Preprints, 10th IFAC International Symposium on Advances in Automotive Control
Columbus, Ohio, USA, August 28-30, 2022

Automatic track guidance of industrial trucks with time-variant vehicle parameters using AI-based controllers

Timm Sauer * **Luca Spielmann** * **Manuel Gorks** *
Klaus Zindler * **Ulrich Jumar** **

* *University of Applied Sciences, Aschaffenburg, Germany (e-mail: Timmanton.Sauer@th-ab.de, Luca.Spielmann@th-ab.de, Manuel.Gorks@th-ab.de, Klaus.Zindler@th-ab.de)*
** *ifak - Institute of Automation and Communication, Magdeburg, Germany (e-mail: Ulrich.Jumar@ifak.eu)*

Abstract:
This paper presents an extension of a self-learning control concept for automatic track guidance of industrial trucks in intralogistic systems. The presented approach is based on Reinforcement Learning (RL), a method of Artificial Intelligence (AI) and is able to adapt itself to different industrial truck variants and the associated specific vehicle parameters. Moreover, time-variant parameters during operation, such as the vehicle's velocity are taken into account. In order to consider the existing a priori knowledge of the controlled system and to avoid starting the whole training process of the controller for each truck variant from scratch, the training process is divided into two steps. In the first step, the controller is trained on a model using parameters of a nominal vehicle variant. Based on this, the control parameters are only fine-tuned in the second step. In this way the controller is adapted to the actual truck variant and the corresponding parameter values. In order to take into account the time-variant vehicle parameters during operation, the Artificial Neural Networks (ANN) of the RL controller and the observation vector are suitably extended. In this way, the varying speed can be considered in both training steps and the control parameters can be optimized accordingly. Thus, in case of the investigated scenarios a stable control loop behavior can be guaranteed for the entire speed range of industrial trucks. In order to demonstrate this, the new approach is compared with a RL control concept, not considering time-variant parameters.

Keywords: Artificial intelligence, Automatic control, Intelligent transportation systems

1. INTRODUCTION

1.1 Problem description and requirements

In times of global economic markets and increasing competition, the automation of logistic processes is a basic requirement for corporate success. An important object of research and development is to increase the internal material flow via an autonomous and intelligent networked fleet, that usually consists of a wide variety of different individual truck variants.
In order to achieve this objective, the track guidance of the industrial trucks has to be implemented autonomously. The classical control design is based on a mathematical model that describes the dynamics of the controlled system as accurately as possible. In case of a heterogeneous fleet, a suitable model has to be derived for each vehicle variant. Based on this model, the control design has to be carried out for every single truck variant, which proves to be time-consuming.
Furthermore, time-variant parameters, such as the vehicle speed, cannot be considered using classical control concepts. However, actually the vehicle speed can vary extremely. For example, the industrial truck creeps with low speed during pick and place operations, but can reach maximum values of up to 5.5 m/s during transfer operations (see Linde Material Handling [2022]).
Consequently, a control concept for automatic track guidance of industrial trucks has to be developed that independently adapts to different industrial truck variants and moreover considers time-variant parameters during operation.

1.2 Related research

The papers Li et al. [2020], Tamba et al. [2008] and Mohammadi et al. [2016] deal with the automatic track guidance of industrial trucks, but each of them is focusing only on a single truck variant. The classical methods of robust control (Ackermann [1993] and Zindler [1994]) take into account unknown but during operation constant model parameters. The consideration of time-variant model parameters is not possible with these approaches.
In addition to the classical methods of adaptive control (Landau et al. [2013] and Aström et al. [1995]), the concepts based on AI are becoming increasingly important. An overview as well as a classification of the different AI approaches is given in Sauer et al. [2021].

The well-known RL-control methods suffer from the fact, that a priori knowledge concerning the dynamic plant behavior is not integrated in the training process (Havenstrom et al. [2020], Sallab et al. [2016]). Therefore, a new approach has been presented in Sauer et al. [2021]. It's basic idea consists of integrating a priori knowledge of the controlled system into the training process. For this purpose, the training is divided into two steps. In the first step the controller is pre-trained on basis of a nominal model representing a priori knowledge of lateral dynamic vehicle behavior. Since this model is derived for an industrial truck with average vehicle parameter values, in the second step a fine tuning of the control parameters is performed in order to adapt to the actual vehicle variant. In this way the efficiency of the whole training process is significantly increased.
However, this approach does not consider time-variant vehicle parameters. Therefore, the simulation results are restricted to a constant vehicle velocity value. Since this parameter actually has a high influence on the dynamic behavior of the vehicle (section 3), this dependency should be taken into account in the design of the controller.

