visualizes its effect. Now the anomaly detection triggers the analyzing step that ends up in case (3), short but intense deviations with random characteristic. Unlike the completely different route, the event is characterized by the fact that the anomaly is only effective for a very limited spatio-temporal area. Therefore, anomaly detection triggers event

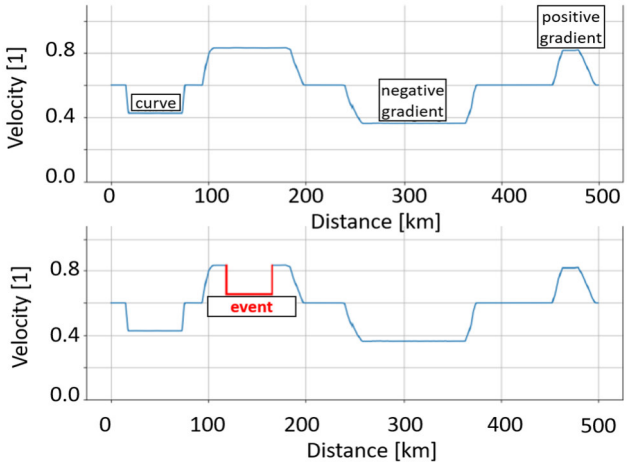

Fig. 5. Detailed view of edge3's velocity with event (b) and without event (a).

diagnosis, which compares the event with the rules from other events uncovered by RUDE+SC. Since the event occurs for the first time, no matching model does exist. Thus, the algorithm adds a new event model. Having instantiated a new event, any similar occurrence triggers rule identification for this event. Fig. 6 visualizes the RUDE+SC-Algorithm identifying the event's rule R.

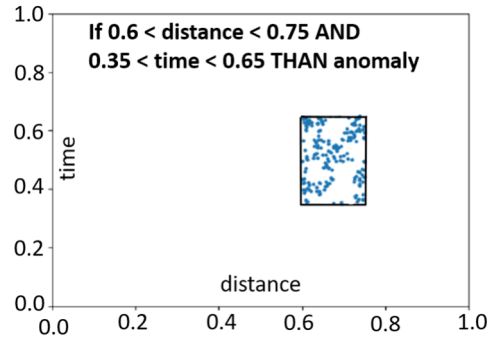

Fig. 6. Rule identification using RUDE+SC algorithm in normed values. The rule refers to traffic jam at rush hour.

In parallel to the rule analysis, the Event Training module performs the impact analysis exploiting neural networks for regression. The heat-map of the training is visualized in Fig. 7.

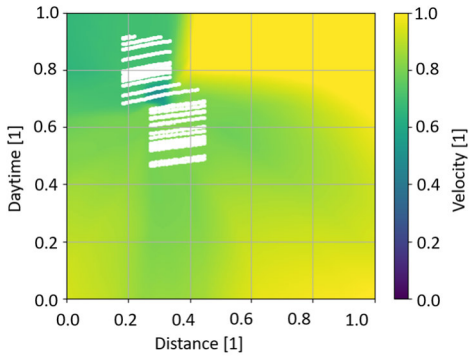

Fig. 7. Impact analysis of the identified events using neural networks. Anomalies are marked with white dots.

In the simulated case, the algorithms identify clear rules and patterns creating good estimation for the specific events, namely the pattern in the process data hinting to the nearby event, the rule when the event occurs and the impact.

Combining all the algorithms as described in the scenario, the simulated results are promising. Fig. 8 shows the development of the perception accuracy.

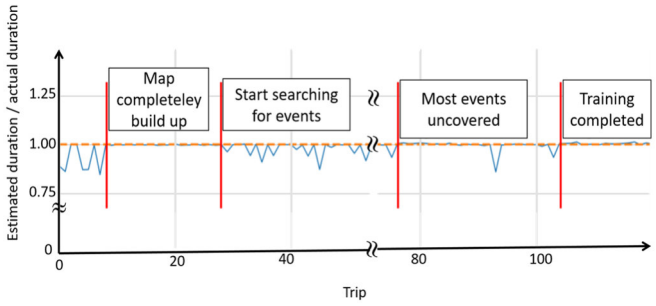

Fig. 8. Percentage of accuracy: accuracy of prediction increases with every new discovered element. When completely trained prediction error is lower than 1%.

