UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Model-free Intelligent Control for Antilock Braking Systems on Rough Terrain

By:

Ricardo de Abreu

Submitted in partial fulfilment of the requirements for the degree

Master of Engineering

(Mechanical Engineering)

in the Faculty of

Engineering, Built Environment and Information Technology (EBIT)

University of Pretoria

June 2022

Executive Summary

Title: Model-free Intelligent Control for Antilock Braking Systems on Rough Terrain

Author: Ricardo de Abreu

Study Leaders: Dr. T.R. Botha, Dr. H.A. Hamersma

Department: Mechanical and Aeronautical Engineering, University of Pretoria

Degree: Masters in Engineering

Advancements made in Advanced Driver Assistance Systems such as the Antilock Braking System (ABS), have made profound improvements in the handling and safety of vehicles. An ABS enhances the braking performance and steerability of a vehicle under severe braking conditions by preventing wheel lockup. However, its performance degrades on rough terrain resulting in an increased wheel lockup and stopping distance. This is largely due to noisy measurements, the type of ABS control algorithm used, and the excitation of complex dynamics such as higher order tyre mode shapes that are neglected in the control strategy.

Model-free reinforcement learning (RL) control techniques do not have modelling constraints, thus enabling them to overcome the pitfalls of un-modelled dynamics and parametric uncertainties that come with braking over rough terrain. It is for this reason that they are proposed for the ABS control algorithm. Initial tests using a simplified quarter car model are completed to determine the aspects of the controller such as the appropriate RL algorithm, reward function, and function approximator. The results show that the Double Deep Q-learning Network (DDQN) algorithm with the Temporal Convolutional Network (TCN) is the best performing algorithm, with a pressure-based reward function that also penalizes high longitudinal slip of the wheel.

Using a validated full-vehicle simulation model and the FTire tyre model in ADAMS, the proposed control algorithm is trained over a rough Belgian paving, with stochasticity introduced in the measurements during training to improve its generality and robustness. Straight line braking, braking-in-turn, and braking on split-mu surface simulations are performed from an initial braking speed of 55km/h; comparisons are then drawn against a Bosch algorithm tuned for off-road braking as well as a conventional braking system (no ABS).

The DDQN-TCN algorithm is able to prevent wheel lockup without significantly deteriorating the vehicle's stopping distance in all simulations over rough terrain. Additional performances over the high friction flat surface and split-mu surface highlight the generality of the algorithm. The robustness of the algorithm to noisy data over rough terrain is underlined by its ability to prevent wheel lockup without the use of a filter. A limitation of the algorithm occurs over the low friction flat surface in which the algorithm is unable to cycle the wheel speed fast enough resulting in wheel lockup. Two possible solutions are proposed: 1) training the model over a low friction surface and/or multiple surfaces, and 2) improving the braking efficiency of the model through a modern ABS modulator with faster response times. These two solutions should allow the algorithm to generalise over multiple surfaces, and produce improved response times over low friction surfaces respectively.

Additional recommendations include experimentally testing the control algorithm over the Belgian paving at the Gerotek testing facilities, making use of an alternative method to overcome the catastrophic forgetting that the algorithm suffers during training. Furthermore, investigating other performance affecting parameters such as tyre carcass oscillation, and revisiting the reward function in which parameters such as the yaw rate and path following error can be explored, possibly improving the performance of the algorithm.

Acknowledgements

I would like to extend my appreciation to the following people:

- Dr. Theunis Botha, for his mentorship, and guidance throughout the last 4 years. Thank you for encouraging me to pursue my interests and placing faith in me to complete this study.
- My parents, Teresa and Jose de Abreu, for their love and support throughout my studies, and for being the epitome of hard work paying off.
- Dr. Herman Hamersma, for his kind assistance, and guidance throughout this research.
- My VDG colleagues, for their advice, criticism, and ability to make research interesting and fun. To mention a few: Matthew, Jesse, Andries, Wiestche, and Prof Els.
- Dakota Saudan, for proof reading, grammar refinement, and her unwavering love and support throughout.

Contents

List of Figures

List of Tables

List of Symbols

Abbreviations

Acronym	**Description**
A2C	Advantage Actor-Critic
A3C	Asynchronous Advantage Actor-Critic
ABS	Antilock Braking Systems
ADP	Adaptive Dynamic Programming
AI	Artificial Intelligence
CNN	Convolutional Neural Networks
CG	Center of Gravity
DQN	Deep Q-Learning Networks
DDQN	Double Deep Q-Learning Network
DOF	Degree of Freedom
ECU	Electronic Control Unit
EWC	Elastic Weight Consolidation
FC	Fuzzy Control
FCN	Fully Convolutional Network
FLC	Fuzzy-Logic Controllers
FTire	Flexible Ring Tire
i.i.d	independent and identically distributed
LSTM	Long Short Term Memory

MEX ..... MATLAB executable

ML ..... Machine Learning

MDP ..... Markov Decision Process

NHTSA ..... National Highway Traffic Safety Administration

NN ..... Neural Networks

NNFBL ..... Neural Network-based Feedback Linearization

ODE ..... Ordinary Differential Equation

PID ..... Proportional Integral Derivative

RL ..... Reinforcement Learning

RNN ..... Recurrent Neural Network

RTMC ..... Road Traffic Management Corporation

SAE ..... Society of Automotive Engineers

SMC ..... Sliding-Mode Control

SAC ..... Soft Actor-Critic

SUVs ..... Sports Utility Vehicles

TCN ..... Temporal Convolutional Network

TCP ..... Transmission Control Protocol

TD ..... Temporal Difference

UDP ..... User Datagram Protocol

VDG ..... Vehicle Dynamics Group

WGN ..... White Gaussian Noise

WHO ..... World Health Organization

Alpha-numeric symbols

Symbol Description ..... **Unit**

---

a Acceleration of a body ..... [m/s^2]

a An action ..... [-]

A Set of all actions available in state s ..... [-]

A_t Action at time t ..... [-]

d Receptive field dilation factor ..... [-]

F_n Normal force ..... [N]

F,F_x Longitudinal force ..... [N]

F_y Lateral force ..... [N]

I Mass moment of inertia ..... [kgs^2]

$J(\boldsymbol{\theta})$ Performance measure for the policy $\pi_{\boldsymbol{\theta})}$ ..... [-]

k Receptive field filter size ..... [-]

m Mass of vehicle ..... [kg]

M_z Moment about z-axis ..... [N m]

P Pressure ..... [MPa]

$\dot{P}$ Instantaneous rate of Pressure ..... [MPa]

P_{set} Backpressure ..... [MPa]

$P(s'|s,a)$ Probability of transition to state s', from state s taking action a ..... [-]

$q_\pi(s,a)$ Value of taking action a in state s under policy π ..... [-]

$q_*(s,a)$ Value of taking action a in state s under the optimal policy ..... [-]

Q Array estimates of action-value function q_π or q_* ..... [-]

r A reward ..... [-]

r_{eff} Effective rolling radius ..... [m]

$R(s,a)$ Reward function ..... [-]

s, s' States ..... [-]

S Set of all nonterminal states ..... [-]

S_t State at time t, typically due, stochastically, to S_{t-1} and A_{t-1} ..... [-]

T_b Brake Torque ..... [N m]

$T(s'|s,a)$ Transition function ..... [-]

v Vehicle velocity ..... [m/s]

$\dot{v}$ Vehicle acceleration ..... [m/s^2]

v_r Relative velocity ..... [m/s]

v_s Stribeck velocity ..... [m/s]

$v_\pi(s)$ Value of state s under policy π (expected return) ..... [-]

$v_*(s)$ Value of state s under the optimal policy ..... [-]

V Array estimates of state-value function v_π or v_* ..... [-]

$y(t_k)$ Unfiltered input signal ..... [-]

$y(t_k)_{filtered}$ Filtered output signal ..... [-]

z Bristle deflection ..... [m]

$\dot{z}$ Instantaneous rate of bristle deflection ..... [m/s]

Greek symbols

Symbol Description ..... Unit

α Angular Acceleration ..... [rad/s^2]

α Slip angle ..... [°]

$\boldsymbol{\alpha}$	Stribeck exponent	[-]
α	Learning rate	[-]
β	Smoothing factor	[-]
ϵ	Exploration parameter	[-]
γ	Discount factor	[-]
λ	Longitudinal wheel slip	[-]
μ_c	Kinematic friction coefficient	[-]
μ_p	Peak coefficient of friction	[-]
μ_s	Static friction coefficient	[-]
π	Policy (decision making rule)	[-]
$\pi(s)$	Action taken in state s under deterministic policy π	[-]
$\pi(a\|s)$	probability of taking action a in state s under stochastic policy π	[-]
σ_0	Rubber longitudinal lumped stiffness	[1/m]
σ_1	Rubber longitudinal lumped damping	[s/m]
σ_2	Viscous relative damping	[s/m]
σ^2	Variance	[-]
$\boldsymbol{\theta}, \boldsymbol{\theta}_t$	Parameter vector of target policy	[-]
τ	Time delay constant	[s]
ω	Angular velocity	[rad/s]
$\dot{\omega}$	Rotational Speed of tyre	[rad/s^2]
ζ	Contact Position	[m]

Chapter 1: Introduction

1.1 Background

Road traffic injuries are a leading cause of death globally, with the World Health Organization (WHO) stating that more than 1.2 million people die on the world's roads per annum. It is expected that by the year 2030, vehicle crashes will become the fifth leading cause of fatalities worldwide - overtaking AIDS and tuberculosis (WHO, 2015). According to the Road Traffic Management Corporation (RTMC), an average of 23.5 people per 100,000 lose their lives on South African roads (RTMC, 2017), compared to the 2015 WHO global status report which puts the global average of road fatalities at 17.42 per 100,000 (WHO, 2015). Vehicle factors account for 14.1% of the road crashes in South Africa, with 11.4% of these factors due to faulty brakes (RTMC, 2017). The conditions of the road and environment present further challenges with poor and wet/slippery road conditions accounting for 8.3% and 19.4% of the fatal casualties respectively (RTMC, 2017).

Significant advances towards vehicle safety have been made with one such system being the development of Antilock Braking Systems (ABS). Originally introduced for trains in the early 1900s, the ABS was developed by Bosch for passenger vehicles and light trucks in 1936 and by 1957 Kelsey-Hayes initiated the "automatic" braking system exploratory development program, which concluded that the braking system should not only reduce the vehicle's stopping distance but also prevent the loss of vehicle control (SAE, 2014). Since then, the implementation of ABS on passenger and light commercial vehicles has become mandatory in most markets, with reports released by the National Highway Traffic Safety Administration (NHTSA) highlighting notable decreases in multi-vehicle crashes and fatal pedestrian crashes with the inclusion of ABS (Garrott and Mazzae, 1999).

However, the performance of an ABS on rough terrain conditions is found to be suboptimal, resulting in wheel lockup, bad stopping distances, and poor lateral control (Penny and Els, 2016; Hamersma, 2017; Van der Merwe et al., 2018). This has inspired the emerging field of study exploring the direct influence of several phenomena such as different ABS algorithms and suspension control, which directly influence the effectiveness of ABS on rough terrain. Braking over rough terrain presents a challenge to the performance of an ABS by introducing effects that

interfere with its normal operation, leading to subpar performances. These effects include noisy measurements, continually varying terrain excitations, vehicle body motions, and tyre carcass oscillations. As a tyre travels over a rough terrain, the undulations in the terrain surface not only change the effective rolling radius of the tyre (Hamersma, 2017), but also cause changes in the vertical and torsional dynamics of the wheel. These variations are often neglected by most braking systems, resulting in the tyre losing contact with the terrain and/or wheel lockup. Two aspects are often addressed to improve off-road braking. The first is to reduce the effect of the road excitation by controlling the suspension system - the majority of available research (Shao et al., 2007; Reul and Winner, 2009; Els et al., 2013; Hamersma and Els, 2014). The second explores improving the ABS algorithm to compensate for the vertical excitations and rotational tyre dynamics. Less research has been conducted on the latter, especially on rough terrain (Adcox et al., 2012; Penny and Els, 2016; Hamersma, 2017), and this study aims to narrow the gap in literature by concentrating on the ABS algorithm.

The majority of ABS controllers are simulated on smooth terrain (dry/wet/icy) with focus on improving the vehicles stopping distance. However, excessive wheel lockup occurs as a result of the additional excitation in off-road conditions, resulting in wheel lockup and loss of directional control.

Developing an ABS controller which prevents wheel-lockup on rough terrain, maintaining directional stability without considerably worsening the stopping distance, forms the main objective of this study. An initial broad problem statement is defined at first, with the intention of refining this statement post the completion of a literature review and investigation into the initial aspects of the proposed ABS control method. The final problem statement is then used throughout the remainder of the study.

1.2 Problem Statement

The development of an ABS control method that is capable of detecting and reducing high slip conditions on rough terrain in an attempt to prevent wheel lockup, while maintaining lateral control and a good stopping distance.

1.3 Overview of Study

This document is structured as follows:

- **Chapter 2: Literature Review**

  The relevant research carried out on the topic at hand is presented in this chapter. This includes tyre force generation, existing tyre models, an introduction to ABSs, and state-of-the-art ABS algorithms.

- **Chapter 3: Reinforcement Learning**

  An introduction into reinforcement learning and its different algorithms are explored. An overview of the study is provided.

- **Chapter 4: Single Tyre Simulation**

  Training and simulations over smooth and rough are performed to determine the initial aspects of the control algorithm using a simplified quarter car model. Comparisons are drawn against a Bosch algorithm tuned for off-road braking, and a conventional braking system.

- **Chapter 5: Full Vehicle Simulation**

  The final ABS algorithm is trained over a rough Belgian paving, and tests according to the SAE requirements are simulated on a validated full vehicle simulation model in ADAMS. A summary of the results concludes this chapter.

- **Chapter 6: Conclusion and Recommendations**

  Conclusion of this study, and future recommendations are detailed in Chapter 6.

Chapter 2: Literature Review

The tyre-road relationship influences the performance of an ABS and is explored through tyre force generation and tyre models. An introduction to ABS and the current state-of-the-art algorithms are provided. A refinement of the research question concludes the chapter.

2.1 Tyre Force Generation

To understand how an ABS works, it is imperative to understand the relationship between the tyre and the terrain. A vehicle's ability to accelerate, brake, and corner are dictated by the forces generated by the interaction between the tyre and the terrain (the tyre-terrain interface). The friction force generation in the contact patch area of a tyre (the area of the tyre that is in direct contact with the terrain) is a result of two mechanisms: adhesion and hysteresis (Gillespie, 1992). Adhesion arises from the intermolecular bonds between the tyre and the aggregate in the terrain, and is the greater of the two mechanisms on dry roads. However, it diminishes on wet surfaces, hence reduced friction on such surfaces (Gillespie, 1992). Hysteresis is generated as a result of energy loss that occurs during the deformation of a tyre over the terrain and does not diminish on wet surfaces. The generation of longitudinal and lateral tyre friction forces are discussed in the sections to follow.

2.1.1 Longitudinal Force Generation

The development of the friction force of a rolling wheel is achieved when there is relative movement between the tyre and the terrain. The longitudinal friction force is generated along the direction of travel of the wheel, and exists when the wheel is accelerated or braked. Slip that occurs in the case of longitudinal friction force generation is defined as the ratio between the contact patch sliding velocity (difference between the vehicle velocity and the tangential tyre velocity) and the vehicle's velocity. This slip ratio is presented in Equation (2.1)

$$\lambda = \frac{|v - \omega r_{eff}|}{v}, \tag{2.1}$$

where λ is the longitudinal wheel slip, v is the velocity of the vehicle, ω is the wheel angular velocity, and r_{eff} is the effective rolling radius of the tyre. The longitudinal friction coefficient is typically

characterised as a function of longitudinal wheel slip as depicted in Figure 2.1. The peak friction coefficient (μ_p) generally occurs within the region of 8-10% wheel slip and indicates the maximum braking force that can be obtained from the particular tyre-terrain friction pair (Gillespie, 1992). Once the longitudinal slip exceeds this peak, the braking coefficient decreases and wheel lockup occurs almost immediately (Hamersma, 2017).

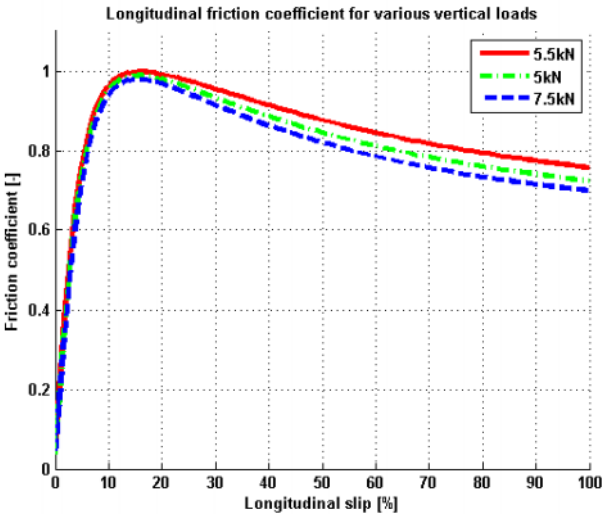


Figure 2.1: Longitudinal friction coefficient versus longitudinal wheel slip (Hamersma and Els, 2014)

The development of the longitudinal braking force is affected by factors such as the vertical load, where a decrease in the vertical load results in an immediate decrease in the longitudinal force. Conversely, an increase in the vertical load results in an increased contact patch area but not an instantaneous increase in the longitudinal friction force. This is due to the run-in effect (or relaxation length), which is when a section of the tyre contact patch remains non-deformed, requiring for the tyre to continue rolling before an increase in the longitudinal friction force is generated (Breuer and Bill, 2008). Thus, over rough terrain where the vertical force changes, there is a nett loss of traction.

2.1.2 Lateral Force Generation

Lateral friction forces are an important aspect when considering the stability and directional control of a vehicle, and essential in achieving one of the aims of an ABS - maintaining directional control of the vehicle without significantly impacting its stopping distance. When steering a vehicle while cornering, lateral forces are generated perpendicular to the direction of travel of the wheel. The tyre contact patch is deflected away from the direction of heading of the wheel, and in the direction of travel of the vehicle (see Figure 2.2). The difference in the angle is referred to as the slip angle (α). Lateral forces have two main sources of generation: the lateral slip of the tyre (sideslip angle),

and the lateral inclination of the tyre (camber angle), with the camber angle being prominent in motorcycles (Gillespie, 1992). The relationship between the longitudinal and lateral friction force is explored in the next section.

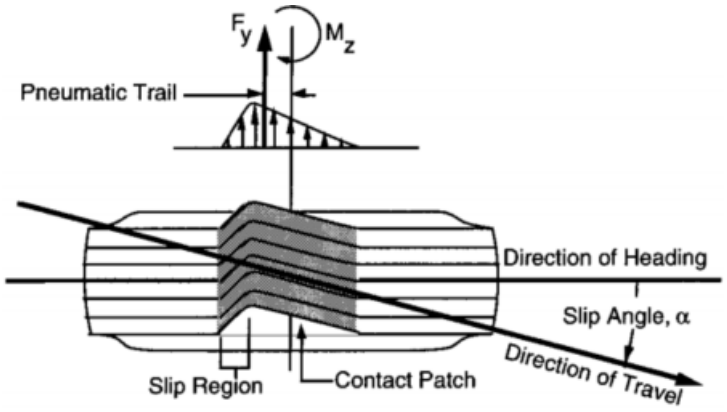


Figure 2.2: Deformation of a tyre under lateral force (Gillespie, 1992)

2.1.3 Combined Lateral and Longitudinal Force Generation

When a tyre undergoes braking and steering at the same time, longitudinal and lateral forces are generated simultaneously. However, these force characteristics share an "inverse" relationship such that as the tyre generates longitudinal force, the maximum lateral force that can be generated diminishes and vice versa. The longitudinal forces take preference over the lateral forces such that if lateral forces are present on a tyre when hard braking is applied, these lateral forces become zero, whilst the longitudinal forces remain present (Gillespie, 1992). Thus, a vehicle that utilises all traction for braking will lose any available traction for steering and therefore lose directional control.

A more accurate depiction of the relationship is the friction circle (see Figure 2.3). Owing to the fact that the friction force generated by a tyre in any direction is equal to its friction coefficient multiplied with the normal load, the dotted line represents the maximum force that can be developed in any direction. This follows the assumption that the contact patch is symmetric and the tyre is isotropic. Point C illustrates the point at which the tyre is generating maximum lateral force at the provided slip angle with zero longitudinal force, while point A represents the maximum longitudinal (braking) force without lateral force. Point D illustrates a combined case where the tyre generates as much steering as braking force and is at its frictional limit.

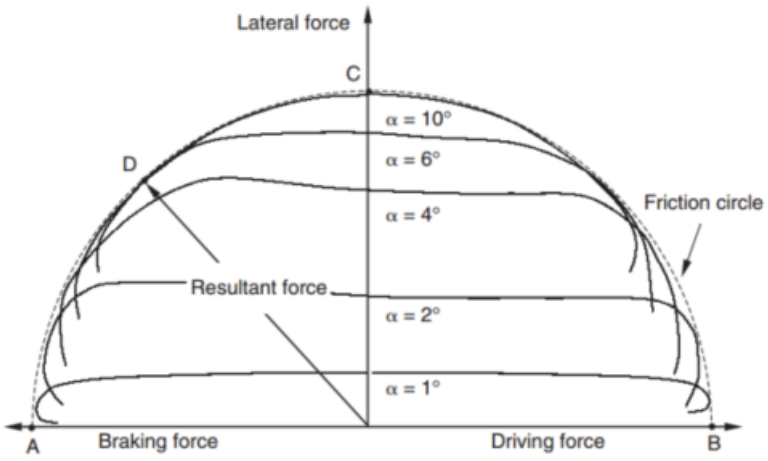


Figure 2.3: Combined lateral and longitudinal friction force (Blundell and Harty, 2004)

2.1.4 Tyre Modelling

Primary control of a vehicle is achieved through the interactions between the tyre and the terrain, and an accurate tyre model is therefore necessary. Tyre models can be broken up into two main categories: empirical models and physics-based models (Pacejka, 2005).

Empirical models are supported by experimental data and are popular for their simplicity and low computational cost. Examples include the semi empirical Magic Formula tyre model developed by Bakker et al. (1989). While these models have the ability to approximate experimental data (Pacejka, 2005), their coefficients have no engineering significance in terms of physical tyre properties, and are unable to accurately capture the transient response of a tyre (Stallmann and Els, 2014). Physics-based models are developed through first principles, allowing for a deeper understanding of the complex behaviour between the tyre and the terrain. These models also allow for easy tuning of model-parameters. Dynamic tyre models are of interest in control studies that involve the analysis of the longitudinal transient tyre dynamics under severe braking scenarios, with one such model being the LuGre tyre model.