1.3 Main contribution and outline of this paper

This paper presents an extension of a self-learning control concept for automatic track guidance of industrial trucks which is based on RL. It adapts itself to different vehicle variants and also takes into account time-variant model parameters during operation. RL is implemented in form of the so-called Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm, as it proves to be suitable for the application of automatic track guidance (Fujimoto et al. [2018]). The method of integrating a priori plant knowledge into the training process presented in Sauer et al. [2021] is extended to the consideration of time-variant vehicle parameters, in this application the highly variable vehicle velocity.
By means of an appropriate extension of the so-called observation vector (section 4) the information of the current vehicle velocity is integrated into the RL control concept. Furthermore, the structures of the RL controller's ANN, have to be adjusted in order to process the information of the observation vector. To take into account the variable speed in the training process of the RL controller, the training is divided into several epochs. During an epoch, the vehicle speed is remaining at a constant value but is changing between the individual epochs. By varying the speed within the training, the parameters of the controller can be adjusted for the entire speed range of the different truck variants and a stable control loop behavior can be guaranteed. To demonstrate this, the control concept proposed in this paper is compared with the RL control concept given in Sauer et al. [2021].
This paper is organized as follows. Section 2 introduces the principle of automatic steering control and the corresponding control structure. In section 3 the modeling of the plant and analysis of the system will be described for different vehicle velocities. The fundamentals of RL, the used TD3 algorithm and the control approaches will be introduced in section 4. Subsequently, the simulation results of both control concepts are assessed (section 5). At the end of the paper, in section 6, the main conclusions are discussed.

2. SYSTEM OVERVIEW

Figure 1 demonstrates the principle of automatic steering control of an industrial truck. First of all, the desired vehicle trajectory (predefined path) is calculated and stored as a data set. The record includes the necessary setpoint information for automated vehicle guidance, such as the Cartesian Coordinates and curvature of the trajectory. The objective of automatic track guidance consists of eliminating the lateral deviation of the vehicle with respect to the path. In Sauer et al. [2021] it is shown that compensating the lateral deviation a_p in a preview point P_p, is resulting in an improved controllability of the system. For this purpose, P_p is defined in the preview distance l_p in front of the industrial truck's center of gravity (CoG) (Tan et al. [1999]). Based on the measured position of the vehicle's CoG, the Cartesian Coordinates of the preview point P_p can be calculated. In order to determine the deviation a_p, it is necessary to calculate a reference point R_p on the predefined path by means of the algorithm given in Zindler et al. [2012]. The lateral deviation a_p finally corresponds to the distance between the preview point P_p and the reference point R_p.

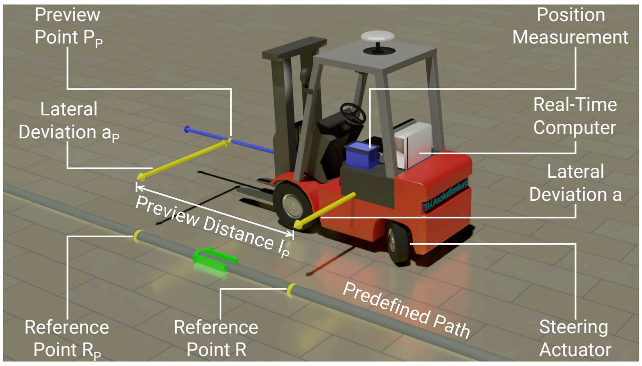


Fig. 1. Principle of the lateral vehicle guidance system

The structure of the proposed vehicle guidance system is provided in figure 2. The plant model (section 3) consists of three parts starting with the position controlled steering actuator. The second part is the so-called single track model, that describes the lateral vehicle dynamics (side slip angle β and yaw rate $\dot{\psi}$) depending on the steering angle δ_r. The last part represents the kinematics of the vehicle, i.e. its relative motion with respect to the predefined path. The curvature χ_p of the path in the reference point R_p represents one input of the controlled system and is considered as disturbance variable. The second input is a control signal δ_{set} which is calculated by the lateral controller in dependence of the lateral deviation a_p (output signal of the controlled system). The lateral controller, shown in figure 2, is implemented using a new method of RL, which is described in detail in section 4.