In the first phase, where the DTuT has no prior information, the deviation from the estimation is quite high. Therefore, the safety margin has to be calculated high as well. After a few journeys, the rough state model converges and the simulation environment activates the events. In the next phase, the events linked to missing information show off. As the events are unknown to the truck, the impact on the estimation of the journey time is quite high. However, with increasing training data, the DTuT knows the events and thus the prediction improves. In the third phase, most predictions are accurate besides some outliers. This outliers refer to *(5) covariance shift*, where few data is known. Having uncovered these seldom events, the DTuT has completely established the event model and the estimation error drops to lower than 1%.

5. SUMMARY

In this paper, we presented an approach for dynamic and adaptive models hosted by the digital twin, which adapt to their environment not only by optimizing their parameters, but also by uncovering entirely new contextual situations or events. Subsequently, the models learn the characteristics of these situations or events, and their impact on other models and goals, specifically the optimal maintenance interval. Automatically derived and human-interpretable rules provide insights into the decision-making process. We showcased the approach on the example of safe maintenance planning of a highly-automated truck and evaluated it in a simulation. The approach shows promising results and is generally valid. It is transferable to all domains of industrial automation.

6. REFERENCES

Allal AA, Mansouri K, Qbadou M, Youssfi M. Task human reliability analysis for a safe operation of autonomous ship. In: 2nd International Conference 2017. p. 74–81.

Ashtari Talkhestani B, Jung T, Lindemann B, Sahlab N, Jazdi N, Schloegl W, et al. An architecture of an Intelligent Digital Twin in a Cyber-Physical Production System. at - Automatisierungstechnik 2019; 67(9): 762–82.

Bajcsy A, Bansal S, Bronstein E, Tolani V, Tomlin CJ. An Efficient Reachability-Based Framework for Provably Safe Autonomous Navigation in Unknown Environments, 2019.

Berkolaiko G, Duffield N, Ettehad M, Manousakis K. Graph Reconstruction from Path Correlation Data, 2018. 25 p.

Chen J, Yuan B, Tomizuka M. Deep Imitation Learning for Autonomous Driving in Generic Urban Scenarios with Enhanced Safety, 2019. 7 p.

Di Franco C, Bezzo N. Interpretable Run-Time Monitoring and Replanning for Safe Autonomous Systems Operations. IEEE Robotics and Automation Letters 2020; 5(2): 2427–34.

Halpern J. The Sequential Covering Problem Under Uncertainty. INFOR: Information Systems and Operational Research 1977; 15(1): 76–93.

Kephart JO, Chess DM. The vision of autonomic computing. Computer 2003; 36(1): 41–50.

Lindemann B, Fesenmayr F, Jazdi N, Weyrich M. Anomaly detection in discrete manufacturing using self-learning approaches. Procedia CIRP 2019; 79: 313–8.

Löcklin A, Müller M, Jung T, Jazdi N, White D, Weyrich M. Digital Twin for Verification and Validation of Industrial Automation Systems – a Survey. In: 25th IEEE International Conference 2020. p. 851–858.

Lud M-C, Widmer G. Relative Unsupervised Discretization for Association Rule Mining. In: Zighed DA, Komorowski J, Żytkow J, editors. Principles of data mining and knowledge discovery. Lecture notes in computer science. 1910 : Lecture notes in artificial intelligence. Berlin, Heidelberg: Springer; 2002. p. 148–158.

Luo Y, Yu Y, Jin Z, Zhao H. Environment-Centric Safety Requirements for Autonomous Unmanned Systems. In: 2019 IEEE 27th International Requirements Engineering Conference (RE); 2019. p. 410–415.

Machin M, Guiochet J, Waeselynck H, Blanquart J-P, Roy M, Masson L. SMOF: A Safety Monitoring Framework for Autonomous Systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2018; 48(5): 702–15.

Magdici S, Althoff M. Fail-safe motion planning of autonomous vehicles. In: IEEE 19th International Conference 2016. p. 452–458.

Müller J, Gabb M, Buchholz M. A Subjective-Logic-based Reliability Estimation Mechanism for Cooperative Information with Application to IV's Safety, 2019. 7 p.

West TD, Blackburn M. Demonstrated benefits of a nascent Digital Twin. INSIGHT 2018; 21(1): 43–7.