2.1.4.1 LuGre Tyre Model

The LuGre tyre model is a brush-type tyre model that is based on the contact-point LuGre friction model developed by Canudas de Wit et al. (1995). By being able to accurately capture the transient behaviour of the friction force observed during transitions between braking and acceleration, the LuGre tyre model has been used successfully to simulate ABS control studies over smooth terrain (Liang et al., 2008; Adcox et al., 2013; Yu et al., 2015).

The biggest advantage of the LuGre tyre model is its representation as a smooth Ordinary Differential Equation (ODE) involving no discontinuities. However in doing so, it neglects the nonsmooth nature of real friction, failing to reproduce the exact static friction of the tyre (Kikuuwe et al., 2006). The LuGre tyre model consists of two types: lumped and distributed. The lumped model describes the state of friction between the tyre and terrain through the friction state (deflection) parameter $z(t)$ and brake tyre force F, and is characterised by normalised lumped stiffness and damping, and viscous damping coefficients: σ_0, σ_1, σ_2. The distributed model assumes the existence of an area at the tyre contact patch and represents the deflection of the bristles as a function of the contact point ζ along the patch, at time t i.e. $z(\zeta, t)$ (Canudas-de Wit et al., 2003).

The vertical forces generated in the tyre are of great significance, and while no basic tyre model allows for rough terrain, the LuGre model is useful for simple simulations where the vertical excitations can be modelled by changing the vertical force on the tyre. Alternatively, a complex and more accurate method such as the multibody FTire model is considered as it allows for the capturing of vertical excitations over rough terrain.

2.1.4.2 Flexible Ring Tire Model (FTire)

The Flexible Ring Tire (FTire) model is an advanced, physics-based, 3-dimensional nonlinear vibration model developed by Gipser (1999). This model represents the tyre belt as a flexible ring modelled with bending stiffness while elastically mounted on the rim. It consists of radial, tangential, and lateral stiffnesses, and makes use of a number of belt elements to approximate the generated tyre forces (Stallmann and Els, 2014). Literature shows that the FTire model can accurately simulate and capture vertical excitations over rough terrain while accurately simulating the friction and transient response of the tyre (Stallmann and Els, 2014).

2.2 Antilock Braking Systems (ABS)

According to the Society of Automotive Engineers (SAE), the application of an ABS to a vehicle is to provide improvements in the vehicle's performance under braking compared to a conventional brake system, with primary focus on the stability, steerability, and stopping distance of the vehicle (SAE, 2014). ABS is a feedback control system that attempts to maintain controlled braking under all operating conditions.

2.2.1 ABS Operation

The basic ABS hardware consists of an Electronic Control Unit (ECU), a solenoid valve for releasing and reapplying pressure to a brake circuit (two valves per wheel), a shuttle valve, a pump, and a wheel speed sensor. A simplified layout of an ABS setup is provided in Figure 2.4, where the solenoid and shuttle valves and pump are located in the modulator. The ECU measures the wheel speed and vehicle speed through sensors (e.g. proximity sensors), and determines the slip and wheel angular deceleration (Gillespie, 1992). Should these parameters exceed a predefined threshold, the system determines that wheel lockup is imminent and control is applied to prevent this.

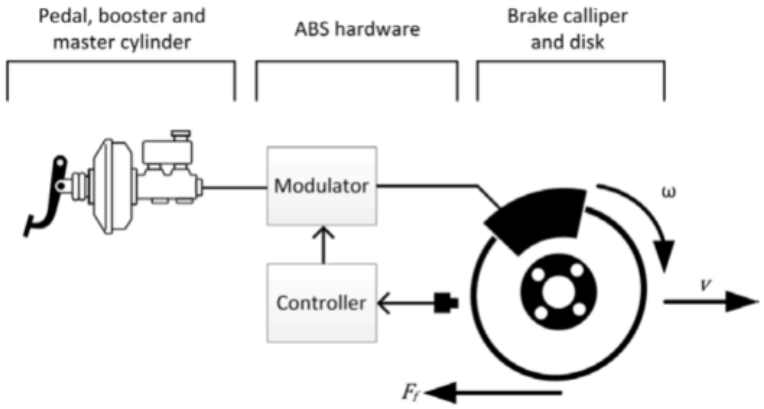


Figure 2.4: Simplified layout of ABS hardware (Penny, 2016)

When wheel lockup occurs the result is reduced braking and no lateral control. It is the aim of an ABS to prevent this. This involves manipulating the wheel slip such that maximum friction coefficient between the tyre and terrain is achieved, whilst maintaining the steerability (lateral stability) of the vehicle. This is accomplished by actuating the solenoid valves in order to increase, reduce, or hold the brake pressure through a control strategy with three braking states: *pump, dump* and *hold*. The state of the brake defines the type of actuation that is to be done to the brake pressure at each wheel, and includes increasing the pressure (*pump*), reducing the pressure (*dump*), or maintaining a constant brake pressure (*hold*) (Penny, 2016).

Figure 2.5 shows the application of the different brake states. The controller attempts to maintain the longitudinal slip within the cycling region (dark blue line) as this allows for the optimum braking distance whilst providing some lateral force generation for steering. Once the longitudinal slip exceeds the pre-defined threshold (state 3), the wheel angular acceleration becomes large and the wheel stops rotating within a few milliseconds, resulting in wheel lockup almost instantaneously. This is prevented by the *dump* state which reduces the brake pressure and

enables the wheel to spin again, allowing for the generation of lateral forces to steer the vehicle while braking. Similarly, once the longitudinal slip decreases below a pre-defined threshold (state 1), the brake pressure is gradually reapplied (*pump* state) and maintained in the cycling process (state 2). This process is repeated until the brake is released by the driver or the vehicle speed reduces below a set value.

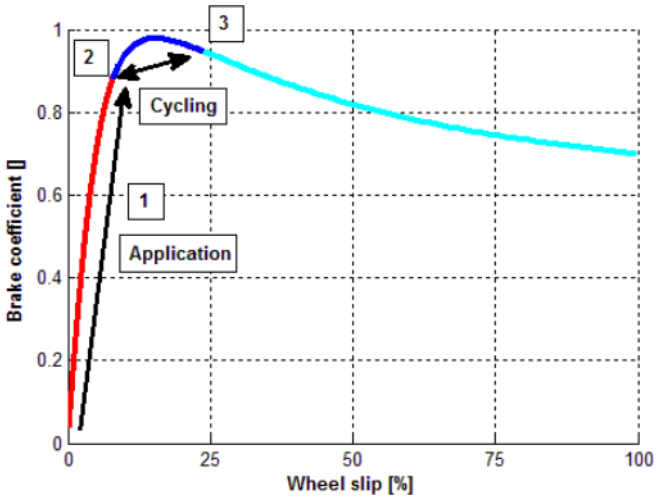


Figure 2.5: ABS cycling procedure (Hamersma and Els, 2014)

2.2.2 ABS Performance on Rough Terrains

As the popularity of Sports Utility Vehicles (SUVs) has increased in recent years, so has the requirements and operating conditions of ABS in various terrain conditions. The study of an ABS operation on rough terrain is a relatively new area of research which is yet to be fully understood; literature reveals that its performance deteriorates on such conditions (Penny and Els, 2016; Hamersma, 2017; Van der Merwe et al., 2018). To better understand why this is, the evaluation method and metrics that are used to measure the performance of an ABS need to be understood, along with the different parameters that affect an ABS when operating on rough terrain.

2.2.2.1 Performance Evaluation of ABS

Guidelines for test procedures to evaluate the performance of an ABS are outlined by the SAE (SAE, 2014) and include:

- Straight line braking on high and low friction surfaces
- Stability and controllability in response to steering inputs
- Braking in a turn
- Split friction coefficient braking

The performance evaluation of an ABS is usually based off the stopping distance of the vehicle, although a more accurate evaluation of its effectiveness would be to consider a combination of the stopping distance and steerability of the vehicle during braking. As a result of the conflicting nature of an ABS requirements, no single metric or evaluation method exists that is capable of comparing one controller to another. Hamersma (2017) proposed an ABS performance evaluation technique which considers four metrics that can be use to better evaluate the performance of a vehicle under braking:

1. Stopping distance and its standard deviation (if more than one test is carried out for the same algorithm)
2. Mean and standard deviation of average vehicle deceleration
3. Mean and standard deviation of absolute vehicle yaw rate error
4. Mean and standard deviation of absolute lateral path following error

Metrics 1 and 2, and 3 and 4, are similar and in competition, such that the controller with the best stopping will not have steering capabilities or vice versa. Using an illustration similar to a radar plot as seen in Figure 2.6, the proposed technique is suitable for all surface types, and provides a clear and logical method to measure the performance of ABS algorithms. The best performing algorithm occupies the outermost corners of the plot (Hamersma, 2017).

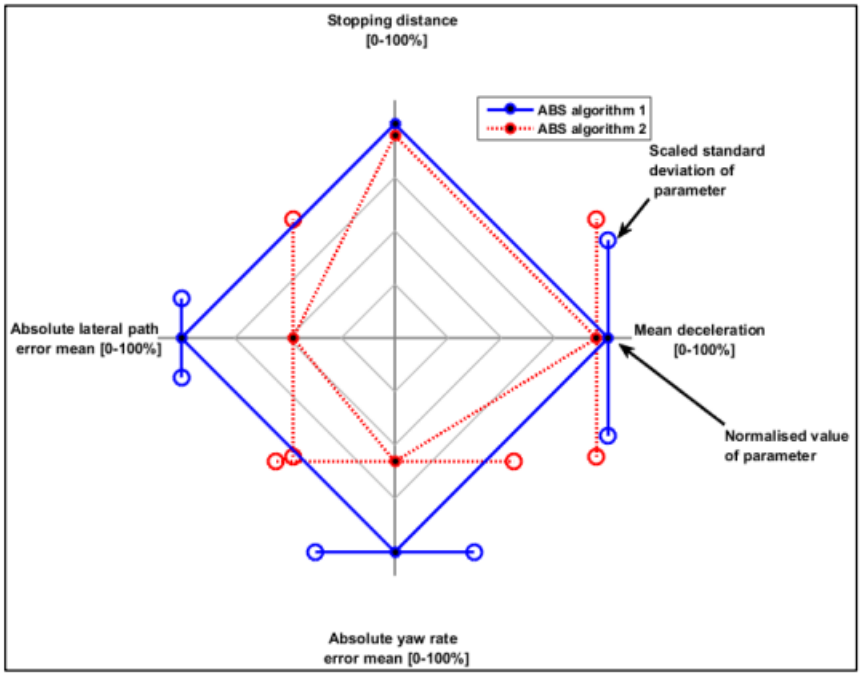


Figure 2.6: ABS evaluation technique (Hamersma, 2017)

To better understand the factors that influence the performance of an ABS on rough terrain as opposed to normal road applications, the effect of the terrain, suspension control, and ABS control algorithms are investigated.

2.2.2.2 Effect of Terrain on ABS

Braking over rough terrain presents a tough challenge to the performance of an ABS by introducing effects that interfere with its normal operation, leading to subpar performances. These effects include noisy measurements, inconsistent terrain input excitations, vehicle body motions, and tyre carcass oscillations (Hamersma, 2017).

As a tyre travels over rough terrain, the undulations in the terrain surface not only change the effective rolling radius of the tyre as proven in tests conducted by Hamersma (2017) - see Figure 2.7 - but also cause changes in the vertical and torsional dynamics of the wheel which often result in loss of contact between the tyre and terrain. Another effect is the tyre carcass may begin to oscillate, causing significant errors in the computation of the tyre's longitudinal slip (Adcox et al., 2013). The undulations result in variations in the normal (vertical) load applied on the tyre, with these variations negatively impacting the run-in effect of the tyre force generation as discussed in Section 2.1.1. The run-in effect occurs as a result of the delayed response of tyre to a step input (Gillespie, 1992), with this delay taking one-half to one full wheel revolution to reach the desired steady state force conditions. An additional consequence of tyre undulations is the generation of vehicle body motions such as roll, pitch, and yaw motions that cause large variations in wheel vertical loads (Hamersma and Els, 2014).

Figure 2.7: Tyre deformation as it rolls over a cleat (Hamersma, 2017)

Traditionally, slip estimation procedures are determined through the wheel rotation speed, vehicle speed, and tyre rolling radius parameters. These parameters are directly affected by the terrain fluctuations and become noisy - often requiring filtering for cleaner signals (Van der Merwe et al., 2018). The slip estimation procedure is further complicated due to phase lag attributed to the transient behaviour of the longitudinal forces in the wheel, causing a delay between the actual longitudinal slip at the tyre terrain contact and the slip measured at the wheel hub. As a result, Botha and Els (2015) proposed a method of determining the tyre slip-ratio directly through digital image correlation techniques at the tyre-terrain interface. This approach avoids the uncertainties caused by measuring inaccuracies of vehicle speed and tyre roll radius due to terrain roughness.

2.2.2.3 Effect of Suspension Control on ABS

The use of suspension control to improve the braking performance over rough terrain has been a significant focus in literature (Shao et al., 2007; Reul and Winner, 2009; Els et al., 2013; Hamersma and Els, 2014). Els et al. (2013) demonstrated that combining a semi-active suspension control with an ABS over rough terrain can result in stopping distance differences of up to 14m. Using a 2013 year model Land Rover Defender 110 and a semi-active suspension system, Hamersma and Els (2014) found that the spring and damper characteristics directly influence the performance of ABS on rough roads, with the best braking results achieved with a lower stiffness at the rear suspension, and lower stiffness at the front suspension yielding the worst results. Moderate to high damping at the front suspension also proved to be advantageous.

As an alternative approach to improving off-road braking, the ABS algorithm can be improved to compensate for the vertical excitations and rotational tyre dynamics, and is explored further.

2.2.3 ABS Control Algorithm

The type of control approach used for controlling an ABS plays a crucial role in its performance and forms the main focus of this study. As a result of the complex nature between the tyre friction and slip, the control of an ABS is a highly nonlinear problem. This control is further impeded by noisy wheel speeds and vehicle parameters that are generally not directly measurable (vehicle speed and longitudinal slip) but estimated, often leading to controller malfunction (Blundell and Harty, 2004). Several different control approaches exist as summarised in Figure 2.8. Few control strategies have been simulated over rough terrain; a summary of the more relevant methods and the terrain simulated on are explored.

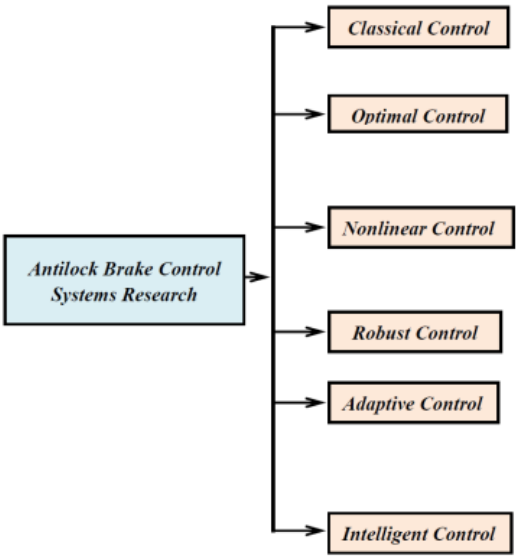


Figure 2.8: Summary of different ABS control methods (Aly et al., 2011)

2.2.3.1 Bang-bang Control

Bang-Bang control is a type of control system that mechanically or electronically turns something on or off when a desired target/threshold has been reached. This approach is one of the earliest and simplest strategies used to modulate the brake pressure, with almost all ABS setups that make use of the three different brake states (*pump*, *dump*, and *hold*) being a bang-bang control of some nature. The most common type of bang-bang control method is the Bosch algorithm (Bauer and Girling, 1999). This algorithm accepts the wheel angular acceleration and longitudinal slip as inputs, and makes use of predefined thresholds to control the brake pressure between the three different brake states. By combining these control modes in various configurations, the Bosch algorithm maintains a repetitive ABS control cycle consisting of eight phases. This approach is very robust, and studies have shown that by fine tuning the parameters of the Bosch algorithm, reasonable results can be obtained when braking over rough terrain (Penny, 2016; Van der Merwe et al., 2018). This algorithm, however, is unable to adapt to different terrain without adjusting the predefined thresholds, while also assuming that the normal force at the tyre contact patch is consistent.

As technology continues to grow and develop, there is the continuous need to design control systems that are capable of maintaining acceptable performance levels when faced with uncertainties. This has lead to the rise in the development of control system schemes that emulate intelligent biological systems. Intelligent methods including fuzzy logic, neural networks, evolutionary algorithms, and machine learning have been incorporated into the design of control systems in an attempt to make them more robust.

2.2.3.2 Intelligent Fuzzy and Sliding-Mode Control

Fuzzy Control (FC) methods or Fuzzy-Logic Controllers (FLC) are capable of controlling non-linear systems such as ABS, with added advantages including robustness and rule-based algorithms (Lee, 1990). Keshmiri and Shahri (2007) implemented an intelligent ABS fuzzy controller capable of adjusting its slipping performance for different friction surfaces (dry, wet and icy asphalt road). This approach uses the vehicle acceleration to identify the type of road and adapt the controller accordingly, yielding reliable results for the different friction surfaces. Despite showing satisfactory performance in the handling of the uncertain dynamics of the slip ratio, the FC methods make use of a large amount of the fuzzy rules, causing a time-consuming design process to select the best parameters and a complex analysis thereof. As an alternative, Sliding-Mode Control (SMC) methods can be used.

SMC is a robust nonlinear control method that alters the dynamics of a nonlinear system by application of a high-frequency switching control. Bhandari et al. (2012) proposed the use of an SMC to control the wheel slip ratio at values predetermined by a surface prediction algorithm, according to the different surface terrains encountered (dry concrete, dry and wet asphalt, dry and wet cobblestone, and snow). The desired slip ratio for maximum friction coefficient is provided as input to the SMC via a Proportional Integral Derivative (PID) controller, from which the SMC outputs the braking torque, as shown in Figure 2.9. Results show that the combination of PID control for estimation and SMC control to achieve the estimated slip ratio for maximum friction coefficient, are capable of reducing braking distances by more than 3% during the different surface changes encountered during braking.

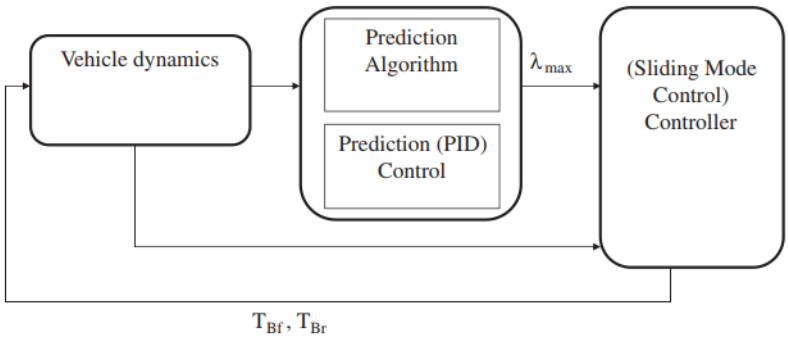


Figure 2.9: Model breakdown of the SMC (Bhandari et al., 2012)

2.2.3.3 Model-free Intelligent Control

Owing to the difficulty involved in developing a mathematical model that is capable of describing all the dynamics of the non-linear ABS control problem, researchers have turned towards the development of intelligent control methods such as Neural Networks (NN) that are capable of mitigating the effects of an inaccurate mathematical model and adapting to different terrains.

John and Pedro (2013) propose a multi-layered Neural Network-based Feedback Linearization (NNFBL) control scheme over dry asphalt road. This approach follows the tracking of a pre-determined slip ratio, with results proving that slip regulation using a NN-based control scheme is feasible and robust to external disturbances. Poursamad (2009) proposes an adaptive NN-based hybrid controller, which combines two NN's to learn the nonlinearities of the ABS associated with the hybrid feedback linearization controller. Results indicate desirable performances under the transition between various flat friction surfaces (dry and icy asphalt road). Lin and Hsu (2003) propose a hybrid control system with a Recurrent Neural Network (RNN) observer, which shows effective braking responses over various flat friction surfaces (dry, wet, and icy asphalt road).

The majority of ABS control approaches, including those mentioned previously, are model-based control approaches that rely on a set of rules and/or require their decision logic to be manually tuned. This has its pitfalls, especially in the application over rough terrain, where the non-linear dynamics are excited further and a large volume of noise is present. Mismatches tend to occur between the control design model and the real process (Radac and Precup, 2018), and as a result, the un-modelled dynamics and parametric uncertainties lead to underwhelming performances of the ABS. As an alternative, research has turned to intelligent model-free control techniques that make use of data (usually input-output) in their control design, instead of modelling constraints.

Reinforcement Learning (RL), also known as Adaptive Dynamic Programming (ADP), is a representative of data-driven techniques that exploit the process structure used when designing the control. This field is a relatively new concept, however its application to non-linear braking applications have been explored with promising results. Fu et al. (2020) propose a deep deterministic policy gradient braking model for autonomous braking in emergency situations. This approach factors in the passengers' comfort using a multi-objective reward function, with results highlighting the efficacy of the controller to make fast decisions in a continuous control environment. Zhao et al. (2013) propose a supervised actor-critic approach for adaptive cruise

control. This approach makes use of two NN's to approximate the optimal control policy for adaptive cruise control and can be adapted to the control of an ABS. Sardarmehni and Heydari (2015) propose the use of ADP for the optimal control of an ABS over smooth terrain. This approach relies on penalizing the braking distance of the vehicle and outperforms many existing solutions in literature. Drechsler et al. (2021) investigate an actor-critic algorithm as a form of traction control. This method tracks the position of the accelerator pedal and wheel slippage, and results show the ability of the controller to maintain the slip ratio under adequate levels on different surfaces (dry, and ice asphalt). Radac and Precup (2018) propose the design and implementation of a model-free slip control of ABS through Q-learning, with the results illustrating the algorithms adaptability and robustness over smooth terrain for a wide operating range.

2.3 Research Question and Conclusion

Tyre-terrain interactions are crucial to the performance of an ABS, and multiple tyre models capable of capturing these interactions are introduced. The LuGre tyre model presents a timesaving solution due to its easy implementation and ability to capture the transient behaviour of the tyre, while the more complex 3D FTire tyre model is found to be the most advanced tyre model that is capable of accurately simulating and capturing the vertical excitations that occur over rough terrain.

The control of an ABS is a highly non-linear problem, and very little literature on off-road ABS control techniques exist. It is the aim of this study to address this. Two model-free reinforcement learning ABS control approaches have demonstrated great adaptability and robustness over smooth terrain, leading to the question: Can a model-free technique be used successfully over rough terrain?