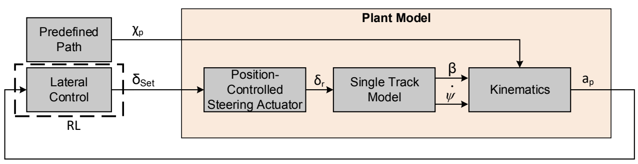


Fig. 2. Structure of the vehicle guidance system

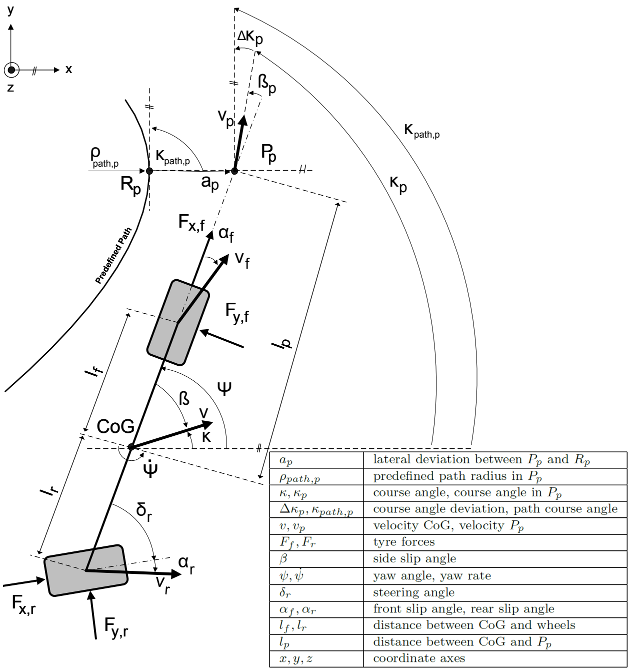
3. PLANT MODEL

3.1 Modeling of the plant

As shown in section 2, the mathematical model consists of three parts. The first part describes the dynamics of the steering actuator and is implemented as a first order delay element with the delay time T_s. Subsequently, the single-track model (Pacejka [2006] and Zindler et al. [2012]) is used and adapted to describe the lateral dynamic of forklifts with rear axle steering (figure 3). It is based on the following simplifications and assumptions:

- Reduction to one wheel per axle
- Neglect of longitudinal dynamic forces like traction forces, braking forces and aerodynamic drag forces
- Small steering angles, slip angles and side slip angles

In contrast to the assumption known from literature, in this paper, the velocity of the vehicle is not assumed as constant or slowly changing. Rather, it is considered as time-variant model parameter $v(t)$.



a_p	lateral deviation between P_p and R_p
$\rho_{path,p}$	predefined path radius in P_p
κ, κ_p	course angle, course angle in P_p
$\Delta\kappa_p, \kappa_{path,p}$	course angle deviation, path course angle
v, v_p	velocity CoG, velocity P_p
F_f, F_r	tyre forces
β	side slip angle
$\psi, \dot{\psi}$	yaw angle, yaw rate
δ_r	steering angle
α_f, α_r	front slip angle, rear slip angle
l_f, l_r	distance between CoG and wheels
l_p	distance between CoG and P_p
x, y, z	coordinate axes

Fig. 3. Single track model with rear axle steering

The third part of the model represents the kinematics of the vehicle. It describes the relative motion of the industrial truck with respect to the predefined path (Söhnitz [2001]). As shown in Sauer et al. [2021], the preview concept is used in order to eliminate the non-minimum phase system behavior, caused by the rear axle steering of industrial trucks. The model can be given in state space representation (equation 1), where $x(t) = \left[\beta(t), \dot{\psi}(t), \Delta\kappa(t), a_p(t), \delta_r(t)\right]^T$ describes the state vector of the system and $u(t) = [\delta_{set}(t), \chi_p(t)]^T$ represents the vector of its input signals. These are the steering angle set-point, calculated by the lateral controller (control signal), as well as the curvature of the predefined path, considered as disturbance variable. The plant model is derived in detail in Sauer et al. [2021], where a validation using real measurement data is described as well.

$$
\begin{bmatrix} \dot{\beta}(t) \\ \ddot{\psi}(t) \\ \Delta\dot{\kappa}(t) \\ \dot{a}_p(t) \\ \dot{\delta}_r(t) \end{bmatrix} = \begin{bmatrix} -\frac{c_f + c_r}{m \cdot v(t)} & \frac{-c_r \cdot l_r + c_f \cdot l_f}{m \cdot v(t)^2} + 1 & 0 & 0 & \frac{c_r}{m \cdot v(t)} \\ \frac{-c_r \cdot l_r + c_f \cdot l_f}{J_z} & -\frac{c_r \cdot l_r^2 + c_f \cdot l_f^2}{J_z \cdot v(t)} & 0 & 0 & \frac{c_r \cdot l_r}{J_z} \\ -\frac{c_f + c_r}{m \cdot v(t)} & \frac{c_f \cdot l_f - c_r \cdot l_r}{m \cdot v(t)^2} & 0 & 0 & \frac{c_r}{m \cdot v(t)} \\ 0 & -l_p & v(t) & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_s} \end{bmatrix}
\cdot \begin{bmatrix} \beta(t) \\ \dot{\psi}(t) \\ \Delta\kappa(t) \\ a_p(t) \\ \delta_r(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & v(t) \\ 0 & 0 \\ \frac{1}{T_s} & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta_{set}(t) \\ \chi_p(t) \end{bmatrix} \tag{1}
$$

3.2 Analysis of the plant model

In real applications, the speed varies frequently and is depending on the current operation mode. For example, in pick and place operations, the industrial truck has to be guided precisely and accordingly very slowly, while in transfer operations higher velocities are achieved. The vehicle speed has a high influence on the dynamics of the system. Figure 4 demonstrates the pole-zero diagram of the industrial truck variant Linde E30. Its model parameters are given in table 1. The diagram shows the location of the poles (x) and zeros (o) for the entire speed range from $v = 1\text{m/s}$ to $v = 5\text{m/s}$.