Literature reveals that the intelligent model-free control techniques do not have modelling constraints, thus are able overcome the pitfalls of un-modelled dynamics and parametric uncertainties that come with a highly non-linear problem. The desired control technique needs to be able to learn, adapt, and make decisions based off an unknown environment, without manually tuning the decision logic. Reinforcement learning may achieve this, leading to a better performance of the ABS over rough terrain, and is therefore investigated further in Chapter 3.

Chapter 3: Reinforcement Learning

Machine Learning (ML) is defined as the application of Artificial Intelligence (AI) to a system, enabling it to automatically learn and improve from previous experiences, without being explicitly programmed. ML algorithms are categorized into three different types: supervised, unsupervised, and reinforcement learning (RL) (Sutton and Barto, 2018).

Supervised and unsupervised learning make use of data that is readily available. The former is presented with labelled inputs along with their desired outputs, and has the objective of learning the rule that maps the inputs to the outputs. The latter is provided with collections of unlabelled data, and the objective of finding hidden structure within the dataset. RL is a method that learns through the interaction between an autonomous active decision-making agent and its dynamic environment as seen in Figure 3.1, with the objective of the agent to maximize a reward signal despite uncertainty about its environment (Sutton and Barto, 2018).

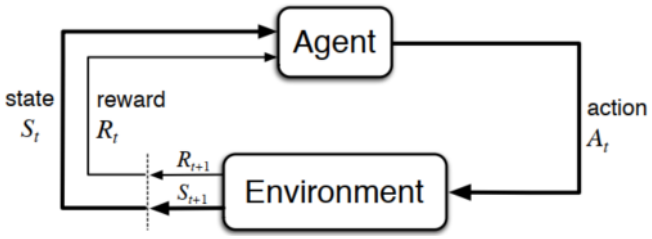


Figure 3.1: The interaction loop between the environment and the agent (Sutton and Barto, 2018)

The application of ML to real-life scenarios is difficult for numerous reasons. These include the expense and time required to create a database, limited information about the problem is available, and the dynamic environment is often unknown - similar to that of the non-linear control of an ABS over a rough terrain. However, through the application of RL - which learns by way of trial and error in unforeseen environments - an optimal solution may be obtained.

An introduction to RL and its different aspects, the benefit of time-based function approximators, and transfer learning are explored in this chapter.

3.1 Background

Formally, the process of the agent-environment interaction is a Markov Decision Process (MDP) problem, and consists of a set of states S, a set of actions A, a transition function $T(s'|s,a)$, and a reward function $R(s,a)$. In each state $s,s' \in S$, the agent takes an action $a \in A$, upon which it receives reward and the next state s' as determined from the transition probability distribution. The action-value function, also known as the Q function $Q^{\pi}(s,a)$, is determined through the Bellman equation, with the optimal state-action value function shown in Equation (3.1). This represents the expected long-term rewards that can be obtained from a specific state-action pair (s,a) (Sutton and Barto, 2018).

$$Q^*(s,a) = R(s,a) + \gamma \sum_{s'} P\left(s' \mid s,a\right) \max_{a'} Q^*\left(s',a'\right), \tag{3.1}$$

where $0 < \gamma < 1$ is the discount factor, and dictates how much the agent cares about future rewards relative to its immediate reward. The goal of the agent is to find the policy π that maximises the expected discounted total reward over the agents lifetime. This is defined as the optimal policy π^*, and is represented in Equation (3.2).

$$\pi^*(s) = \text{argmax}_a \, Q^*(s,a). \tag{3.2}$$

Beyond the agent and environment, definitions of the four subsystems can be defined as (Sutton and Barto, 2018):

1. A policy - a strategy the agent follows to discern its next action based on the current state.
2. A reward function - an instant return awarded when the agent executes a particular action; influenced by the current state, the action just performed, and the next state.
3. A value function - expected reward by an agent for being in a state s. Also known as state-value function.
4. A model of the environment - which simulates the behaviour of the environment.

RL methods can be divided into two broad classes: *model-based* and *model-free* learning, with both considered.

3.2 Model-based RL

Model-based RL methods are a deductive problem-solving approach that applies a model, given or learned, from which the agent will use its general understanding to construct a best action. The agent learns the model of the domain by approximating $R(s,a)$ and $T(s'|s,a)$, and from this it

determines the optimal policy through a method such as value iteration. Despite allowing for quicker learning, model-based RL algorithms typically take too much time between each action for practical on-line learning (Hester et al., 2012), and the model agent is only as good as the model learnt. This approach is unsuitable, as the perfect ABS controller that is able to overcome the un-modelled dynamics and parametric uncertainties that occur over rough terrain is not known.

3.3 Model-free RL

Model-free RL methods are an inductive problem-solving approach where the agent makes use of empirical evidence from past experiences; they are applicable in situations where $R(s,a)$ and $T(s'|s,a)$ are unknown. Despite not being able to approximate off past actions, model-free approaches have been shown to have strong convergence guarantees (Sutton and Barto, 2018). Basic model-free approaches update according to Temporal Difference (TD) learning. This enables the agent to learn from its environment through episodes with no prior knowledge of the environment, and is defined by Equation (3.3):

$$V_{i+1}(s) = V_i(s) + \alpha\left(r + \gamma V_i\left(s'\right) - V_i(s)\right), \tag{3.3}$$

where $V(s)$ is the value function, and α the step size. Agent training and representation with model-free RL can be approached in two main ways: *Q-Learning* and *Policy Optimization* (OpenAI, 2018). A structural outline of the two classes of RL as well as the two categories of model-free methods are presented in Figure 3.2. Summaries of the more relevant examples are explored further.

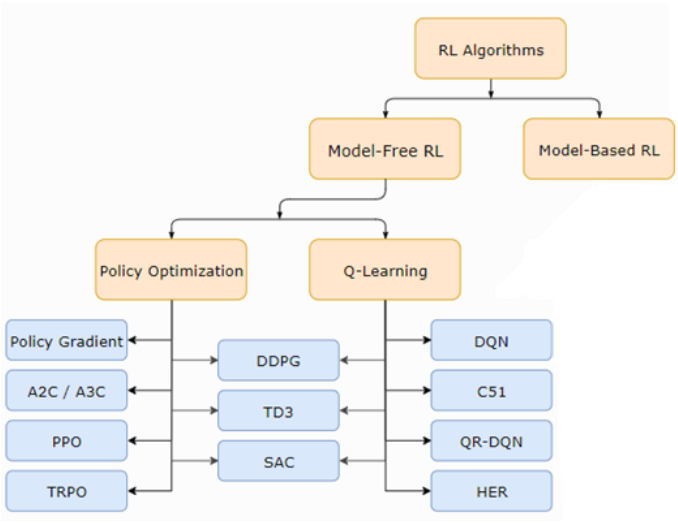


Figure 3.2: Taxonomy of model-free RL algorithms (Adapted from OpenAI (2018))

3.3.1 Q-Learning

Q-Learning is a basic form of RL that makes use of Q-values Q (action value functions as defined in Equation (3.1)) to iteratively improve the behaviour of the agent. This approach is almost exclusively achieved through off-policy methods, meaning that each update is not dependent on the agent, and data obtained at any stage during training can be used. The estimation of Q is achieved iteratively through the TD update rule (Equation (3.3)), until the optimal action-value function $Q^*(s,a)$ is achieved (Watkins and Dayan, 1992). Q-learning is easy to implement and has a high sample efficiency, but is an unstable technique plagued by the complexities of high dimensional environments (Mnih et al., 2013). Deep Q-Learning Networks (DQN) are used to overcome this disadvantage, and are explored below.

3.3.1.1 Deep Q-learning Network (DQN)

DQN's make use of a deep NN - a NN with multiple layers - which map state-action pairs to Q-values, and use two important techniques: 1) a target network that makes use of the same parameters as the online NN and acts as an error measure, and 2) an experience replay that is used to store the agent's experiences at every time-step $e_t = (s_t, a_t, r_t, s_{t+1})$ into a dataset D. After each e_t, the agent executes an action according to an ϵ-greedy policy, which initially encourages the agent to explore the environment, before exploiting the best policies (Mnih et al., 2013). DQN is simple to implement, and its structure of learning is provided in Figure 3.3.

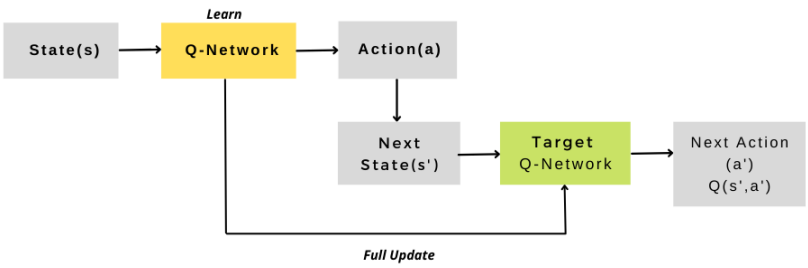


Figure 3.3: Structure of DQN

3.3.1.2 Double Deep Q-learning Network (DDQN)

A disadvantage of the DQN is its overestimation of Q-values during learning, leading to unstable learning and underwhelming performances (Van Hasselt et al., 2016). To overcome this, a simple change is proposed - the use of two Q-networks. The existing Q-network is used to evaluate the greedy policy and select its action, while the target network is used to estimate its Q-value. This approach is known as the Double Deep Q-Learning Network (DDQN), and not only reduces the

overestimations while keeping the rest of the DQN algorithm intact, but also leads to better performance (Van Hasselt et al., 2016). The structure of learning is provided in Figure 3.4.

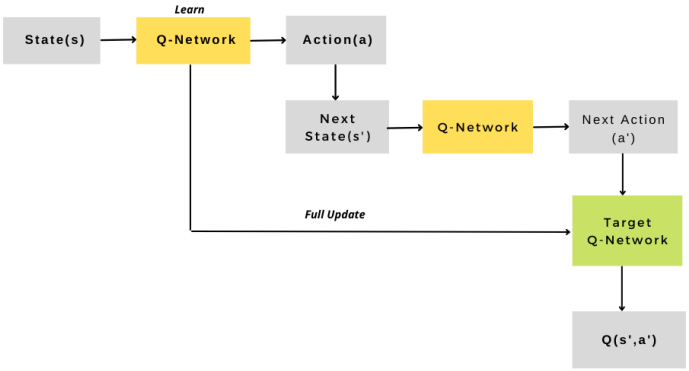


Figure 3.4: Structure of DDQN

3.3.2 Policy Optimization

The RL methods considered thus far have been action-value methods that choose an optimal policy based on the values of the actions. Policy Optimization methods learn a parameterized policy that can select actions without consulting a value function (Sutton and Barto, 2018). The value function may still be used to learn the policy parameter, however it is not required to select the action. These methods are referred to as policy gradient methods, and they directly parameterize the policy $\pi(a|s,\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the policy's parameter vector, and is updated according to Equation (3.4). Policy gradient methods maximise the performance measure of $J(\boldsymbol{\theta})$ with respect to the policy parameter through gradient ascent.

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \alpha\widehat{\nabla J(\boldsymbol{\theta}_t)}. \tag{3.4}$$

Methods that learn approximations to both policy and value functions are referred to as *actor-critic* methods. The policy structure is known as the *actor* as it is used to select actions, and the estimated value function as the *critic* because it criticizes the actions made by the actor. Updates are done each step (TD learning), with the critique taking the form of the TD learning error (second half of Equation (3.3)).

The biggest advantage of policy optimization is its ability to inject prior knowledge about the desired form of the policy into the reinforcement learning system (Sutton and Barto, 2018). Additional advantages that policy optimization methods have over action-value methods include approximating the policy as a simpler function, and enabling the selection of actions with arbitrary probabilities. The synchronous and asynchronous advantage actor-critic, and soft actor-critic algorithms are explored.

3.3.2.1 Advantage Actor-Critic

The advantage actor-critic makes use of an advantage function defined in Equation (3.5), which calculates the agents TD error, and allows the agent to seek out good states and avoid bad states (Mnih et al., 2016).

$$A(s,a) = r + \gamma V\left(s'\right) - V(s). \tag{3.5}$$

The Advantage Actor Critic has two main variants: the Asynchronous Advantage Actor-Critic (A3C) and Advantage Actor-Critic (A2C).

A3C introduced by Mnih et al. (2016), operates by asynchronously executing different agents in parallel on multiple instances of the environment, where the multiple workers update a global value function independently. The A3C is shown to outperform conventional RL algorithms such as the DQN on several Atari games (Mnih et al., 2016). Disadvantages of the A3C are its asynchronous nature, in which some workers will play older versions of the parameters (meaning aggregating the update is not optimal), and its unstable run-to-run learning. This implies that the algorithm produces inconsistent results with each training simulation.

As an alternative, the A2C is a synchronous implementation that waits for each actor to finish its segment of experience before updating the global network, averaging over all of the actors. The A2C is found to produce comparable performances to the A3C but with added efficiency, resulting in faster and more cohesive training. However, it too suffers from the unstable run-to-run learning.

3.3.2.2 Soft Actor-Critic (SAC)

Soft Actor-Critic (SAC) is an off-policy algorithm that makes use of entropy regularization. This feature encourages the agent to maximise a trade-off between expected reward while also maximizing entropy - a measure of randomness. That is, to succeed at the task while acting as randomly as possible (Haarnoja et al., 2018). SAC concurrently learns a policy as well as two Q-functions. Despite being suited for continuous action settings, changes to the policy update rule can result in a SAC for discrete settings (Christodoulou, 2019). The SAC presents a number of advantages that include: encouraging the policy to explore more widely while giving up on clearly unpromising avenues, and improving the learning speed over state-of-art methods that optimize the conventional RL objective function (Haarnoja et al., 2018).

Examples of both Q-learning and policy optimization have been used in vehicle control, specifically during braking manoeuvres (Zhao et al., 2013; Sardarmehni and Heydari, 2015; Radac and Precup,

2018), rendering the aforementioned examples applicable to this study.

3.4 Function Approximators

Current state of the art RL algorithms make use of deep NNs as their function approximators, however real-life applications generally involve time series data, as is the case with the control of an ABS. This data contains an added complexity of sequence dependency amongst the input variables. Knowing this, it may be beneficial to store this information over a long period of time and learn from it, adding an aspect of predictive modelling to the control problem. Two types of NNs that are capable of handling sequence dependency and the modelling of time series data are Recurrent Neural Network (RNN) and Convolutional Neural Networks (CNN). Recent research highlights the Temporal Convolutional Network (TCN), a variation of the CNN, as a promising solution for the sequence modelling tasks (Bai et al., 2018).

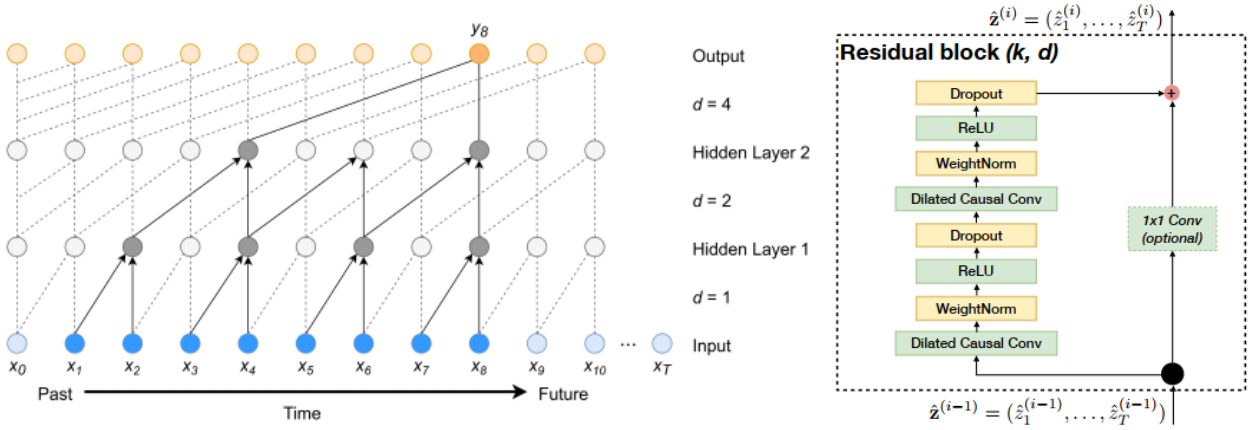
3.4.1 Temporal Convolutional Network (TCN)

TCN is a CNN architecture that is based on two principles:

1. The architecture is capable of producing an output the same length as the input
2. The convolutions in the architecture are causal - no information leakage between future and past

The first point is accomplished by using a 1D Fully Convolutional Network (FCN) architecture where the filters exhibit only one dimension (time). The second point is achieved by making use of dilated causal convolutions, where an output at time t is convolved with elements from time t and earlier in the previous layer. Figure 3.5(a) demonstrates this, where y_8 is predicted on the combination of inputs $x_1,...,x_8$. The dilated convolutions enable an exponentially larger receptive field, however, since this receptive field depends on the network depth n as well as filter size k and dilation factor d, it is important to ensure stabilization of deeper and larger TCNs. This is achieved by a residual block as seen in Figure 3.5(b), which effectively allows layers to learn modifications to the identity mapping rather than the entire transformation (Bai et al., 2018).

Results show that the TCN not only exhibits longer effective memory but consistently outperforms general RNNs such as the Long Short Term Memory (LSTM), on a vast range of tasks (Bai et al., 2018).



(a) (b)

Figure 3.5: Architectural elements in a TCN. (a) A dilated causal convolution with dilation factors d = 1; 2; 4 and filter size k = 3. The receptive field is able to cover all values from the input sequence. (b) TCN residual block (Bai et al., 2018).

3.5 Transfer Learning

Conventional ML and deep learning algorithms are traditionally designed to work in isolation, where they solve a specific task, and have to be rebuilt once the feature-space distribution changes. This may not hold in many real world applications. Transfer learning is the idea of overcoming the isolated learning paradigm, utilizing knowledge acquired for one task to solve related tasks (Pan and Yang, 2009). This enables utilizing existing knowledge from a relatively simple task such as knowing how to ride a bicycle, to a more complex task such as learning how to drive a car. Transfer learning can be separated into three categories: inductive transfer learning, transductive transfer learning, and unsupervised transfer learning. Inductive transfer learning involves the source and the target having the same domains, with related but different tasks, and is representative of deep learning algorithms.

Benefits of transferring prior knowledge include overcoming the need for large volumes of new data, using a pre-trained model to speed up the process of training on new tasks, and increasing the overall accuracy and efficiency of the model. A limitation to transfer learning is negative transfer - transferring negative/unwanted information across tasks (Pan and Yang, 2009).

3.6 Conclusion

An introduction to reinforcement learning is provided, and due to its ability to learn from an environment with no prior knowledge nor model constraints, it is proposed that a model-free

intelligent controller be used to overcome the challenges that an ABS experiences over rough terrain. Focus is placed on the following model-free algorithms: Q-learning (DQN and DDQN), and policy gradient methods (A2C, A3C and SAC), and an investigation/comparison of each algorithm is recommended to identify the best suitable algorithm for the problem.

The advantages of time-series based function approximators, as well as the transfer of prior knowledge from a simple task (such as a single tyre model) to a more complex task (such as a full vehicle model) are explored, and may be beneficial to be used in conjunction with the aforementioned RL algorithms.

3.6.1 Updated Problem Statement

The original problem statement as defined in Section 1.2 read: *The development of an ABS control method that is capable of detecting and reducing high slip conditions on rough terrain in an attempt to prevent wheel lockup, while maintaining lateral control and a good stopping distance.* This can be expanded into the final problem statement:

The development of a model-free intelligent ABS control method that is capable of detecting and reducing high slip conditions on rough terrain in an attempt to prevent wheel lockup, while maintaining lateral control and a good stopping distance.

3.6.2 Study Overview

An overview of the structure and approach followed in this study is provided by the organisational chart in Figure 3.6, and can be split into two main sections: 1) the parameters that affect the performance of an ABS which are considered in this study, and 2) the different types of simulations considered.

Among the many parameters that are associated with the effectiveness of an ABS, the terrain, ABS algorithm, tyre model, and suspension setup used are considered in this study. The type of terrain directly affects the performance of an ABS controller, and in this case a combination of smooth surfaces (high and low friction) and rough terrain (Belgian paving and ISO 8608 class D road) are considered. Emphasis is placed on rough terrain and how the different braking configurations respond over it. The type of ABS algorithm forms the main focus of this study, with the proposed intelligent algorithm compared to a Bosch algorithm tuned for off-road braking and a conventional braking system (no ABS). The FTire tyre model is found to be the most accurate and effective over rough terrain (Stallmann and Els, 2014), and is proposed as the tyre model to be used in ADAMS in

conjunction with a validated full vehicle simulation model. The final parameter considered is the type of suspension setup used in the vehicle. Hamersma and Els (2014) found that a hard and soft damper and suspension setup, in the front and rear respectively, resulted in the best braking performance over rough terrain. It is proposed that this suspension setup be used in all simulations.

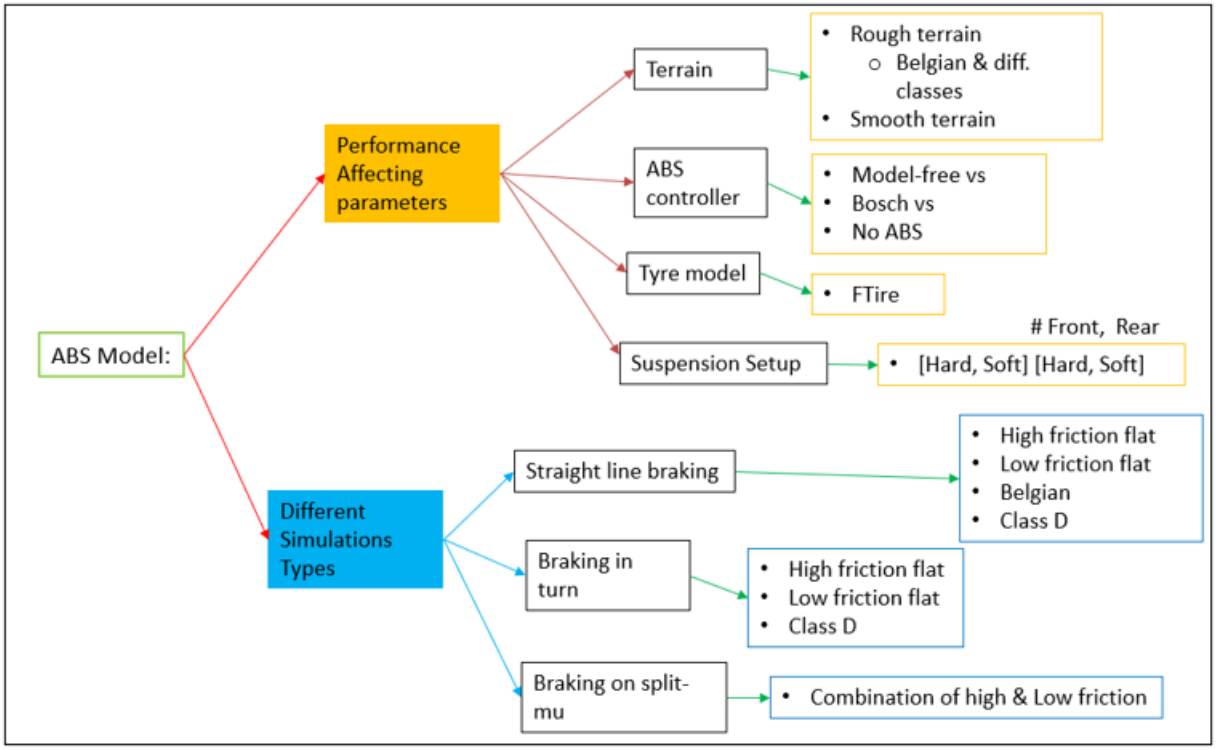


Figure 3.6: Organisational overview of this study

A future recommendation is the addition of performance affecting parameters such as the effect of tyre carcass oscillation and the run-in-effect of tyre force generation, with the ideal ABS controller being able to accommodate for all aspects. The second section relates to the different simulation tests run according to the requirements set by SAE (2014). Eight braking scenarios are to be simulated through three main types: straight line braking, braking-in-turn, and braking on a split-mu surface. The robustness of the control algorithm against noisy measurement is also investigated.

Prior to performing these simulations, it is necessary to explore the initial aspects of the intelligent control algorithm such as the appropriate RL algorithm, reward function, and function approximator. Chapter 4 investigates these aspects through multiple simulations using a simplified quarter car model to reduce complexity and solution times.

Chapter 4: Single Tyre Simulation

In this chapter the baseline controller, a Bosch algorithm tuned for off-road braking, is introduced. A simplified quarter model with tyre characteristics modelled by the LuGre tyre model is introduced and used in an initial study to determine the aspects of the proposed intelligent control method.

4.1 Bosch Algorithm Modelling

The best-documented ABS algorithm was published by Horst Bauer and Peter Girling in 1999 (Bauer and Girling, 1999), and is known as the Bosch algorithm. Studies (Penny, 2016; Van der Merwe et al., 2018) have shown that by fine tuning the parameters of the Bosch algorithm, reasonable results can be obtained when braking over rough terrain, emphasizing its suitability as a baseline algorithm. The Bosch algorithm is a bang bang type, rule-based control strategy that takes the wheel angular acceleration and longitudinal slip as inputs, and makes use of three different pressure control modes to control the brake pressure at each wheel: *pump*, *dump*, and *hold*. Dictated by two upper and one lower threshold for the angular acceleration, and an upper longitudinal slip threshold, the Bosch algorithm combines these three pressure control modes in various configurations resulting in a repetitive ABS control cycle consisting of eight phases as illustrated in Figure 4.1.

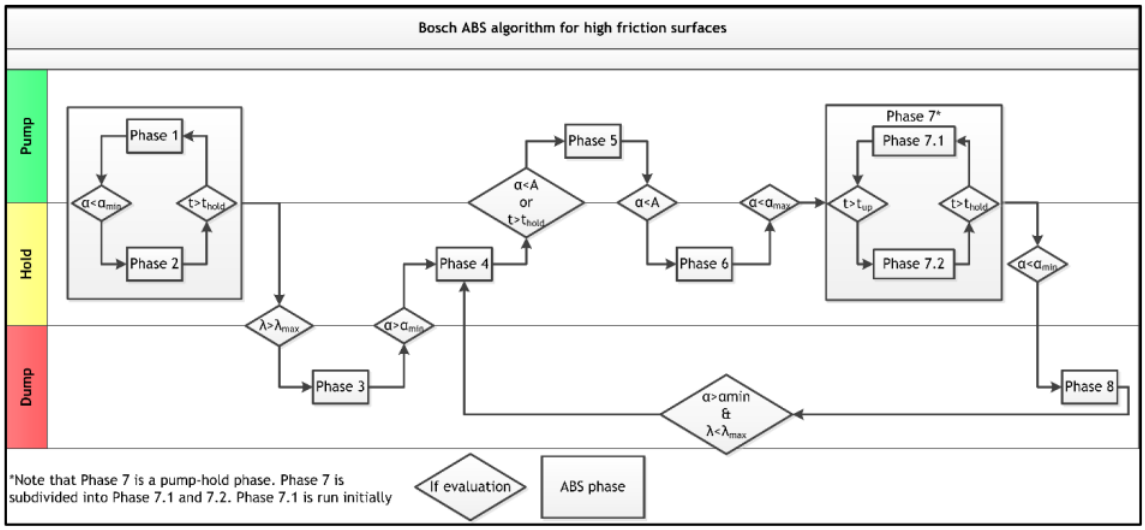


Figure 4.1: Schematic of Bosch algorithm (Hamersma, 2017)

The Bosch control cycle begins at Phase 1, which is the initial application of the brake pedal by the

driver. Phase 2 is activated once the minimum angular acceleration threshold of the wheel has been exceeded, during which the pressure is held constant and is time limited to prevent the controller from remaining in a hold-phase. The brake pressure is gradually built up through cycling between phases 1 and 2 until the slip threshold is exceeded - initiating Phase 3. The brake pressure is reduced until the angular acceleration is below the minimum acceleration threshold - initiating Phase 4 where the pressure is held constant (and is time limited) until the angular acceleration exceeds the maximum acceleration threshold. Once the maximum acceleration threshold has been exceeded, Phase 5 begins and the brake pressure is increased as rapidly as possible. Phase 6 commences once the angular acceleration drops below the maximum acceleration threshold, and involves holding the pressure until the angular acceleration falls below the lesser of the two upper acceleration thresholds. A time limit is applied to Phase 6 as well. The brake pressure is increased in a stepwise manner in Phase 7, and continues until the angular acceleration drops below the minimum acceleration threshold. The final phase is similar to Phase 3, and the control cycle is repeated.

An example of the ideal Bosch control cycle is presented in Figure 4.2. Using the thresholds defined by Penny (2016) as reference, the thresholds used in this study are summarised in Table 4.1. The values A and α_{max} denote the two upper angular acceleration thresholds, α_{min} the lower angular acceleration threshold, and λ_{max} the upper slip threshold. t_{up} and t_{down} are time limits used in the control cycle to prevent the algorithm from becoming stuck in a certain phase.

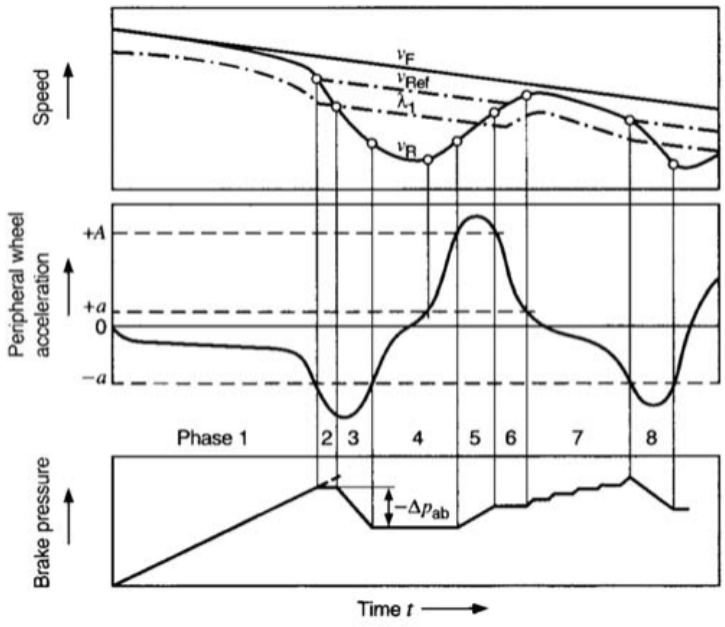


Figure 4.2: Idealized Bosch control cycle (Bauer and Girling, 1999)

Table 4.1: Thresholds used for Bosch Algorithm

Threshold	Value	Unit
A	25	rad/s^2
α_{max}	5	rad/s^2
α_{min}	-100	rad/s^2
λ_{max}	0.2	-
t_{up}	10	ms
t_{down}	10	ms

4.2 Single Tyre Model

To reduce solution times and determine the initial aspects of the ABS controller such as which RL algorithm to use, the appropriate reward function, and function approximator, the simplified motion dynamics of a quarter-vehicle model are used in the initial stages of this study. The single tyre model is demonstrated in Figure 4.3.

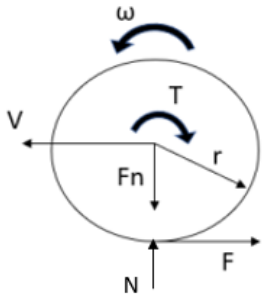


Figure 4.3: A single tyre free-body diagram with lumped friction

This model only considers longitudinal motion and takes into account the normal and friction forces acting on the tyre, the brake torque, and angular velocity of the wheel. Suspension dynamics and effects such as the torsional deflection of the sidewall, rolling resistance, and transportation delay caused by the brake hydraulic systems are neglected. The single tyre model provides a link between tyre braking dynamics and brake hydraulic dynamics, and although simple, retains the essential characteristics of the actual system.

Applying first principles to the free-body diagram, with the sum of the forces acting positive in the left direction and the sum of moments acting positive in the anticlockwise direction, a system of the

form is obtained:

$$m\dot{v} = -F, \tag{4.1}$$

$$I\dot{\omega} = rF - T_b, \tag{4.2}$$

where m is a 1/4 of the vehicle mass, I the moment of inertia of the wheel about the spin axis, and r is the radius of the wheel. The vehicle velocity, angular velocity, brake torque, and tyre-road friction force are represented by v, ω, T_b, and F respectively. For the sake of simplicity, the vertical force is assumed to act in line with the centre of the wheel, neglecting effects of rubber pneumatic trail. The tyre-road friction force is described by the lumped LuGre model:

$$\dot{z} = v_r - \frac{\sigma_0 |v_r|}{g(v_r)} z. \tag{4.3}$$

$$F = (\sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v_r) F_n, \tag{4.4}$$

where

$$g(v_r) = \mu_c + \left(\mu_s - \mu_c\right) e^{-|v_r/v_s|^{\alpha}}. \tag{4.5}$$

The brake torque is the force applied at the brake disc to stop its motion and is defined as a linear relationship with the brake pressure in Equation (4.6) (Penny, 2016).

$$T_b = 271 \times P, \tag{4.6}$$

$$\dot{P} = \frac{P_{set} - P}{\tau}. \tag{4.7}$$

Control of the pressure as a function of time ($\dot{P}$) is derived from the Laplace transform of a first order system with a time delay constant τ, and is equal to the difference between the backpressure (P_{set}) and the current pressure in the valve, divided by τ. The time delay constant defines how quickly the pressure changes and is dependant on many factors including the size of the valves. Therefore τ is different for the pump and dump phases, and is set according to the thresholds used in (Penny, 2016):

- Pump: P_{set} = 10 MPa & τ = 200 ms
- Dump: P_{set} = 0 MPa & τ = 50 ms
- Hold: $P_{set} = P$

The relationship between the braking friction coefficient and wheel slip at steady state conditions is investigated to establish reasonable values for the LuGre tyre model parameters. These parameters are defined in Table 4.2 and a plot of the steady state friction coefficient μ as a function of the slip is provided in Figure 4.4.

This figure plots the relationship between the braking coefficient and wheel slip, and correlates with literature (Canudas-de Wit et al., 2003; Aly et al., 2011). It can be seen that the optimal slip occurs within the region of 3-20% with the peak brake coefficient μ_p around 5%. Therefore considering the four nonlinear differential equations $[\dot{P}, \dot{z}, \dot{v}, \dot{\omega}]$, the relationship of the system dynamics and how they interact with each other is established.

Table 4.2: Parameters used in the LuGre model

Parameter	σ_0	σ_1	σ_2	μ_c	μ_s	v_s
Value	160	3.5	0.0018	0.8	1.55	6.57

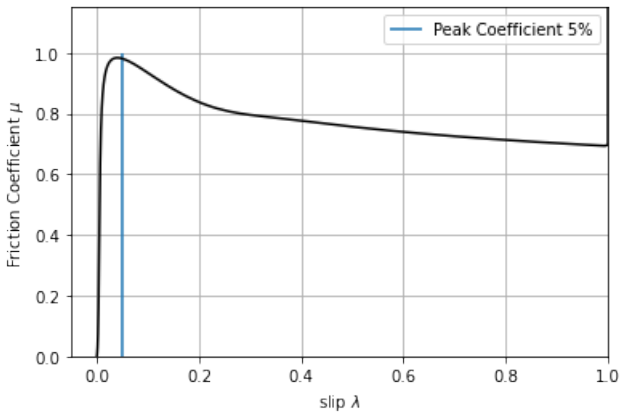


Figure 4.4: Relationship between the steady state brake friction coefficient as a function of the slip

4.3 Model of Controller

From the literature review (Section 3.6) it is proposed that the different Q-learning (DQN and DDQN) and policy gradient methods (A2C, A3C and SAC) be compared as potential ABS control algorithms in order to find the best suitable method for this problem. Due to the availability of baseline implementations for RL algorithms, the GitHub repository of Deep-Reinforcement-Learning-Algorithms-with-PyTorch as provided by Christodoulou (2017) is used as the main source of implementation and comparison of the above RL algorithms. Adjustments are made where necessary to the make the algorithms suitable for the aforementioned environment described in Section 4.2, with the different hyperparameters used during training of the different algorithms summarised in Appendix A.

In order to model the controller, a state space and action space are to be defined. Several combinations of states can be used as the input such as: pressure, velocity, angular velocity, bristle deflection, and/or longitudinal slip. These inputs however need to be easily accessible/measurable in a physical system and should avoid redundancy, for example, using the vehicle velocity, angular velocity, and longitudinal slip (which is a function of the two previous parameters). Therefore a state space of $[v/r, w]$ is proposed. These two parameters are able to sufficiently detail when wheel lockup is imminent as well as determine the longitudinal slip of the wheel. The vehicle velocity is divided by the wheel radius for two reasons: firstly, to normalise it with respect to the angular velocity, and secondly, to ensure that both inputs are independent of wheel radius. A discrete action setting of [pump, dump, hold] is provided as output. A schematic of the mapping from the state space to the action space is provided in Figure 4.5.

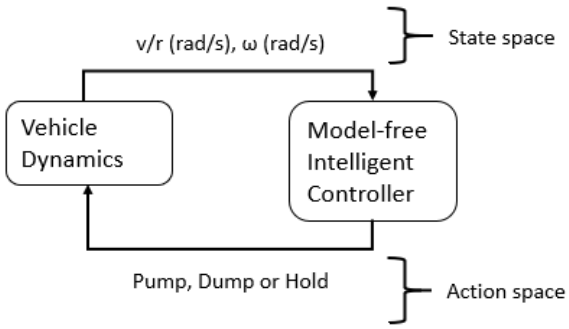


Figure 4.5: Model breakdown of the proposed controller

4.3.1 Reward Function

The reward function determines the behaviour of the brake control, therefore to ensure reliability of the brake control it is crucial to use a properly defined reward function. Different aspects in determining the ideal reward/penalty function exist, such as 1) regulating the longitudinal slip by means of predetermined slip thresholds, and/or 2) monitoring the stopping distance of the vehicle. The first approach is the most commonly used penalty system for ABS systems, however it restricts the model, thus no longer represents a model-free approach, and is limited by the type of terrain the vehicle travels over, requiring it to adapt accordingly. The second approach neglects the importance of the lateral control of the vehicle and generates rewards that are sparse, rendering it undesirable.

When considering the reward function for the ABS model, there is conflict between two intuitive objectives of the ABS: 1) the vehicle is encouraged to brake hard to achieve the shortest possible stopping distance, and 2) the wheels may not lock/slide, which is common during hard braking. If

it is unbalanced, the agent becomes either too conservative or reckless, leading to slow braking or locked wheels respectively. Therefore two reward functions are proposed with the aim of identifying the most suitable reward function for this study.

To meet the first objective, both functions serve as a penalty function which always returns a negative reward, resulting in the agent trying to minimize the total penalties received. The penalty function also penalizes the duration of the braking process; the longer the vehicle brakes - the longer the amount of penalties received. This encourages the agent to slow down the vehicle as quickly as possible. To further reduce the stopping time, the braking force can be considered via either pressure or longitudinal acceleration. Equation 4.8 involves the use of pressure, and knowing that the maximum pressure pumped from the brake value is 10 MPa, the negative rewards are achieved by vertically shifting the rewards to below 0 with the constant value -10. Equation 4.9 makes use of the longitudinal acceleration, which alternatively to the longitudinal friction, is directly measurable and ensures high deceleration of the vehicle since $F_{max}\ \alpha\ a_{max}$. The longitudinal acceleration is scaled by a factor of 0.5 and vertically shifted below 0 by the constant value of -10, to allow for direct comparison between the two reward functions.

To satisfy the second objective, the function $j(\lambda)$ is introduced. This function is dependant on the longitudinal slip and penalizes the model as the wheels begin to slip. Maximum braking forces tend to occur between the region of 10-30% slip (Aly et al., 2011), however the peak braking force is closer to the 10% region, with lockup occurring almost instantly past 30% slip. Therefore slip greater than 20% is penalized by a constant of 15. This should prevent lockup occurring and keep the friction oscillating around the peak. Due to the inherent system response delay, perfect control of the slip will not be achieved, resulting in some excess traction to steer. Both of these reward functions assume nothing of the environment except that the peak friction is below 20% slip.

$$R = P - 10 - j(\lambda), \quad j(\lambda) = \begin{cases} 15\lambda & \lambda > 20\% \\ 0 & \lambda < 20\% \end{cases} \tag{4.8}$$

$$R_2 = \frac{1}{2}a - 10 - j(\lambda), \quad j(\lambda) = \begin{cases} 15\lambda & \lambda > 20\% \\ 0 & \lambda < 20\% \end{cases} \tag{4.9}$$

A theoretical representation of the two reward functions is provided in Figure 4.6 (which is a close up of Figure 4.4), where the two opposing objectives are shown trying to achieve the desired ABS slip control region which would allow for the greatest lateral control and shortest stopping distance.

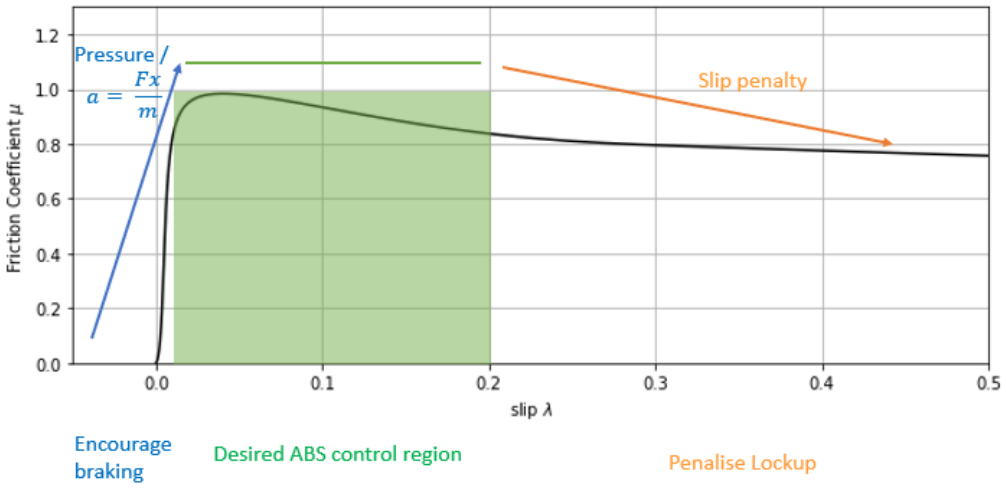


Figure 4.6: Direct visualisation of the different reward functions

4.3.2 Training and Results on Smooth Terrain

Braking over rough terrain causes an increase in complexity and difficulty in the control of an ABS. Initial tests are conducted using smooth terrain to ensure the viability of the different RL algorithms over such terrain. If the RL algorithms cannot prevent wheel lockup over smooth terrain, they will be unable to do so over rough terrain. To reduce solution times and complexity, the single tyre model is trained doing straight-line braking using the different Q-learning (DQN and DDQN) and policy gradient methods (A2C, A3C and SAC), and the use of a TCN function approximator is neglected. The pressure-based reward function (Equation 4.8) is used, and the simulation is terminated below speeds of 2 m/s as ABS algorithms tend to stop working at low speeds. To ensure that the agents are learning generalization and not simply memorizing the specifics of the environments, the following stochastic measures are introduced and explained further:

1. The starting speeds of the vehicle are randomised between 15 and 19m/s
2. Randomised samples of uniformly distributed values between 0 and 0.5, and 0 and 5 are added to the moment of Inertia I, μ_s and μ_c of the tyre model, and mass of the body respectively

By randomising the starting speed of the vehicle at each episode, the model learns to adapt and prevent wheel lockup at various braking speeds. Randomised samples of uniformly distributed values are added to different parameters of the single tyre model at each episode to encourage the model to adapt to varying braking conditions such as different friction conditions and different body masses. These conditions force the agent to learn the robust policies to generalize instead of simply overfitting the environment, and also allow for the model to generalize in new environments such as over rough terrain. Setting the ABS controller to make decisions at 100Hz, and training for 750

episodes, Figure 4.7 plots the mean rolling episode scores against the episode number for the five different RL algorithms.

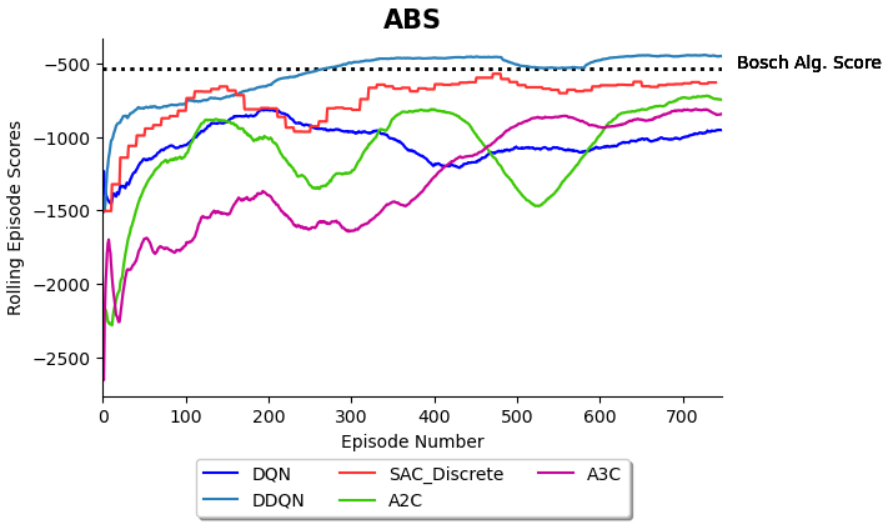


Figure 4.7: Comparison of rolling episode scores for each episode number

An average score of -510 for the Bosch algorithm was achieved over the same environment, and is plotted as a dotted horizontal line to allow for easy comparison. The ideal control algorithm is one that is capable of matching or exceeding this performance measure, and the DDQN algorithm successfully manages to do so. This indicates that it was able to achieve less penalties for each braking episode, achieving better braking performances than the Bosch algorithm. To further validate this, additional comparisons of the simulation results are made. These results are computed by simulating ten episodes of the best performing model (DDQN in this instance) against ten episodes of the Bosch algorithm under the same starting conditions (17m/s). Two performance measures are conducted based on these runs: 1) the vehicle and wheel speed profile for a single episode, and 2) the average stopping distance and longitudinal slip ratio across all ten episodes.