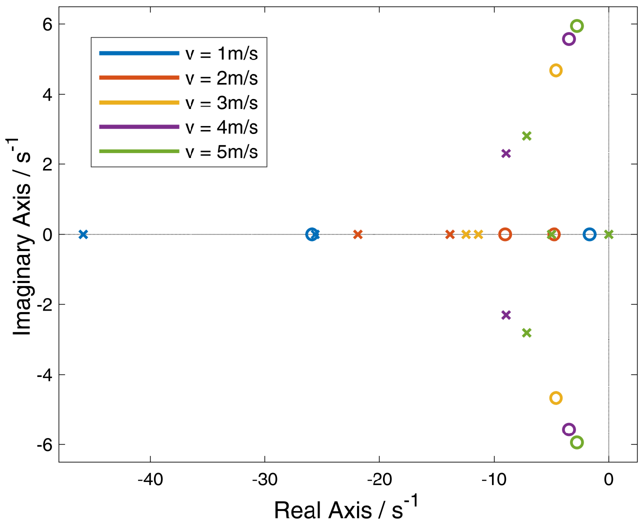


Fig. 4. Pole-Zero-Map for different vehicle speed values

Obviously, the position of the poles strongly varies in dependence of the speed values. For vehicle velocities $v \geq 4\text{m/s}$, there is a conjugated complex pole pair. For increasing velocities this pole pair shifts towards the imaginary axis, indicating a decreasing damping of the system. The high variability of the dynamic system behavior points out that the velocity has to be considered as a time-variant parameter in the control concept.

4. CONTROL METHODS

4.1 Reinforcement Learning basics

RL in the domain of control systems is a well-known approach (Vogt [2018], Lillicrap et al. [2016] and Sutton et al.

[2018]). Due to the analogy to the human learning process, the self learning characteristics of RL offers potential for solving complex control problems.

Table 1. Vehicle parameters (Linde Material Handling [2022])

	Description	Linde E30	Linde E80
m	vehicle mass	4981 kg	15720 kg
l	wheelbase	1.665 m	2.400 m
c_f	cornering stiffness	62000 N/rad	62000 N/rad
c_r	cornering stiffness	122000 N/rad	122000 N/rad
l_f	axial distance to CoG	0.858 m	1.181 m
l_r	axial distance to CoG	0.807 m	1.219 m
J_z	moment of inertia	3624 kgm^2	26490 kgm^2
T_s	delay time constant	0.2 sec	0.2 sec

The principle of closed-loop operation process of RL is displayed in figure 5. It essentially consists of three blocks. The lowest block (vehicle) represents the controlled system, in this case the industrial truck. Its current state Φ_k is provided to the RL controller. This block describes the lateral controller that calculates the control signal u_k in order to affect the controlled system. The third block (reward function) evaluates the control signal u_k based on the current state Φ_k and the following state Φ_{k+1}, in form of a feedback, called reward r_k. It is a measure of control quality. In analogy to the human learning process, the control strategy is adapted in order to optimize the reward.