Figure 4.8 plots a comparison of the vehicle and wheel speed for the DDQN algorithm against the Bosch algorithm. It can be seen that wheel speed is better cycled with the DDQN model compared to the Bosch algorithm (achieving a maximum slippage of 15% compared to 80% respectively), resulting in no wheel lockup and a shorter stopping distance. Using the relationship between the braking coefficient and longitudinal slip (Figure 4.6) as a point of reference for the different slip ratios, the average stopping distances and longitudinal slip ratios are recorded in Table 4.3. Results show that the average stopping distance of the DDQN model is shorter than without ABS, while the average slippage is lower than the Bosch algorithm, with no slip exceeding 20%.

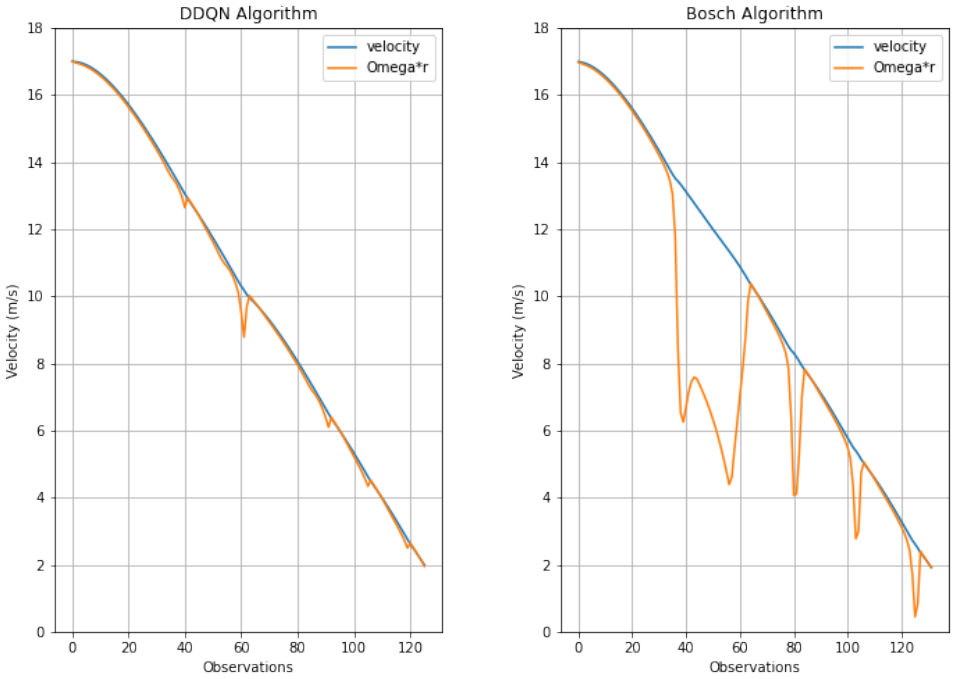


Figure 4.8: Comparison of the ABS configurations over smooth terrain

Table 4.3: Comparison of the stopping distance and longitudinal slip

Stopping Distance (m)	**BOSCH**	**DDQN**	**No ABS**
Mean	20.82	20.80	23.67
Std. dev	1.85	1.16	1.96
Best	20.32	19.84	24.85
Longitudinal Slip (%)			
Mean	8.98	1.33	70.31
Std. dev	17.28	1.91	44.98
$\lambda < 3\%$	71.43	86.92	27.86
$3\% < \lambda < 20\%$	13.64	13.08	1.36
$\lambda > 20\%$	14.93	0.0	70.79

This initial study highlights the application of a model-free intelligent control algorithm as an ABS control approach using a quarter car model under simple conditions. The RL algorithms demonstrate the ability to learn and prevent wheel lockup without compromising the stopping distance over smooth terrain, thus, an increase in the complexity of the simulation environment is proposed. To further identify the best suited algorithm for this control problem an investigation into the better of the two reward functions previously discussed along with the effect of using the

TCN as the function approximator is also proposed. The procedure taken to introduce the varying normal forces at the tyre contact patch is explained.

4.4 Normal Force Generation

In order to simulate braking over rough terrain, a varying normal force at the tyre contact patch is required. Previous studies have been conducted at the Gerotek Test Facilities' Belgian paving (Penny, 2016; Hamersma, 2017; Van der Merwe et al., 2018), see Figure 4.9, and this terrain is used as vertical excitation as it provides a rough, randomly excited road.

Figure 4.9: A test vehicle performing a brake test on Belgian Paving (Penny, 2016)

Due to the difficulty in efficiently modelling a tyre that can capture the tyre dynamics over rough terrain, it is proposed that the vertical force be obtained from a statistical representation of the expected vertical force experienced over the Belgian paving. The approach is to generate reasonably realistic vertical excitations by simulating a quarter-car model with the FTire model over the Belgian paving at varying speeds (Stallmann and Els, 2014). A 2 Degree of Freedom (DOF) quarter-car model with an additional longitudinal DOF at the tyre, see Figure 4.10, is used to generate a series of vertical force profiles during braking. These profiles are customised over a fitted range to ensure that they are evenly spaced irrespective of the initial braking speed. The tyre used is the Michelin LTX A/T2 235/85R16. This is an all-terrain tyre designed to resist chipping and tearing, especially on off-road conditions, and has an upgraded tread compound for tough off-road endurance (Michelin, 2021).

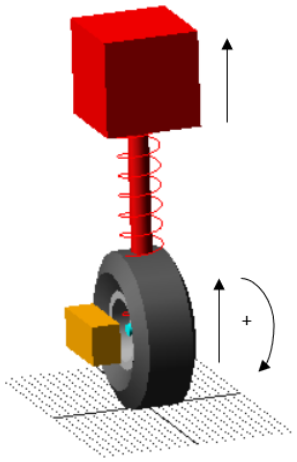


Figure 4.10: 2 DOF Quarter car model with an additional longitudinal DOF at the wheel

For better generalization of the algorithm as it trains over these vertical force profiles, stochasticity is added by simulating the quarter car model braking at different speeds (15-19m/s) and locations along the paving. These individual vertical force profiles are measured at the tyre in ADAMS and captured to allow for easy implementation into the single tyre model as well as quick solution time during simulations. They capture the speed dependency of the model and are evenly spaced, allowing for a realistic representation of the vertical forces as the vehicle begins to slow down. The vertical force profiles are used as input to the LuGre tyre model and are randomly sampled every five episodes during training. This ensures the model learns the general phenomena of braking over rough terrain, instead of a specific tyre over a specific terrain. Figure 4.11 plots different vertical force profiles experienced at the tyre contact patch.

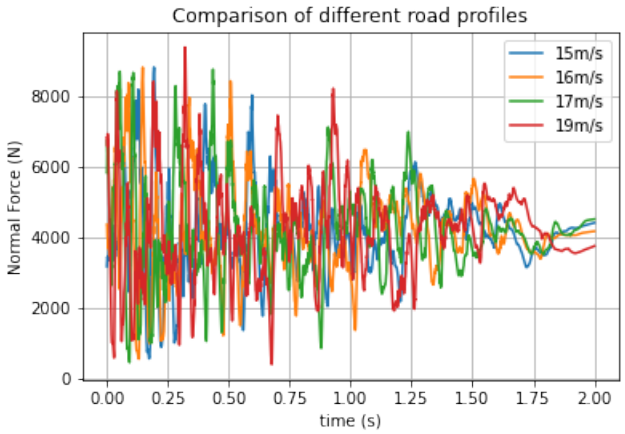


Figure 4.11: Comparison of different vertical force profiles generated over the Belgian paving

4.5 Reward Function

Two reward functions are proposed in Section 4.3.1 to identify their feasibility as the best reward function for the control algorithm. When braking over rough terrain the longitudinal acceleration based reward function is found to be sensitive to changes in the vertical force acting at the tyre contact patch. This sensitivity is demonstrated in Figure 4.12, which shows the longitudinal acceleration profile after a straight line braking simulation using the single tyre model and Bosch algorithm.

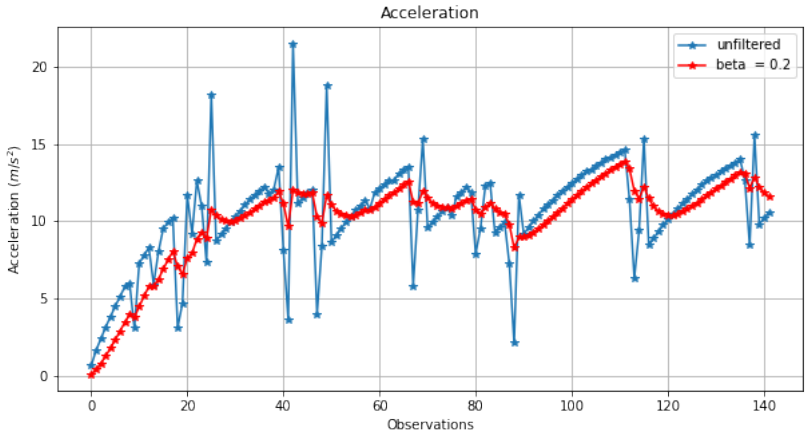


Figure 4.12: Comparison of longitudinal acceleration with and without a filter of β = 0.2 over rough terrain

This is a consequence of the large differences between the vertical forces at the tyre contact patch as the wheel brakes over the rough terrain. As a result, the notion of the reward function serving as a penalty function is compromised, with the tall spikes in the acceleration resulting in positive rewards when large enough. This is undesired during training as the intelligent model will attempt to achieve greater spikes in the longitudinal acceleration as opposed to preventing wheel lockup as the objective of ABS control. A first order filter is proposed to smoothen the acceleration signal and is defined as:

$$\begin{aligned} y(t_k)_{\text{filtered}} &= (1-\beta)y(t_{k-1})_{\text{filtered}} + \beta y(t_k), \\ \beta &= \frac{\Delta t}{T_f + \Delta t}, \end{aligned} \tag{4.10}$$

where $y(t_k)_{\text{filtered}}$ is the filtered output at the current time step, $y(t_k)$ is the unfiltered input value at the current time step, and $y(t_{k-1})$ is the filtered value from the previous time step. β is the smoothing factor between zero and unity, where a value closer to zero produces a greater filtered output, albeit with larger delays. Using β = 0.2, the filtered longitudinal acceleration is plotted on top of the unfiltered longitudinal acceleration in Figure 4.12. Despite a small delay existing in the

filtered acceleration, it captures the desired motion of the acceleration - ensuring no sensitivities to the changes in the normal force.

Figure 4.13 shows a straight line braking simulation of the Bosch algorithm braking from 17m/s on the Belgian paving. Focusing on the reward profile, the pressure, unfiltered and filtered longitudinal acceleration reward functions are plotted on top of each other. Both the pressure and filtered longitudinal acceleration reward functions are capable of achieving the desired objectives of the controller with their rewards increasing up until the slip exceeds 20%, resulting in a decrease in the rewards. The unfiltered longitudinal acceleration reward function produces peaks that approach positive rewards as the wheel travels over the terrain, making it undesirable. Thus, the pressure based and filtered longitudinal acceleration based reward functions are capable of meeting the desired control objectives and are analysed further.

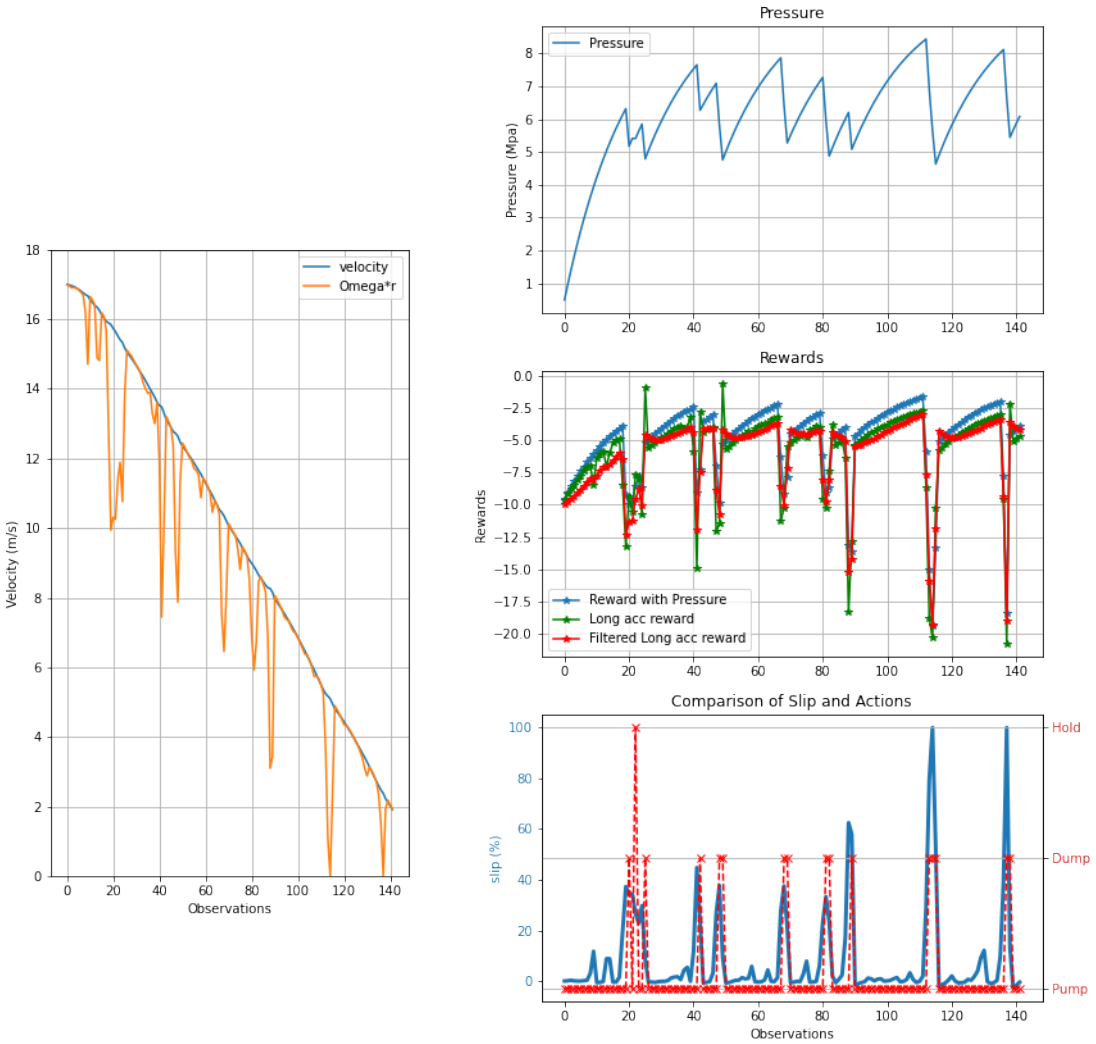


Figure 4.13: A comparison of the different reward functions using the Bosch Algorithm

4.6 Training and Results on Rough Terrain

Simulations over rough terrain are performed. Increasing the complexity of the simulation environment from smooth terrain to rough terrain has two purposes. The first is to ensure the viability of the different RL algorithms over rough terrain. If they are unable to prevent wheel lockup using a simplified quarter car model over rough terrain, they will be unable to do so on a more complex agent such as a full vehicle model. The second is to narrow down the best performing RL algorithm which will serve as the proposed ABS control algorithm in further simulations, involving a more realistic and complex agent and environment (such as a full vehicle model in ADAMS).

To reduce solution times, the single tyre model is trained doing straight-line braking using the three most consistent performing algorithms over smooth terrain: DDQN, A2C, and A3C. Using 100 previous step intervals (past data), the TCN function approximator is used in conjunction with these algorithms with the following model settings: kernel size of 5, dilation of d = 128, and filter size of k = 2. The algorithms are trained separately using the pressure and filtered longitudinal acceleration reward functions respectively. As with training over smooth terrain, the following measures are taken to ensure that the agents are learning generalization and not simply memorizing the specifics of the environments:

1. The starting speeds of the vehicle are randomised between 15 and 19m/s
2. Randomised samples of uniformly distributed values between 0 and 0.5, and 0 and 5 are added to the moment of Inertia I, μ_s and μ_c of the tyre model, and mass of the body respectively
3. Randomised vertical force profiles are randomly sampled every 5 episodes

These steps force the agent to learn the general phenomena of braking over rough terrain, in addition to allowing the model to generalize in more complex environments. Setting the ABS controller to make decisions at 100Hz and training for 500 episodes, Figure 4.14 plots the performance of the different RL algorithms using the pressure and filtered longitudinal acceleration reward function respectively. An average score of -680 was achieved for the Bosch algorithm and is plotted as a dotted horizontal line to allow for easy comparison to the intelligent algorithms.

Analysing Figure 4.14 and focusing on the DDQN algorithm; the pressure based reward function achieves fewer penalties per episode than the longitudinal acceleration based reward function, and can therefore exceed the performance of the Bosch algorithm. Both the A3C and A2C algorithms demonstrate unsuccessful learning and run-to-run instability.

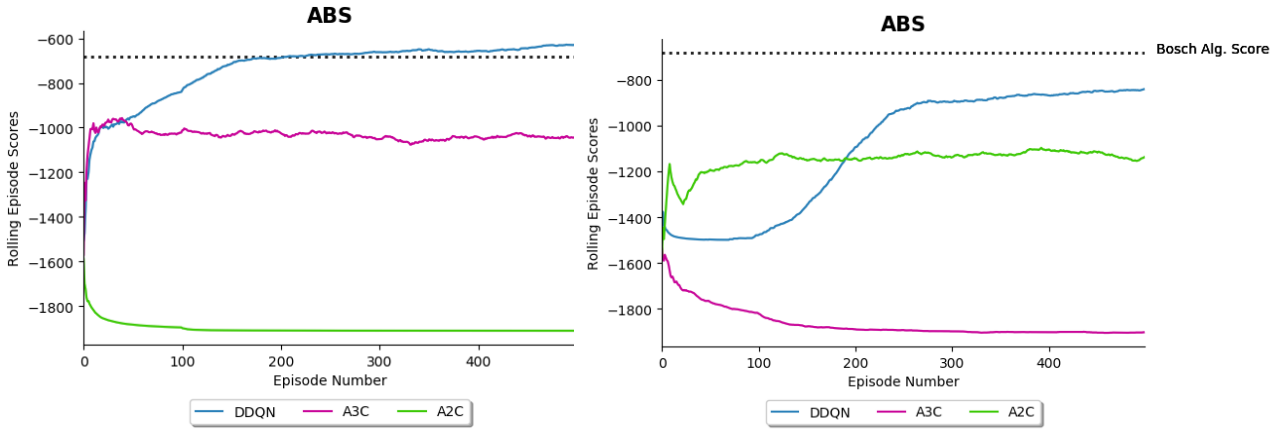


Figure 4.14: A comparison of the DDQN, A2C, and A3C, with TCN as the function approximator, and the pressure (left) versus filtered longitudinal acceleration (right) as the reward function over rough terrain

The advantage of using the TCN as the function approximator is highlighted in Table 4.4 which records the average stopping distances and longitudinal slip percentages of the Bosch, DDQN with and without a TCN algorithm (labelled DDQN-TCN and DDQN-NN), and braking with no ABS. The TCN in conjunction with the DDQN algorithm provides an additional aspect of predictive modelling to the ABS controller through storing the wheel speed and vehicle speed over a period of time, which is advantageous in unknown complex environments.

Table 4.4: Summary of different braking performance measures over rough terrain

Stopping Distance (m)	**BOSCH**	**DDQN-NN**	**DDQN-TCN**	**No ABS**
Mean	24.32	26.14	23.65	22.17
Std. dev	1.69	2.10	1.98	2.13
Best	22.45	25.74	22.17	20.24
Longitudinal slip (%)				
Mean	9.53	3.84	5.56	64.61
Std. dev	19.64	10.87	15.97	42.42
$\lambda < 3\%$	66.47	77.91	75.94	18.81
$3\% < \lambda < 20\%$	16.76	15.73	17.19	8.90
$\lambda > 20\%$	16.77	6.36	6.88	72.29

Analysing this summary, braking without ABS yields the shortest stopping distance at the cost of

maintaining the slip in the desired region, with an average wheel slippage of 65%. As expected, the ABS algorithms achieve lower average slippage with the DDQN algorithms (with and without TCN) producing the least percentage of slip values greater than 20%, highlighting their ability to prevent wheel lockup on rough terrain. Theoretically this results in better lateral control of the vehicle. The DDQN-TCN algorithm ensures the best wheel slippage control while maintaining an average stopping distance within 1m of braking without ABS.

Comparisons between the vehicle and wheel speed for the DDQN-TCN algorithm and Bosch algorithm respectively are presented in Figure 4.15. The DDQN-TCN algorithm cycles the wheel speed better than the Bosch algorithm within the same amount of step intervals (observations), demonstrating that better wheel slip control is achieved within the same stopping distance.

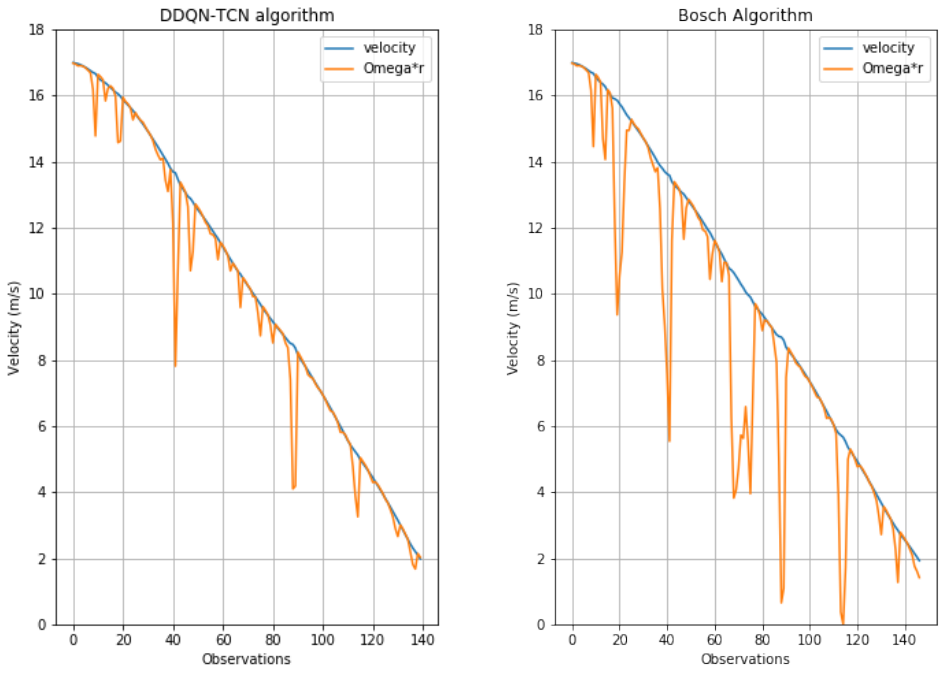


Figure 4.15: Vehicle and wheel speed comparison of each ABS algorithm over the Belgian Paving

This lower wheel slippage is confirmed in Figure 4.16, where the DDQN-TCN algorithm cycles through the three different brake states more frequently, thus achieving maximum wheel slippage of 50% compared to 100% wheel lockup for the Bosch algorithm. Due to the consistency and ability of the DDQN-TCN model to prevent wheel lockup over rough terrain without deteriorating the stopping distance, it is proposed that the DDQN-TCN algorithm and pressure-based reward function be used as the ABS control algorithm.

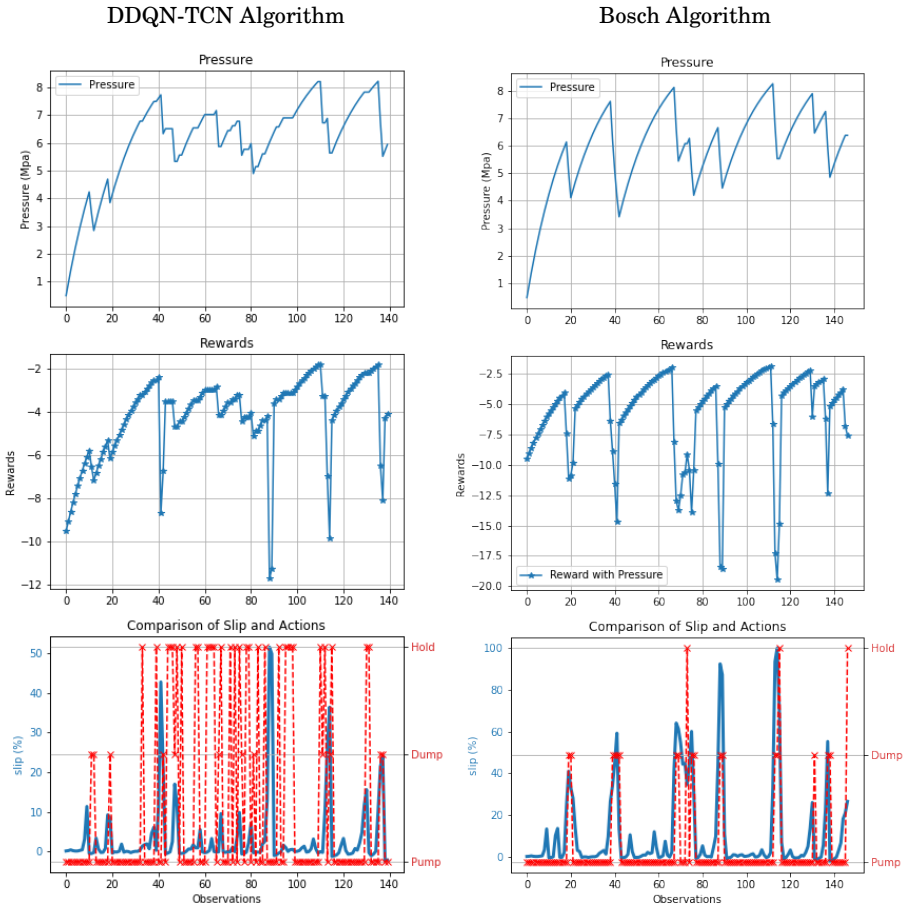


Figure 4.16: A comparison of the pressure, rewards and the slip and actions chosen

4.7 Conclusion

A simplified quarter-vehicle model is introduced to reduce solution times and determine the initial aspects of the ABS controller including which RL algorithm to use, the best suited reward function, and the effect of the TCN function approximator. Training and straight line braking simulations over smooth and rough terrain (modelled after the Belgian paving) are performed, and comparisons are drawn against a Bosch algorithm tuned for rough terrain.

The DDQN algorithm, in conjunction with the TCN function approximator (labelled DDQN-TCN) and pressure-based reward function, demonstrates the ability to prevent wheel lockup over rough terrain, without significantly deteriorating the stopping distance. It is proposed as the ABS control algorithm to be used in the final tests. Chapter 5 looks at increasing the complexity of the model from a single tyre system to a validated full-vehicle simulation model using the FTire tyre model in ADAMS. Additional tests according to the requirements of SAE (2014) are performed, with the performance evaluated using the ABS performance evaluation metric proposed by Hamersma (2017).

Chapter 5: Full Vehicle Simulation

This chapter increases the complexity of the environment from a single tyre braking system to a full vehicle simulation model. Using the DDQN-TCN algorithm, training and simulations are performed over the Belgian paving using the FTire model in ADAMS. Comparisons are made to a Bosch algorithm tuned for rough terrain, and a conventional braking system (no ABS). A summary of the results concludes the chapter.

5.1 Full Vehicle Simulation Model

To further investigate the performance of the proposed model-free controller, a fully nonlinear multi-body dynamics model, developed in ADAMS, is used as the basis of this study. The vehicle was modelled after the 1997 Land Rover Defender by Thoresson (2007), and improved by Uys et al. (2007) and Cronje and Els (2010). The model includes non-linear spring and damper characteristics, non-linear bump stops and bushes, and body torsion about the longitudinal axis allowing for better model accuracy. All mass properties were also experimentally determined. The parameters influencing the longitudinal response i.e. the braking and aerodynamic drag, were modelled by Hamersma (2013), and the hydro-pneumatic suspension unit is modelled as an adiabatic process, thus considering the thermal effects inside the system, assuming no heat transfer to and from the surrounding environment. An isometric representation of the full vehicle model is provided in Figure 5.1.

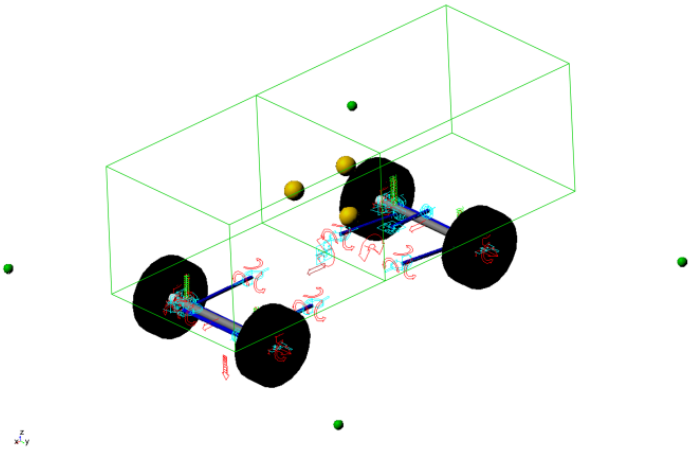


Figure 5.1: Isometric view of the modelled simulation vehicle in ADAMS (Botha, 2011)

The model is run in co-simulation with MATLAB/Simulink and consists of 15 unconstrained degrees of freedom and 16 moving parts as summarised in Table 5.1.

Table 5.1: ADAMS vehicle model's unconstrained degrees of freedom (Thoresson et al., 2014)

Body	**Degrees of Freedom**	**Associated Motion**
Vehicle body (2 rigid bodies)	7	Body torsion Longitudinal, lateral, vertical roll, pitch, yaw
Front axle	2	Roll,vertical
Rear axle	2	Roll,vertical
Wheels	4x1	Rotation

The model consists of many subsystems of which the following are briefly explored:

1. Hydro-pneumatic suspension struts
2. Brake system and ABS algorithm
3. Path-following driver model

The suspension system consists of hydro-pneumatic struts, modelled by Van der Westhuizen and Els (2015), and dampers modelled through the use of a lookup table populated with experimental data by Thoresson (2007). Both spring and damper settings allow for two discrete volume settings (high or low) depending on the application, thus a combination of setups is possible e.g., "ride comfort mode" which makes use of soft springs and low damping. The spring and damping values, in conjunction with the relative suspension displacement and velocity, are used as input for the ADAMS model to calculate the corresponding strut force as output i.e. the sum of spring and damping forces.

The longitudinal demand forces, including the brake system, were modelled by Hamersma (2013), and used extensively by Penny (2016) and Van der Merwe et al. (2018) to investigate the ABS braking performance of the vehicle on rough terrain.

Path-following of the vehicle is accomplished by a driver model developed by Thoresson et al. (2014), and refined by Botha (2011). By defining the desired path as a series of x-y coordinates, from which a path yaw angle is determined, two separate controllers are used to follow the prescribed path. This allows for multiple types of braking manoeuvres to be performed such as straight line braking and braking-in-turn.

The vehicle is fitted with Michelin LTX A/T2 235/85R16 tyres, previously used to generate the vertical force profiles. The FTire model is used as the tyre model in ADAMS. Simulations conducted by Hamersma (2017) have validated the ability of the FTire model to accurately model the force generation characteristics of the Michelin LTX A/T2 235/85R16 tyre on smooth and undulating road surfaces, as seen by the longitudinal force validation in Figure 5.2. The desired slip range for optimal control is between 5-20%, with the sparsity of the points after 20% indicating instantaneous wheel lockup. An important aspect to notice is the FTire model has a flat response at higher slippage compared to experimentally, which decreases at higher slippage. This implies that there is very little difference in the friction coefficient experienced between the peak friction and 100% lockup using the FTire model, which will influence the simulated results.

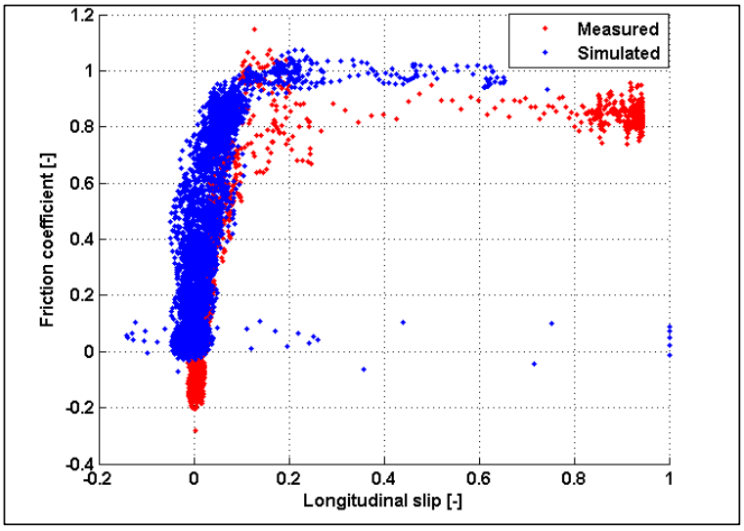


Figure 5.2: Longitudinal tyre force experimental validation (Hamersma, 2017)

The full vehicle simulation model has been validated for vertical dynamics (Els et al., 2007), lateral dynamics (Botha, 2011), and longitudinal dynamics (Hamersma, 2013), highlighting its complexity and reasonable representation of real-world simulations. Numerous tests have been conducted by the Vehicle Dynamics Group (VDG) at the University of Pretoria using this fully validated vehicle model, including ABS tests conducted by (Penny, 2016; Hamersma, 2017; Van der Merwe et al., 2018), in which the simulated results correlate well against vehicle test results, confirming this reasonable real-world representation. It is assumed that this representation will hold for the simulations in this study.

5.1.1 Python-Simulink Interface

With the vehicle model developed in Simulink/MATLAB, and the ABS algorithm written in Python due to the variety of programming frameworks available, the problem of communication between

these two programming languages needs to be solved. The approach of message passing is used through a User Datagram Protocol (UDP). This is a simple, connectionless, internet protocol in which error-checking and recovery services are not required. Not only does this make it comparatively faster than alternative connections such as the Transmission Control Protocol (TCP), it is more efficient, requires no overhead for opening, maintaining, or terminating a connection, and is largely preferred for real-time communications (Cook, 2017). In this study the vehicle model in MATLAB/Simulink acts as the server and the ABS model in Python as the client.

5.2 Training and Results

Training of the DDQN-TCN algorithm as the ABS control approach on the full vehicle simulation model is achieved by simulating straight line braking over the Belgian paving. As with the single tyre model training, the following measures are taken to ensure the generality of the agents, with the benefits of doing so explained after.

1. The braking speed of the vehicle is randomised between 50km/h and 80km/h
2. Randomised samples of uniformly distributed values between 0 and 200 are added to the drive force which brakes the vehicle
3. The front and rear spring, and damper suspension settings are randomised according to their volume settings (high or low)

By randomising the braking speeds and drive force of the vehicle, not only is the vehicle braking at different speeds but also at different locations along the Belgian paving, forcing the agent to adapt and learn the general phenomena of braking over rough terrain. Multiple combinations of suspension settings are possible and they directly impact the performance of an ABS braking over rough terrain (Hamersma and Els, 2014), thus randomising these setups forces the agent to generalize braking over rough terrain regardless of the suspension settings.

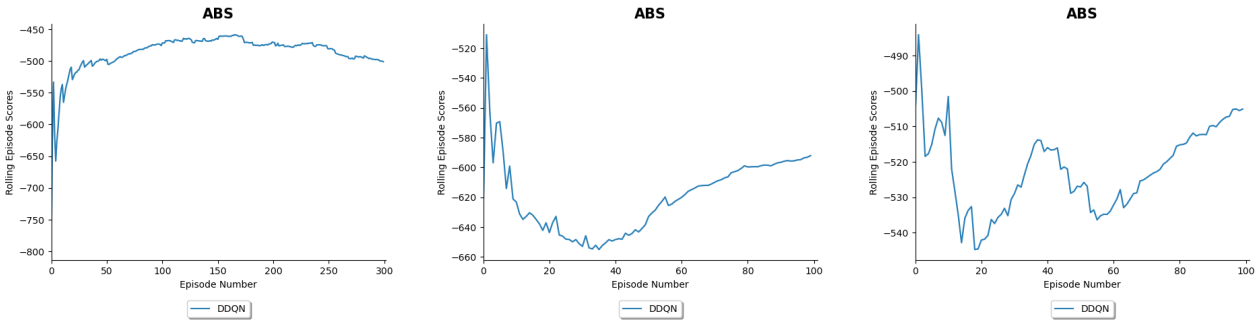
Two separate training setups are investigated: with transfer learning, and without. Transfer learning is introduced in Section 3.5, and involves the passage of knowledge from one task to another. The best performing model trained in the single wheel simulations (500 episodes) is saved and trained further here such that this creates continual learning for the intelligent ABS agent, aiding it to achieve its desired objectives within a more complex agent and environment, without forgetting the knowledge it learnt from the simplified single wheel model and environment.

Upon training with transfer learning for 800 episodes (500 transferred, 300 new episodes), it was discovered that the DDQN-TCN model suffered from a phenomenon known as catastrophic

forgetting (Kirkpatrick et al., 2017), see Figure 5.3a, where the performance decreases after a certain maximum. This is a common occurrence with deep NNs that make use of continual learning, and involves the agent forgetting previously learnt tasks/information as it attempts to incorporate information relevant to the current task, impairing its performance. Despite recent advances on deep NNs, little progress has been made to achieve continual learning. Parisi et al. (2019) identified three main mechanisms for mitigating forgetting in NNs:

1. Replay of previous knowledge
2. Regularization mechanisms to constrain parameter updates
3. Expanding the network as more data becomes available

Kirkpatrick et al. (2017) propose the use of Elastic Weight Consolidation (EWC) to avoid the catastrophic forgetting of old abilities by selectively slowing down learning on the weights of previously learnt tasks. Hayes et al. (2020) propose the use of experience replay through a method called REMIND, in which they continuously remind the NN of previously learnt tasks.



(a) (b) (c)

Figure 5.3: Training across three different experiments with specified settings. (a) After 800 episodes. (b) After 600 episodes and ϵ = 0.5. (c) After 600 episodes and α = 0.001

In this study the model demonstrated great learning capabilities and achieved low penalties in the initial stages of learning before suffering from catastrophic forgetting. The approach of hyperparameter-tuning and early stopping is proposed to overcome the catastrophic forgetting. Two main hyperparameters are targeted: the learning rate α, which controls how fast the model learns, and epsilon ϵ, which introduces randomness into the algorithm and is related to the epsilon-greedy action selection procedure of the Q-learning algorithm.

Despite the advantage of being the quickest solution to catastrophic forgetting, this approach does not always guarantee a solution and may be time consuming as a result of the infinite number of

changes that can be made, often unsuccessfully, to the hyperparameters as seen in Figures 5.3b-5.3c. In both cases the learning rate and epsilon value were decreased from the initial conditions in Appendix A respectively. It is recommended for future transfer learning training that an alternative method be used to overcome the catastrophic forgetting, such as experience replay through REMIND, which may allow for longer training episodes and a superior performing model.

Using a reduced learning rate of 0.005 (from 0.01) and early stopping, the best transfer learning model was achieved after 100 episodes (and 500 transferred) and is plotted in Figure 5.4. This model is directly compared to a model trained without transfer learning using the same hyperparameters, where the advantage of using transfer learning is visually compelling. Simulations using the final DDQN-TCN algorithm are achieved by saving the model as a traced model and calling it as a MATLAB executable (MEX) file in MATLAB. This allows for the simulations to be based solely in MATLAB/Simulink.

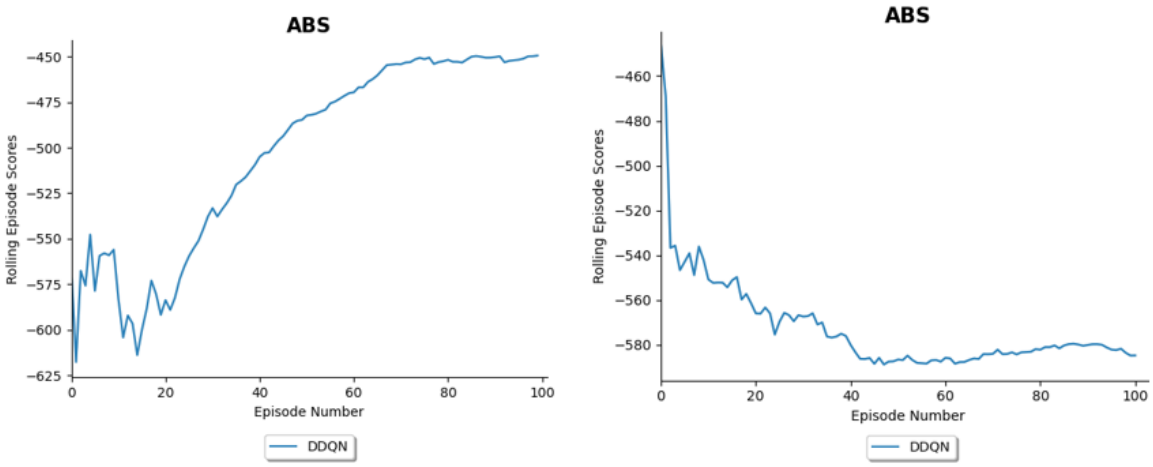


Figure 5.4: Training results achieved with transfer learning (left) vs. without (right)

Eight braking scenarios are simulated according to the SAE (2014) guidelines. These include:

1. Straight line braking:
   (a) Belgian paving
   (b) Class D road (Alternative rough terrain)
   (c) High friction surface
   (d) Low friction surface
2. Braking in a turn:
   (a) Class D road (Alternative rough terrain)
   (b) High friction surface
   (c) Low friction surface
3. Split friction coefficient braking

These simulations are repeated for three different vehicle configurations: 1) DDQN-TCN algorithm, 2) Bosch algorithm, and 3) conventional braking system (no ABS). Five iterations of each braking manoeuvre are simulated from an initial speed of 55km/h. The following suspension configuration is used: hard spring and damper settings in the front, and soft spring and damper settings in the rear, as recommended by (Hamersma and Els, 2014). The high and low friction surfaces have a friction coefficient of 1 and 0.4 respectively. The high friction surface and class D road braking-in-turn simulations have a constant radius of 70m, and the low friction surface has a constant radius of 120m. The split-mu surface has a coefficient of friction of 1 under the left side, and 0.4 under the right side of the vehicle. All of the simulations are terminated when the vehicle speed is below 15km/h, as typical ABS algorithms stop working at low speeds (Hamersma, 2017).

5.2.1 Straight Line Braking

Straight line braking simulations are performed to investigate the ability of the DDQN-TCN algorithm to prevent wheel lockup and maintain directional control (brake in a straight line).

5.2.1.1 Belgian Paving

Figure 5.5 shows a comparison of the different wheel speed profiles. The DDQN-TCN algorithm is able to prevent wheel lockup, achieving lower slippage than the Bosch algorithm and no ABS.

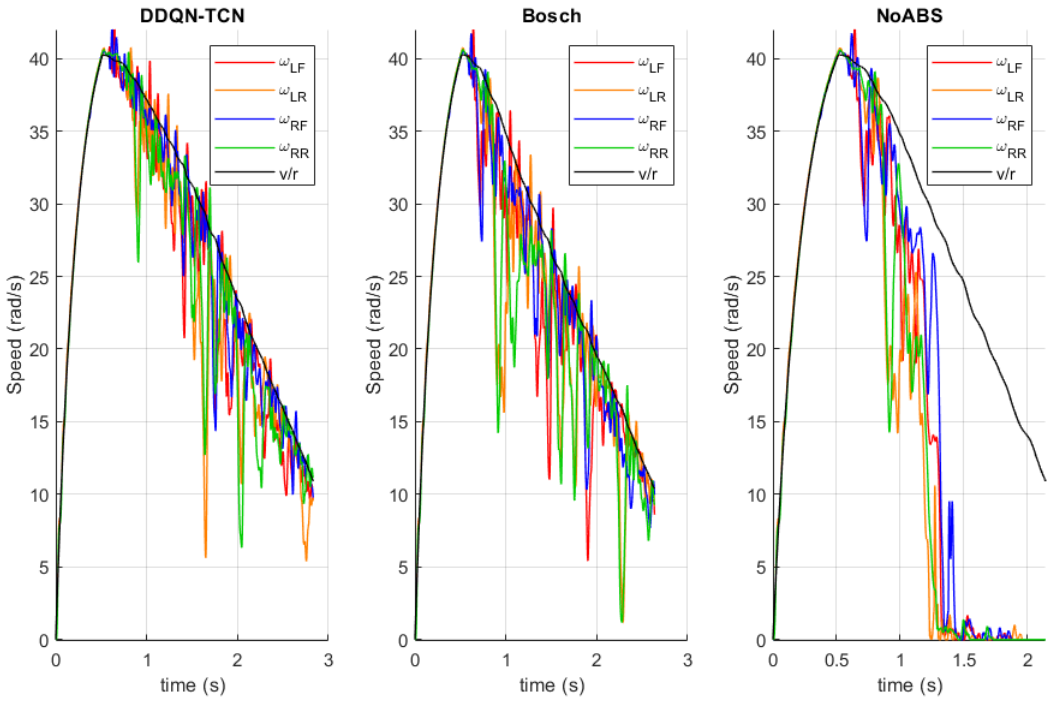


Figure 5.5: Wheel speed comparison of the different vehicle configurations over the Belgian paving

This is confirmed by the slip profiles in Figure 5.6, where the path following error and yaw rate error is also provided.