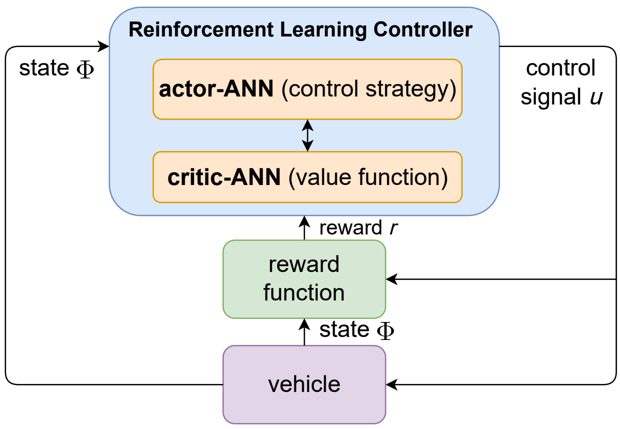


Fig. 5. Principle of Reinforcement Learning

The implementation of the described basic idea of RL can be done by different methods. In this paper the TD3 algorithm is used, which is an extension of the Deep Deterministic Policy Gradient algorithm, given in Lillicrap et al. [2016]. It is a so-called Actor-Critic method that uses separate memory structures to differ between the control strategy $\mu(\Phi)$ (actor) and the value function $Q(\Phi, u)$ (critic). $Q(\Phi, u)$ is a function to calculate the expected reward $\hat{r}$, based on its input signals Φ and u. Both, the critic and the actor, are implemented in form of an ANN. The optimization of the parameters ϕ of the critic-ANN is done by supervised learning, based on the obtained reward (Gurney [1997], Hagan et al. [2014]). The task of the actor-ANN consists of calculating the control signal u_k in dependence of the current system state Φ_k and is indicated as a function of the actor parameters θ. The optimization of the parameters θ of the actor-ANN should be done in order to maximize the output of the critic-ANN and thus the reward. To implement this, a criterion J which is equal to $Q(\Phi, u)$ is maximized using a gradient method given in Lillicrap et al. [2016]. The resulting update function (equation 2) is calculated by applying the chain rule to the expected reward $Q(\Phi, u)$ with respect to the actor-ANN parameters θ:

$$\nabla_\theta J \approx \frac{1}{N} \sum_i^N \nabla_u Q(\Phi, u|\phi)|_{\Phi=\Phi_i, u=\mu(\Phi_i)} \nabla_\theta \mu(\Phi|\theta)|_{\Phi=\Phi_i} \tag{2}$$

The observation vector Φ, reflecting the state of the system, is depending on the chosen methodology. Whether the time-variant parameter of the vehicle speed is taken into account or not, the observation vector is composed differently (subsection 4.2 and subsection 4.3).

4.2 Approach without consideration of the vehicle velocity

In this subsection, the RL control concept without consideration of time-variant parameters is introduced. This approach has been published in Sauer et al. [2021] and is restricted to the assumption of a constant vehicle speed. Therefore, the used observation vector Φ (equation 3) is formed identical to the state vector x of the model, described in section 3:

$$\Phi = x = [\beta, \dot{\psi}, \Delta\kappa, a_p, \delta_r]^T \tag{3}$$

The behavior of the RL controller can be specified by the definition of the reward function. Sauer et al. [2021] demonstrated that closed-loop behavior of optimal state control can be approximated by choosing the reward function r_k in analogy to the quadratic cost function of classical LQR (Ichikawa et al. [1992]). In this application the reward function is defined to focus on minimizing the lateral deviation a_p of the vehicle with respect to the path. Therefore, the weighting factor of $a_{p,k}^2$ is chosen significantly lager than the weightings of the other signals.

$$r_k = -(\beta_k^2 + \dot{\psi}_k^2 + \Delta\kappa_k^2 + 10000 \cdot a_{p,k}^2 + \delta_{r,k}^2 + 5 \cdot \delta_{set,k}^2) \tag{4}$$

4.3 Approach with consideration of the vehicle velocity

In order to take into account the high influence of the vehicle speed on the dynamics of the controlled system (section 3) the observation vector Φ (equation 3) is extended by the vehicle velocity signal, leading to Φ_{ext} (equation 5).

$$\Phi_{ext} = \begin{bmatrix} \Phi \\ v \end{bmatrix} = [\beta, \dot{\psi}, \Delta\kappa, a_p, \delta_r, v]^T \tag{5}$$

Since the signals of the observation vector form the inputs of the actor-ANN and the critic-ANN of the RL controller, the structure of these networks has to be adjusted by extending the input layers of the ANN. A further neuron is integrated in the ANN of the RL controller, in order to process the information of the enlarged observation vector. Figure 6 depicts a simplified representation of the structure of the actor-ANN (left) and the critic-ANN (right). In the first hidden layer of both fully connected feed-forward ANN, 400 neurons are inserted. Therefore, the extension of the input layer with an additional neuron results in a large number of ANN parameters.

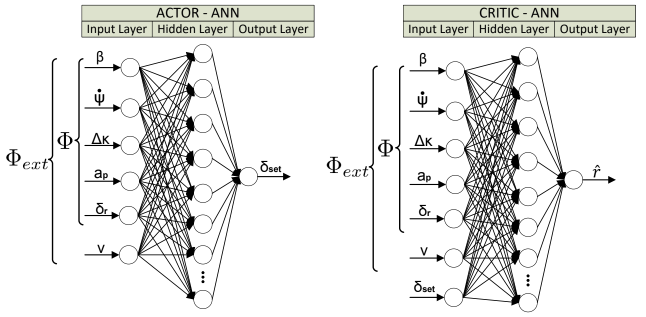


Fig. 6. Simplified representation of the extended ANN structure of the RL controller

In order to compare the different RL control concepts with each other, the reward function given in equation 4 is used for this approach as well.