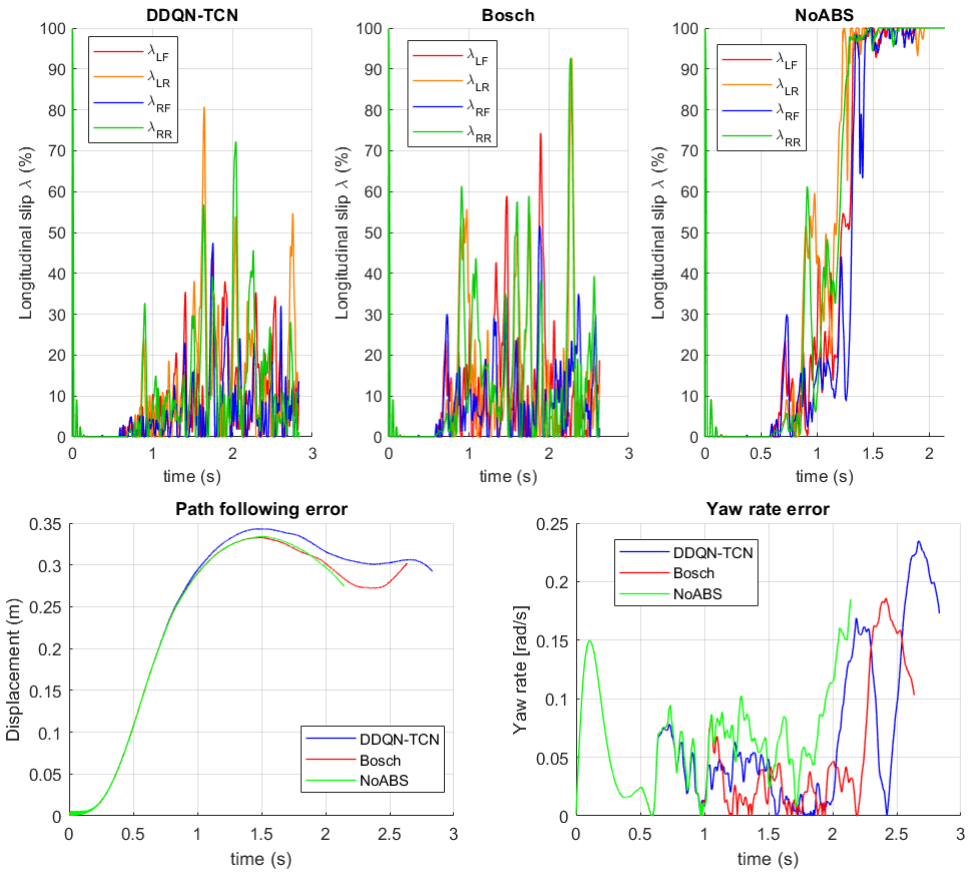


Figure 5.6: Longitudinal slip percentage, path following error, and yaw rate error comparison over the Belgian paving

The lateral path errors are small and similar for the different configurations, as one would expect in a straight line braking simulation. The yaw rate deviated slightly from the desired yaw rate, however the errors are similar. Figures 5.5 and 5.6 provide an insight into the performance of the different braking configurations over the Belgian paving. It is difficult to quickly compare the average performance of each configuration to identify the best braking configuration, hence the ABS performance evaluation technique proposed by Hamersma (2017), introduced in Section 2.2.2.1, is used to directly evaluate the performance of the different braking configurations.

This metric takes into account the following four evaluation metrics:

1. Mean and standard deviation of stopping distance
2. Mean and standard deviation of average vehicle deceleration
3. Mean and standard deviation of absolute vehicle yaw rate error
4. Mean and standard deviation of absolute lateral path following error

Plotting the mean and standard deviation of each metric in a radar-like plot gives an indication of the performance of the brake system with regard to that metric, and of the system's stability (or repeatability) respectively. To support this evaluation technique, a bar graph of the percentage wheel slippage in certain slip ranges is presented. This grants valuable insight into the amount of wheel lockup that occurs, and at what slip ratio the tyre operates in most of the time. From the tyre longitudinal slip relationship it is also possible to determine the longitudinal force produced. Figure 5.7 compares the performance of the different braking configurations using this proposed evaluation technique.

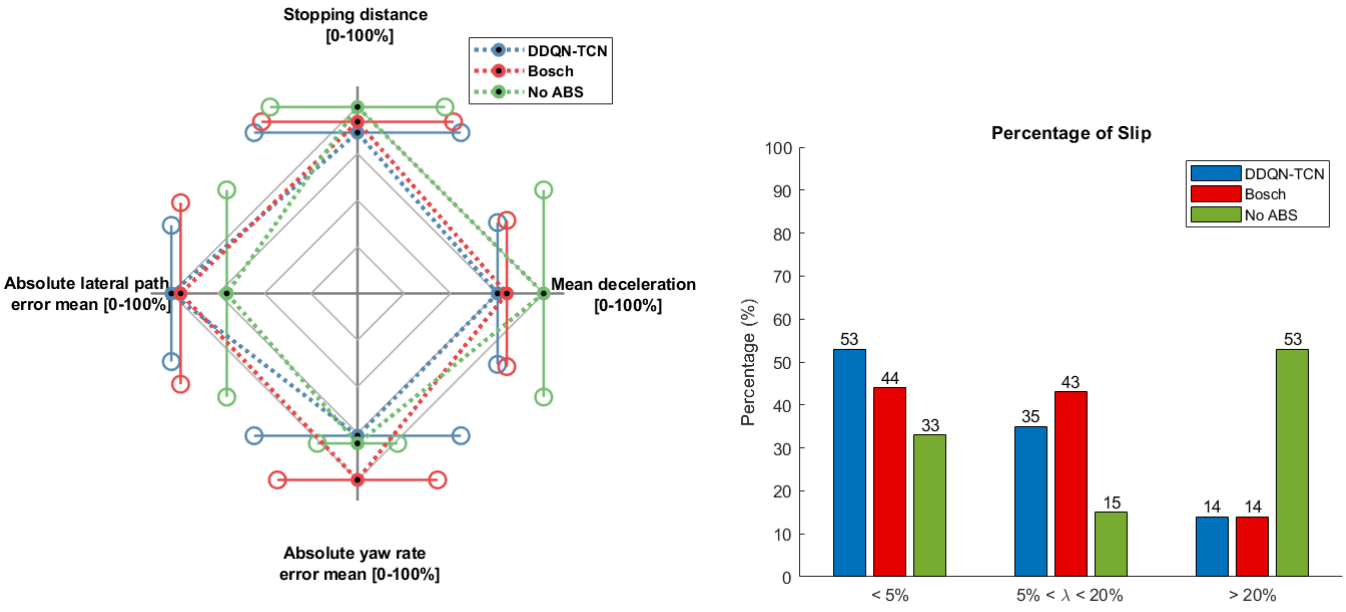


Figure 5.7: Comparison of ABS performance over Belgian paving

It can be seen that the conventional braking system results in the best stopping distance and highest mean deceleration. This is owed to the flat response of the FTire longitudinal slip relationship as wheel lockup occurs, causing no loss of braking from 20-100%, ergo lockup is not penalised. Despite this, both ABS algorithms generate better directional stability of the vehicle as expected. The two ABS algorithms are almost equivalent in vehicle stability, with the DDQN-TCN achieving greater lateral control, while the Bosch achieves a lower absolute yaw rate mean error. The bar graph further highlights their similarity in performance, with the same amount of time spent in slip regions greater than 20%. The performance of the DDQN-TCN algorithm highlights its ability to prevent wheel lockup over rough terrain without significantly compromising the stopping distance (18% further than braking without ABS).

Despite no numeric indication of how close the scaling of each metric is, the advantage of this evaluation technique is recognised by its ability to provide a quick and clear way of comparing different braking configurations. As such it is used as the evaluation technique for the remaining simulations, unless otherwise specified.

5.2.1.2 Class D Road

To investigate the generality of the performance of DDQN-TCN algorithm over rough terrain, an alternative rough terrain that it was not trained on is required. Multiple different classes of roads exist, with the Belgian paving being classified as an ISO 8608 Class D road (ISO, 2016), thus a randomly generated class D road is proposed as the alternative rough terrain. Figure 5.8 plots a comparison of the ABS performance over this surface.

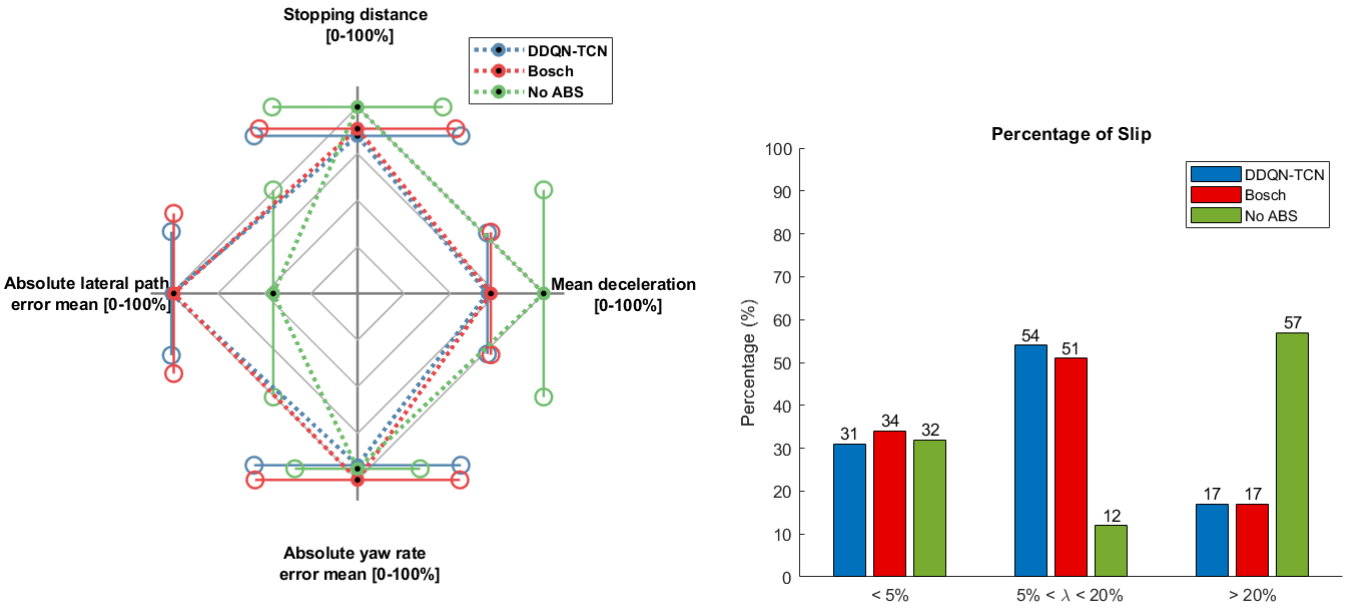


Figure 5.8: Comparison of ABS performance over class D road

As with the Belgian paving simulation, the conventional braking system results in the best stopping distance and highest mean deceleration, while both ABS algorithms generate greater directional stability of the vehicle, as expected. There are few differences between the two ABS algorithms with regard to better vehicle stability, with the bar graph highlighting that both ABS algorithms spend majority of their time between the desired region of 5-20%. The generality of the DDQN-TCN algorithm is highlighted by its ability to prevent wheel lockup and ensure directional stability over previously unencountered rough terrain.

5.2.1.3 High Friction Flat Road

To further investigate the generality of the DDQN-TCN algorithm, straight line braking over a high friction (μ = 1) flat road is conducted. A summary of the performance of the different braking configurations is provided in Figure 5.9.

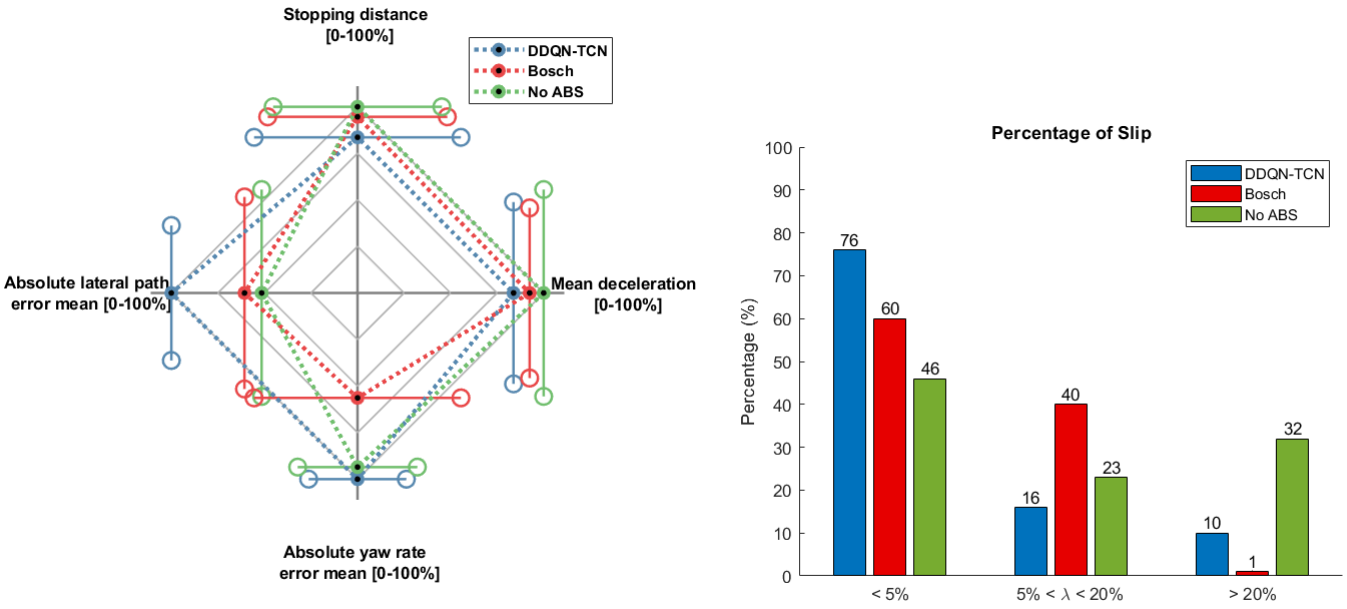


Figure 5.9: Comparison of ABS performance over high friction (μ = 1) flat road

Analysing Figure 5.9, the top performing configuration is the DDQN-TCN algorithm, achieving the best performance in 2 of the metrics while being close in the remaining two. This emphasises its generality in being able to prevent wheel lockup over smooth terrain without significantly compromising the stopping distance (12% further than braking without ABS). However, 76% of the wheel slippage occurs in the region less than 5%, indicating the algorithm is not braking to its full potential over this smooth terrain. By braking harder, a shorter stopping distance can be achieved, in addition to greater lateral control, when the longitudinal slip is cycled in the desired region. One solution is to train the algorithm over smooth terrain. This can be achieved by transfer learning or randomly sampling the smooth terrain with the rough terrain, allowing for training over multiple terrains.

Despite not achieving as much slip within the 5-20% region when compared to the Bosch algorithm, the DDQN-TCN algorithm was able to achieve greater directional stability of the vehicle, indicating that the Bosch algorithm utilised the peak longitudinal friction more optimally. This is highlighted by its reduced mean stopping distance and greater mean deceleration, however to the detriment

of the directional stability of the vehicle. The conventional braking system yields a good yaw rate mean error although a very poor lateral path mean error, indicating wheel lock up with the vehicle skidding to a halt, significantly reducing the steerability.

5.2.1.4 Low Friction Flat Road

Straight line braking over a low friction (μ = 0.4) flat road is conducted, and the performance of the different braking configurations are summarised in Figure 5.10.

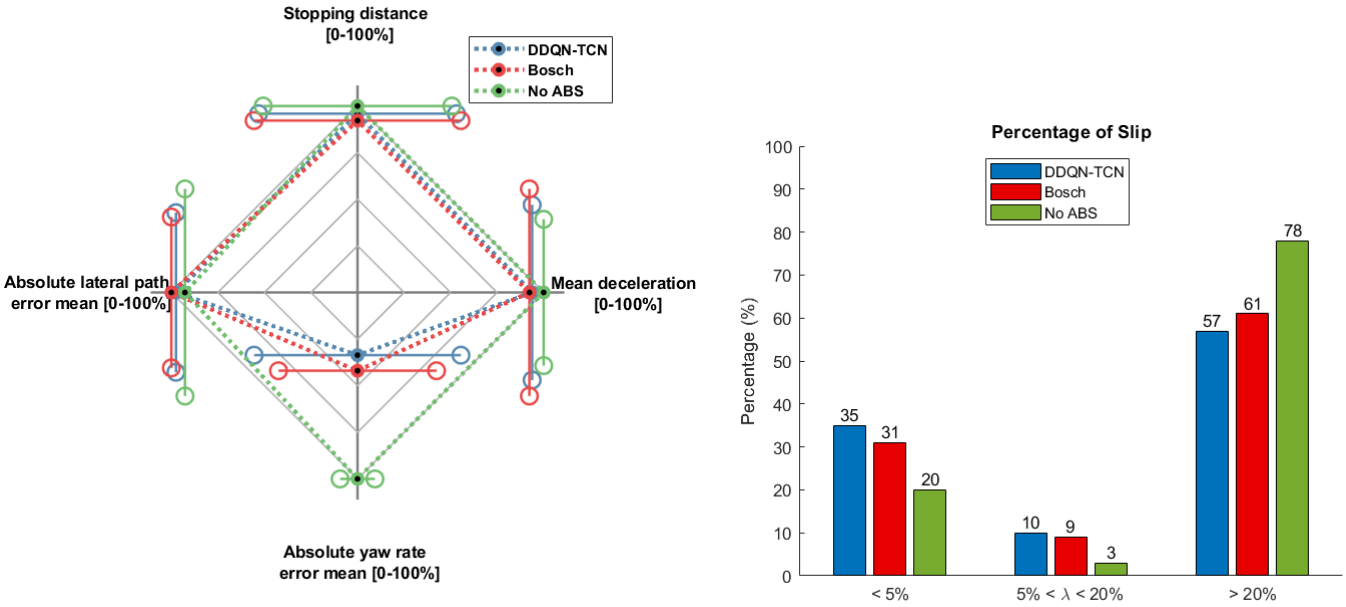


Figure 5.10: Comparison of ABS performance over low friction (μ = 0.4) flat road

Analysing Figure 5.10, a limitation of the DDQN-TCN and Bosch algorithm is identified - both ABS algorithms are unable to prevent wheel lockup over a low friction flat surface. Consequently, the best braking configuration for low friction terrain is the conventional braking system (braking without ABS). Despite no numeric indication, lack of directional control of the vehicle is shown by the scaling of all three configurations in the lateral path error and absolute yaw rate error, and the amount of time spent in the slip region greater than 20%. To overcome this limitation two possible solutions are explored:

1. Train the DDQN-TCN algorithm over low friction smooth terrain
2. Train using a modern ABS modulator with decreased brake delays

The first solution involves incorporating low friction smooth terrain in the training of the algorithm, and presents opportunities such as training over multiple terrains, as well as training the algorithm

with a road classifier.

Penny (2016) found that at least a half-second delay on the modulator, which is based off the four channel Wabco model 478 407 022 0 and is an add-on modulator first introduced on the Land Rover Defender in 1999 (WABCO, 2003), was needed to accommodate for the delay in the measurement and solenoid response as well as mechanical relay in the brake line for the Bosch algorithm. As a result, the ABS modulator is unable to cycle the wheel speed fast enough on a low friction surface where wheel lockup occurs almost immediately. Additionally, a general argument can be made that the Bosch and DDQN-TCN algorithms have similar results due to these modulator limitations. Thus, a faster, modern modulator that can mitigate these delays and result in better braking efficiency is recommended, after which re-training of the algorithm can be completed and simulations re-run.

The effect of decreasing the brake delay is investigated using the the single tyre model mentioned in Section 4.2. The delay τ constants are decreased from 200ms for pump and 50ms to dump, to 10ms each. Figure 5.11 plots a direct comparison against the best model achieved in Figure 5.4. The effect of decreasing the brake delays results in a decrease in penalties received for the DDQN-TCN algorithm, improving the braking performance. This serves as evidence that better brake efficiency and steering controllability can be obtained with a decrease in the brake delay.

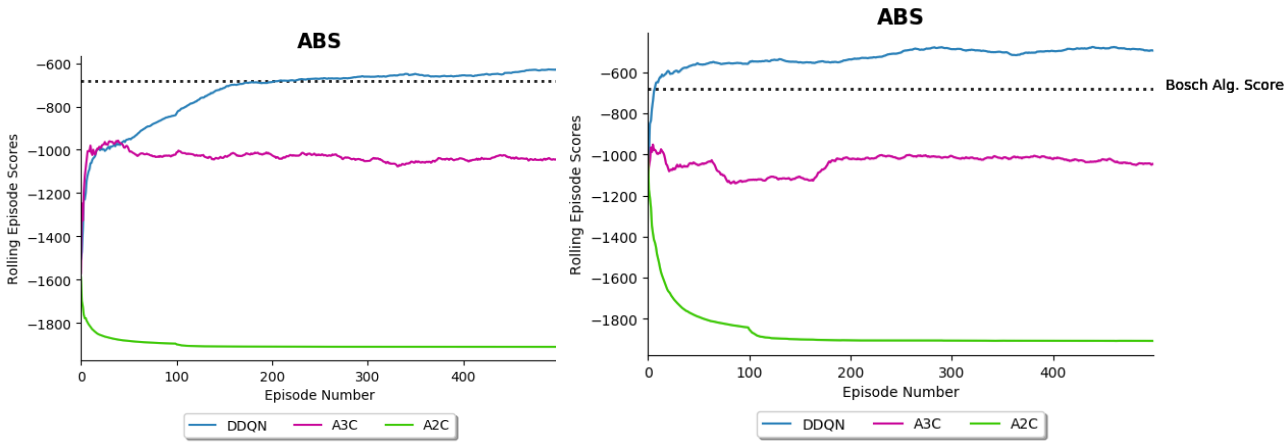


Figure 5.11: Comparison of the best final model with original brake delays (left) vs. with decreased brake delays (right) over the Belgian paving

5.2.2 Braking in Turn

Braking-in-turn simulations are performed to investigate the generality of the DDQN-TCN algorithm to prevent wheel lockup and maintain directional control whilst braking on a curve.