5. CONTROL DESIGN AND SIMULATION RESULTS

In this section, the simulation results of both RL concepts are presented and evaluated. Subsection 5.1 focuses on the results after the first training step (pre-training). This first training step is performed using the model parameters of a nominal industrial truck variant (Linde E30).
Subsequently, the adaptability of the RL concepts to another vehicle variant, such as the Linde E80 will be discussed. For this purpose, the second training step (fine tuning) is performed based on the pre-trained controller (subsection 5.2). Both training steps are carried out in simulation using the model described in section 3, considering the following scenario. The vehicle starts with an initial lateral deviation of the preview point of $a_p = 0.2$m, i.e. offset from the path. The curvature of the path is applied as a disturbance variable illustrated in figure 7. The path initially runs as a straight line [0sec - 5sec] and then merges into a curve with a constant radius ρ_{path} [6sec - 10sec]. The transition between the mentioned segments is realized as a clothoid, where the curvature is linearly increased until it reaches the final value [5sec - 6sec].

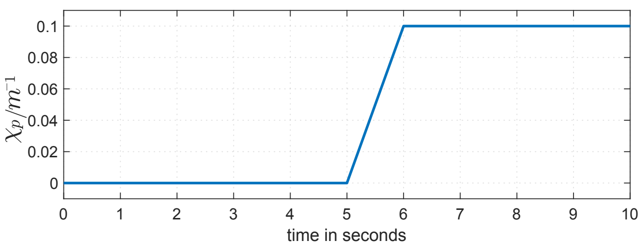


Fig. 7. Path curvature in the reference point R_p (used in subsection 5.1 and subsection 5.2)

The maximum path curvature value is defined to $\chi_p = 0.1m^{-1}$ which corresponds to a curve radius of 10 meters. This scenario is defined in order to test both control concepts within the entire speed range of industrial trucks. Thus, it can be ensured that both industrial truck variants can automatically be guided through the scenario even at a high vehicle velocity of $v = 4$m/s.
However, in order to evaluate the closed loop behavior even in scenarios with higher path curvatures, an additional evaluation using a modified scenario is given in subsection 5.3. The initial lateral deviation of the preview point P_p with respect to the path is set to $a_p = 0.3$m and the curvature is linearly increased [5sec - 6sec] to a final value of $\chi_p = 0.2m^{-1}$ [6sec - 10sec] (figure 8).

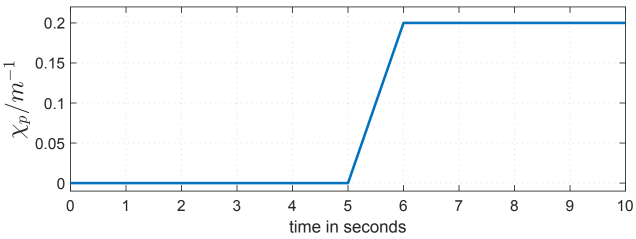


Fig. 8. Path curvature in the reference point R_p (used in subsection 5.3)

5.1 Simulation results after the first training step

This subsection compares the two RL control approaches after completion of the first training step. The simulation results of the nominal vehicle variant (Linde E30) are presented. It is to be shown, how the different control concepts can deal with a time-variant vehicle speed during operation. Since the concept given in Sauer et al. [2021] is restricted to the assumption of a constant vehicle velocity, the RL controller is trained on a medium speed of $v = 2$m/s, which remains constant in all training epochs.
For the concept considering time-variant model parameters during operation, the structure of the ANN of the RL controller is adjusted, as discussed in subsection 4.3. The training process of the RL controller is divided into several epochs, each of them with a different vehicle velocity within the possible speed range [1m/s - 5.5m/s] of industrial trucks.

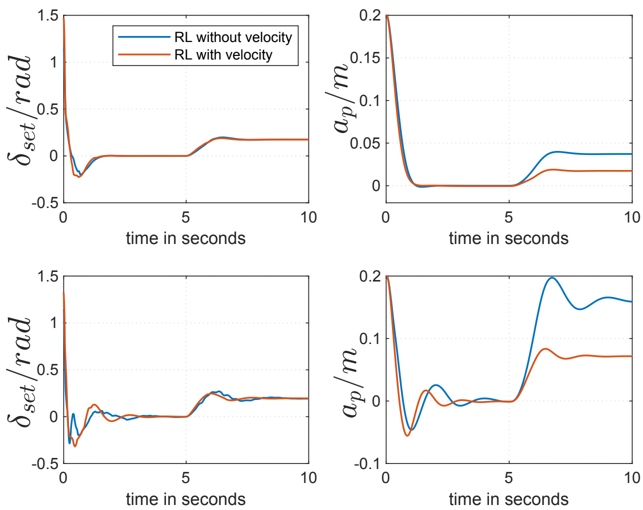


Fig. 9. Steering angle and lateral deviation (Linde E30) for $v = 2$m/s (top) and $v = 4$m/s (bottom)

Figure 9 shows the simulation results of both control concepts. In the upper part of figure, the time courses of the control variable (δ_{set}) and the controlled variable (a_p) at the vehicle speed $v = 2$m/s are depicted. Obviously, the simulation results of both control concepts are comparable. When the speed is increased to $v = 4$m/s, the approach proposed in this paper can show its advantages (lower part of figure 9). This is due to the fact, that the whole speed range of industrial trucks is considered within the training process and the control parameters can be optimized accordingly. Obviously, neither control scheme ensures steady-state accuracy in the curve sector of the maneuver [6sec - 10sec]. The extension of the input layer in the ANN of the RL controller in combination with the high number of neurons of the hidden layer, leads to a more complex ANN with a large number of additional ANN parameters. This results in a higher degree of freedom in regard to the design and already improves the control quality for a speed of $v = 2$m/s. However, it has a negative effect on the training efficiency, since 3.6 times as many optimization steps have to be performed (table 2).

Table 2. Training efficiency

Control Concept	Training step	Optimization steps
No consideration of v	1st step E30	116822
No consideration of v	2nd step E80	14000
Consideration of v	1st step E30	420558
Consideration of v	2nd step E80	66459

5.2 Investigation of the controller's adaption

In this subsection, the adaption of both control concepts to a new industrial truck variant is investigated. To avoid starting the entire training process for another vehicle variant from scratch the pre-trained RL controllers of subsection 5.1 are used. Both controllers have to be adapted within the second training step (fine tuning) to the actual vehicle variant, in this case the Linde E80 and the associated parameters (table 1). By this method, the number of optimization steps can be significantly reduced compared to a training that has to be started from scratch (table 2).

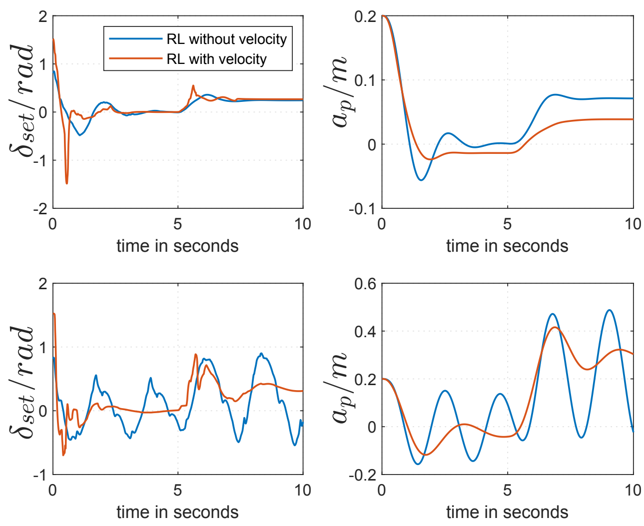


Fig. 10. Steering angle and lateral deviation (Linde E80) for $v = 2$m/s (top) and $v = 4$m/s (bottom)

The re-trained controllers are evaluated in terms of their performance for different vehicle velocities in figure 10. At a speed of $v = 2$m/s, both controllers are able to compensate the initial deviation and guarantee a stable control loop behavior. The increase of the vehicle velocity value to $v = 4$m/s results in an unstable closed-loop behavior applying the retrained controller of Sauer et al. [2021]. Evidently, the new approach ensures a significantly higher control quality and guarantees a stabilization of the industrial truck within the entire speed range.

5.3 Validation of the test results on a modified scenario

Finally, both control concepts are tested on a modified scenario with a constant vehicle velocity of $v = 2$m/s. For this purpose, the presented control approaches will be evaluated both after the first training step on the nominal truck variant (subsection 5.1) and after the adaption to the larger vehicle variant (subsection 5.2).
In contrast to the previous simulation tests, the initial lateral deviation, i.e. the lateral offset of the vehicle with respect to path is significantly increased. At the beginning of the scenario, the controllers have to compensate a lateral deviation of $a_p = 0.3$m, which results in a high control signal (figure 11). The controller considering time-variant parameters during operation, calculates even the maximum value of the control signal ($\delta_{set} = 1.57$rad).
The influence of the tighter curve of this scenario can also be seen in the figure. The resulting lateral deviation during the cornering operation [5sec - 10sec] increases compared to the scenario evaluated in subsection 5.1 and subsection 5.2. Nevertheless, both control concepts are capable of automatically guiding both the Linde E30 and the Linde E80 even through the modified simulation test scenario.

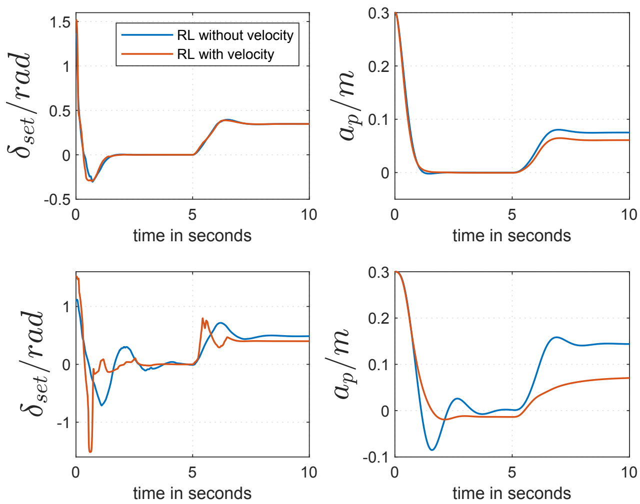


Fig. 11. Steering angle and lateral deviation of Linde E30 (top) and Linde E80 (bottom) for $v = 2$m/s

6. CONCLUSION AND FUTURE WORK

This paper presents a new control concept for the automatic track guidance of industrial trucks based on AI. By separating the training process into two steps, existing a priori knowledge regarding the controlled system can be integrated. In the first training step, the RL controller's

experience is built up using a linear model and the parameters of a nominal, average vehicle variant. Since the fundamental lateral dynamic behavior is comparable for all vehicle variants, this experience does not have to be performed again from scratch. Therefore, based on the pretrained controller, an adaptation to other vehicle variants can be performed by fine-tuning the control parameters in a second training step.
A central result of this paper is the extension of the method in order to consider varying vehicle velocities in the design procedure. This is realized by adjusting the observation vector and the ANN of the RL controller. Thus, a stable behavior of the closed loop can be ensured for an entire speed range of different industrial truck variants.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Bavarian Ministry of Economic Affairs, Regional Development and Energy for the financial support of the project *Cooperative Autonomous Intralogistic Systems* under Grant No. DIK-1910-0016 and the industrial project partner: Linde Material Handling GmbH.

REFERENCES

L. Li et al. Path-Following Control for Self-driving Forklifts based on Cascade Disturbance Rejection with Coordinates Reconstruction. *39th CCC*, 2020.
T. Tamba et al. Trajectory Generation of an Unmanned Forklift for Autonomous Operation in Material Handling System *SICE Annual Conference*, 2008.
A. Mohammadi et al. Model Predictive Motion Control of Autonomous Forklift Vehicles with Dynamics Balance Constraint *14th ICCARV*, 2016.
J. Ackermann Robuste Regelung: Analyse und Entwurf von linearen Regelungssystemen mit unsicheren physikalischen Parametern *Springer-Verlag, Berlin, Heidelberg, 1. Auflage*, 1993.
K. Zindler Nutzung der Freiheitsgrade für den Entwurf robust stabiler Regelungen bei strukturierter und undstrukturierter Modellunsicherheit *Shaker Verlag, Aachen*, 1994.
I. Landau et al. Adaptive Control: Algorithms, Analysis and Applications *Springer, London*, 2013.
K.J.Aström et al. Adaptive Control *Mass: Addison-Wesley, Reading, 2. Edition*, 1995.
T. Sauer et al. Automatic track guidance of industrial trucks using self-learning controllers considering a priori plant knowledge *5th ICCAD*, 2021.
A. Sallab et al. End-to-End Deep Reinforcement Learning for Lane Keeping Assist *30th Conference on NIPS*, 2016.
H. Tan et al. Developement of an Automated Steering Vehicle Based on Roadway Magnets - A case Study of Mechatronic System Design *IEEE/ASME TOM*, 1999.
H. Pacejka Tyre and vehicle dynamics *Oxford: Butterworth-Heinemann, 2. Edition*, 2006.
I. Söhnitz Querregelung eines autonomen Strassenfahrzeugs *Fortschr.-Ber. VDI Verlag Düsseldorf*, 2001.
K. Zindler et al. Querdynamische Fahrzeugführung zur reproduzierbaren Erprobung von Sicherheitssystemen *at-Automatisierungstechnik, Oldenbourg-Verlag*, 2012.
Linde Material Handling Homepage Linde Material Handling *Accessed on 03.01.2022*.
M. Vogt An overview of deep learning techniques *at - Automatisierungstechnik, Oldenbourg-Verlag*, 2018.
T. Lillicrap et al. Continuous Control with Deep Reinforcement Learning *ICLR*, 2016.
R. Sutton et al. Reinforcement Learning: an introduction. *The MIT Press*, 2018.
S. Havenstrom et al. Proportional integral derivative controller assisted reinforcement learning for path following by autonomous underwater vehicles. *arXiv preprint arXiv:2002.01022*, 2020.
S. Fujimoto et al. Addressing Function Approximation Error in Actor-Critic Methods *ICML, Stockholm*, 2018.
K. Gurney An introduction to neural networks *UCL Press Limited*, 1997.
M. Hagan et al. Neural Network Design *Martin Hagan*, 2014.
Y. Ichikawa et al. Neural Network Application for Direct Feedback Controllers *IEEE Transactions on neural networks, VOL. 3, NO. 2*, 1992.