5.2.2.1 Class D Road

Braking-in-turn on a class D road with a constant radius of 70m is presented in Figure 5.12.

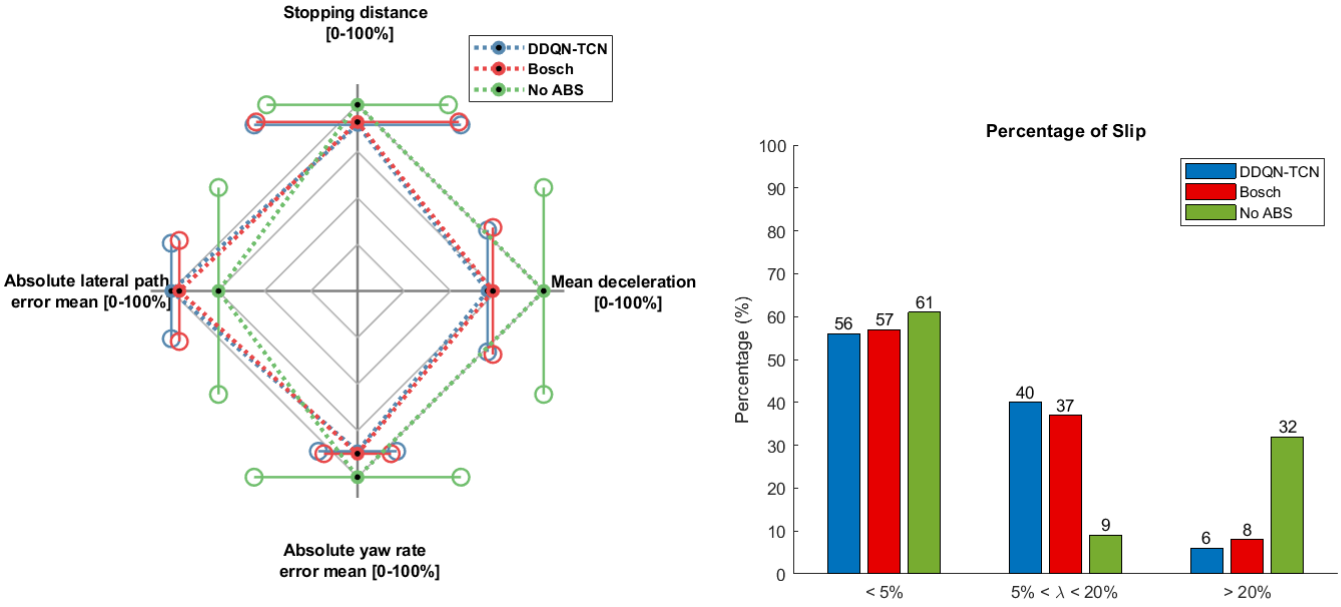


Figure 5.12: Comparison of ABS performance for braking-in-turn on class D road

Analysing Figure 5.12 it appears that the conventional braking system achieves the greatest performance when braking-in-turn over the class D road by being the best in 3 out of 4 metrics. This is not reflected by the comparison of the wheel speed and Center of Gravity (CG) of the vehicle in Figures 5.13 and 5.14 respectively. Figure 5.14 highlights the vehicle's loss of directional stability with the conventional braking system as it continues travelling straight due to its wheels being locked, while the two ABS algorithms are able to return to the desired radius path. This also explains the good absolute yaw rate error mean achieved, as the vehicle stopped earlier, resulting in a shorter error. A recommendation is that a higher braking speed be used (e.g. 65km/h instead of 55km/h) and/or a smaller turning radius, which would allow for bigger differences in the performance metrics such as the absolute lateral path following mean error and yaw rate mean error.

As with the straight line braking simulations, the conventional braking system achieves the best stopping distance and mean deceleration. This is due to the flat response of the FTire longitudinal slip relationship at higher wheel slippage, causing no loss of braking from 20-100% slip and consequently wheel lockup not being penalised. Focusing on the DDQN-TCN algorithm, it is able to prevent wheel lockup when braking-in-turn over rough terrain, slightly exceeding the performance

of the Bosch algorithm, without significantly compromising the stopping distance (13% further than braking without ABS).

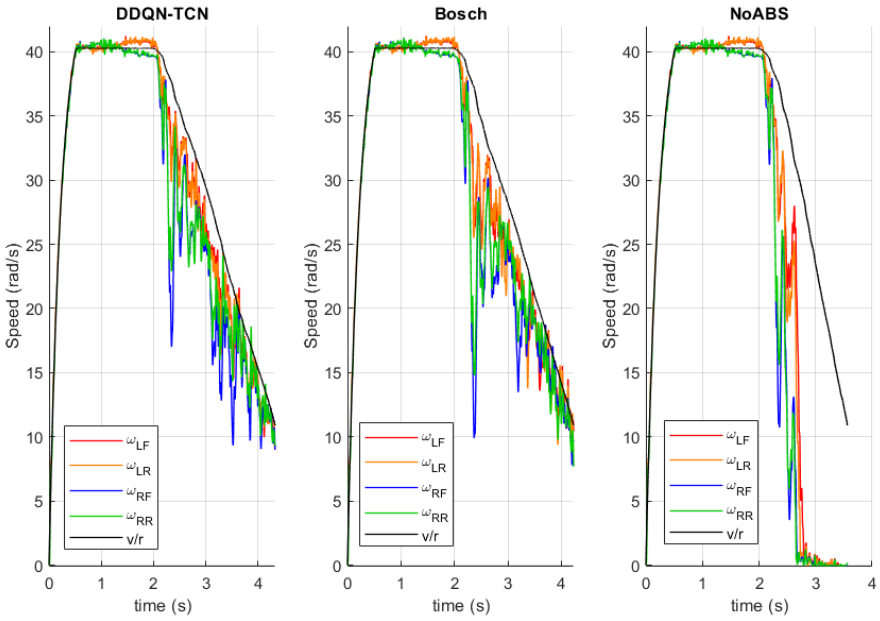


Figure 5.13: Wheel speed comparison of during braking-in-turn on class D road

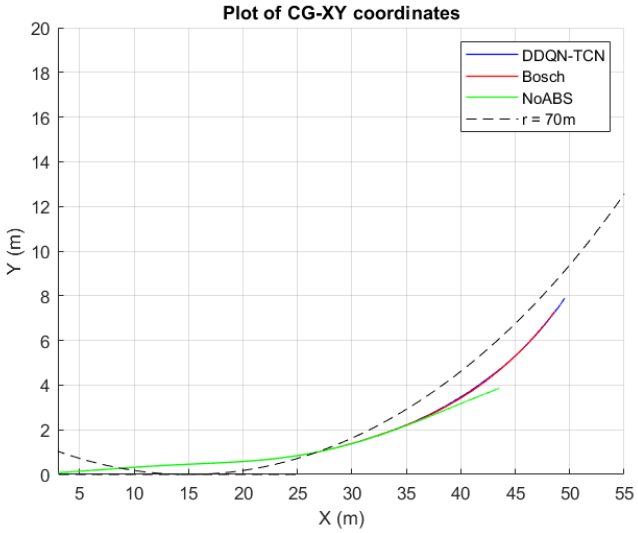


Figure 5.14: Comparison of the CG-XY during braking-in-turn on class D road

5.2.2.2 High Friction Flat Road

Braking-in-turn on a high friction (μ = 1) road with a constant radius of 70m is presented in Figure 5.15.

The advantage of having an ABS equipped vehicle is demonstrated by the lack of significant variation in the stopping distances and mean deceleration as compared to the directional stability results. Similarly to braking in a straight line over the high friction flat road, the DDQN-TCN algorithm is able prevent wheel lockup and achieve the best lateral control of the vehicle, however 92% of the wheel slippage occurs in the region less than 5% indicating the algorithm is not braking in the optimal slip ratio i.e. not as hard as it could. A solution to overcome this is to train the algorithm over smooth/multiple terrain, as recommended in Section 5.2.1.3.

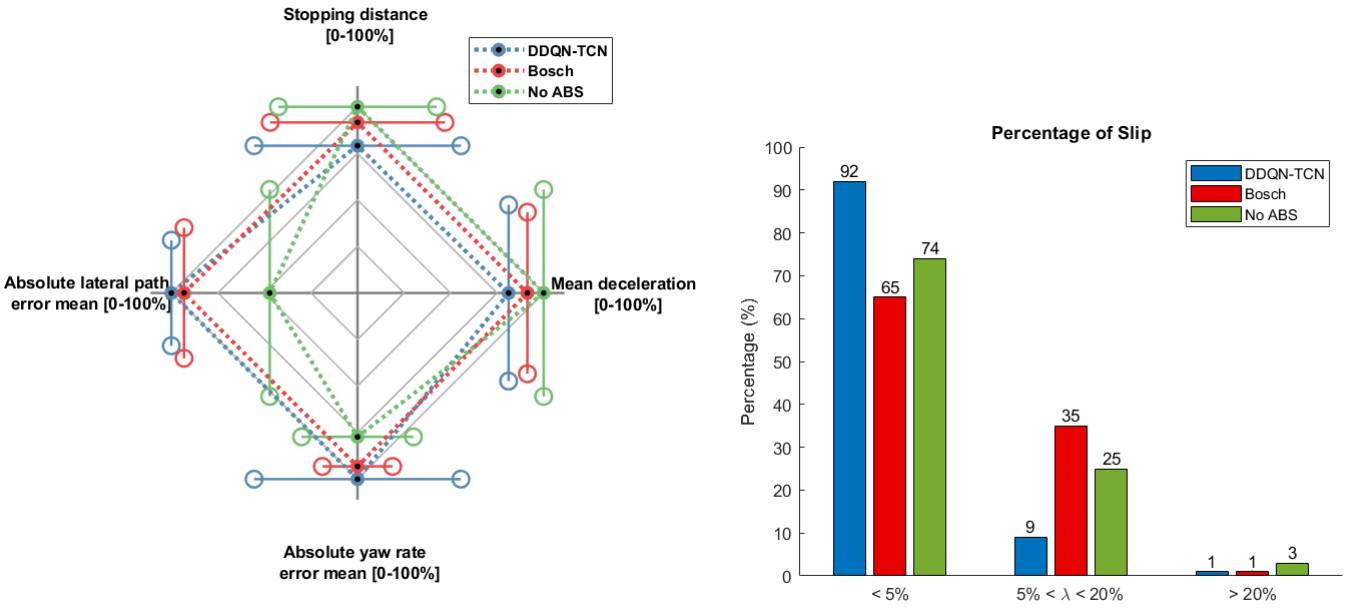


Figure 5.15: Comparison of ABS performance for braking-in-turn on high friction surface

5.2.2.3 Low Friction Flat Road

Braking-in-turn on a low friction ($\mu = 0.4$) road with a constant radius of 120m is presented in Figure 5.16.

The same limitation that both ABS algorithms experienced in straight line braking over a low friction surface (Section 5.2.1.4) is observed in this braking scenario. Braking with no ABS results in the best braking performance in 3 out of the 4 metrics, while the DDQN-TCN algorithm achieves the greatest lateral control of the vehicle. However the close scaling is indicative of how poor all three configurations are. Closer performances are seen in the stopping distance and mean deceleration which is expected with the poor directional stability. The solutions presented in Section 5.2.1.4 are recommended to overcome the limitation that the DDQN-TCN algorithm suffers over low friction surface.

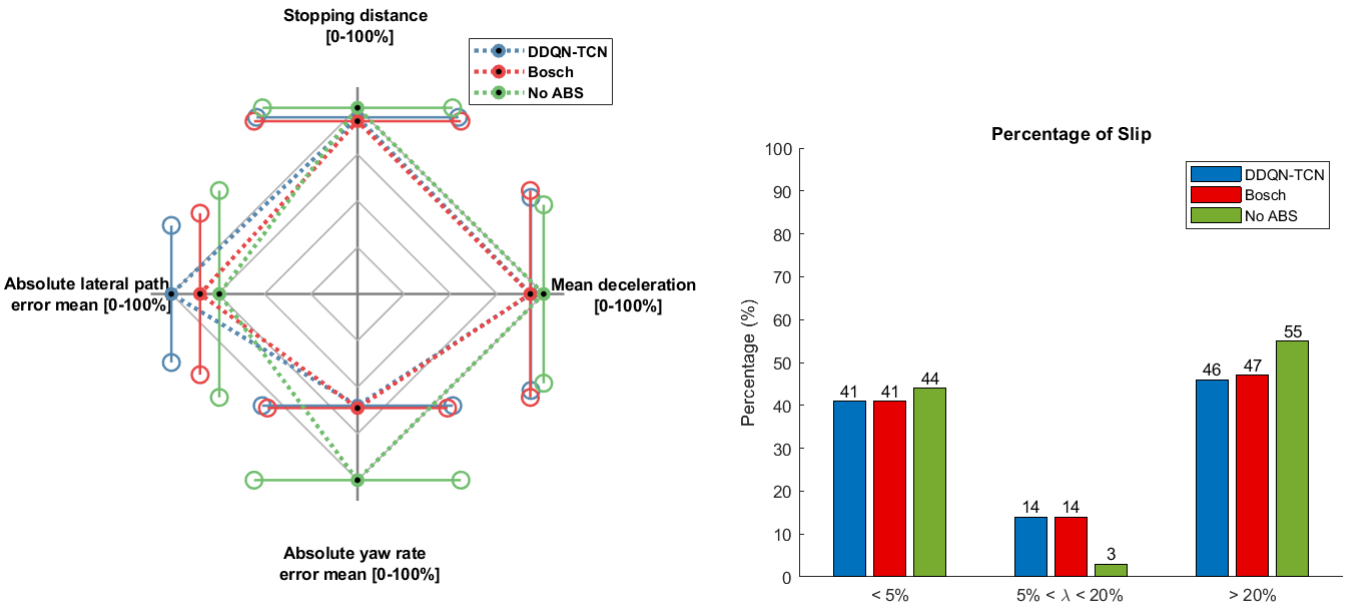


Figure 5.16: Comparison of ABS performance for braking-in-turn on low friction surface

5.2.3 Braking on Split-mu

Braking on a split-mu surface, which has a coefficient of friction of 1 under the left side and 0.4 under the right side of the vehicle is simulated, and the results are summarised in Figure 5.17.

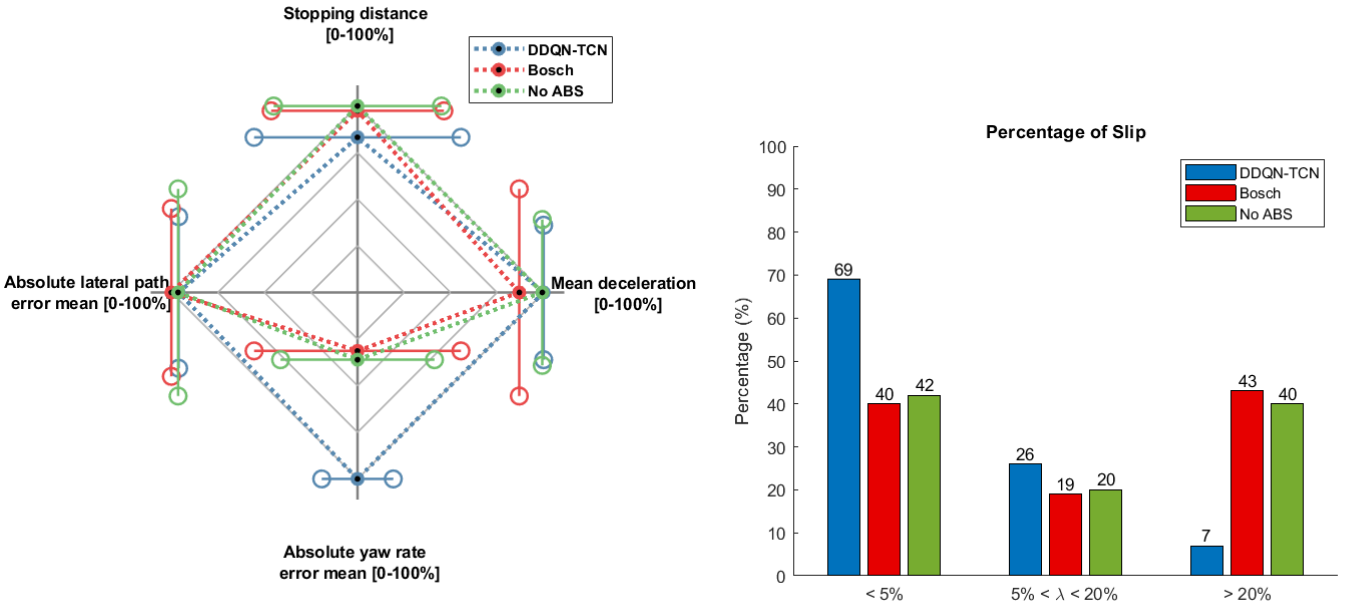


Figure 5.17: Comparison of ABS performance on split mu surface

Braking in split-mu conditions causes the vehicle to rotate, and can lead to total loss of control. Simulation results in Figure 5.17 demonstrate the advantage of braking with ABS under such

conditions as they ensure stability of the vehicle. This is seen by the DDQN-TCN algorithm, which is the best performing braking configuration in 3 out of the 4 metrics, and achieves the lowest yaw rate mean error, reducing the chance of any unwanted yaw torque acting on the vehicle. This may be due to it not braking as hard as the other configurations, with 69% of the slip less than 5%. The absolute lateral path mean error for the conventional braking simulation scales closely to braking with ABS, implying little difference in directional stability of the vehicle whether it is braking with or without ABS. This results from the limitation that the ABS algorithms suffer over the low friction side of the surface, as seen in the straight line braking and braking-in-turn simulations, and shown in Figure 5.18.

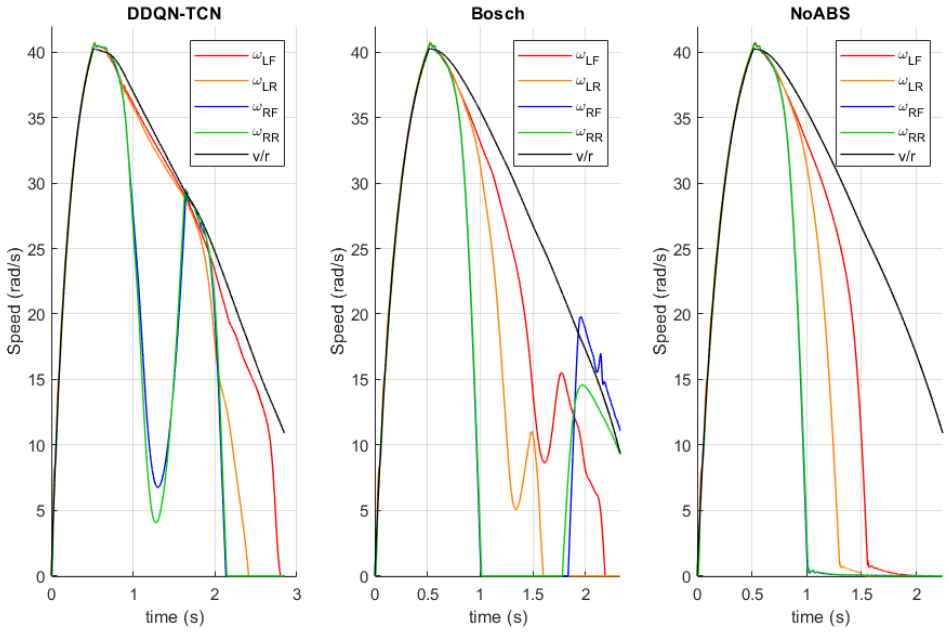


Figure 5.18: Wheel speed comparison of braking on a split-mu surface

Figure 5.18 compares the wheel speeds of the different braking configurations, where the limitation that the ABS algorithms suffer from can be seen as the tyres on the right side (low friction) lockup almost immediately, followed by the tyres on the left side. The solutions presented in Section 5.2.1.4 are recommended to overcome the limitation that the ABS algorithms suffer over low friction surface. Likened to the case of braking-in-turn on a class D road in Section 5.2.2.1, a longer braking simulation is recommended as it may revel a greater difference in the braking performance of the different braking configurations, as seen in a split-mu braking simulation completed by Hamersma (2017), which shows the poor performance of braking without ABS braking from 80km/h.

All previous simulations illustrate the generality of the DDQN-TCN algorithm across different braking scenarios and terrains using a realistic simulation model, however a key aspect of the

physical world, noisy measurement, is not considered. Section 5.2.4 investigates the robustness of the algorithm to noisy measurement data.

5.2.4 Sensitivity to Noisy Data

Two inputs are considered for the DDQN-TCN algorithm: wheel speed and vehicle speed. Both measurements are susceptible to noise, however the terrain effects and vibration of the wheel speed sensors are sufficient enough to hamper the performance of an ABS over rough terrain, thus the response of the algorithm to White Gaussian Noise (WGN) added to the wheel speed measurement is investigated.

A white noise signal, or process, is governed by a set of independent and identically distributed (i.i.d) random variables, and consequently WGN is randomly generated numbers that are independent and follow a zero-mean Gaussian distribution with some variance σ^2 (Viswanathan, 2013). This variance at any sample time reflects the intensity of the underlying white-noise process. The effect of increasing the noise variance (σ^2 = 0, σ^2 = 1, and σ^2 = 5) is investigated, and comparisons are drawn to the Bosch algorithm and conventional braking system. Using the same braking conditions as in Section 5.2.1, two straight line braking simulations are performed over the Belgian paving and class D road.

5.2.4.1 Belgian Paving

The effects of introducing noise to the wheel speed measurement are shown in Figure 5.19, which plots the left tyre of the vehicle braking with no noise, a variance of 1 and 5 respectively. This figure allows for visualisation of the scaling of different noise variances, with a variance of 5 being particularly noisy. Even with increasing the noise, the algorithm is able to prevent wheel lockup.

This is confirmed in Figure 5.20 which evaluates the performance of the DDQN-TCN algorithm using the proposed evaluation technique. Increasing the noise variance results in a general increase in the amount of wheel slippage. By achieving similar braking performances as the zero noise signal without the use of a filter, the robustness and adaptability of the algorithm to noisy data is underlined. This performance is compared to the Bosch algorithm and conventional braking system in Figure 5.21.

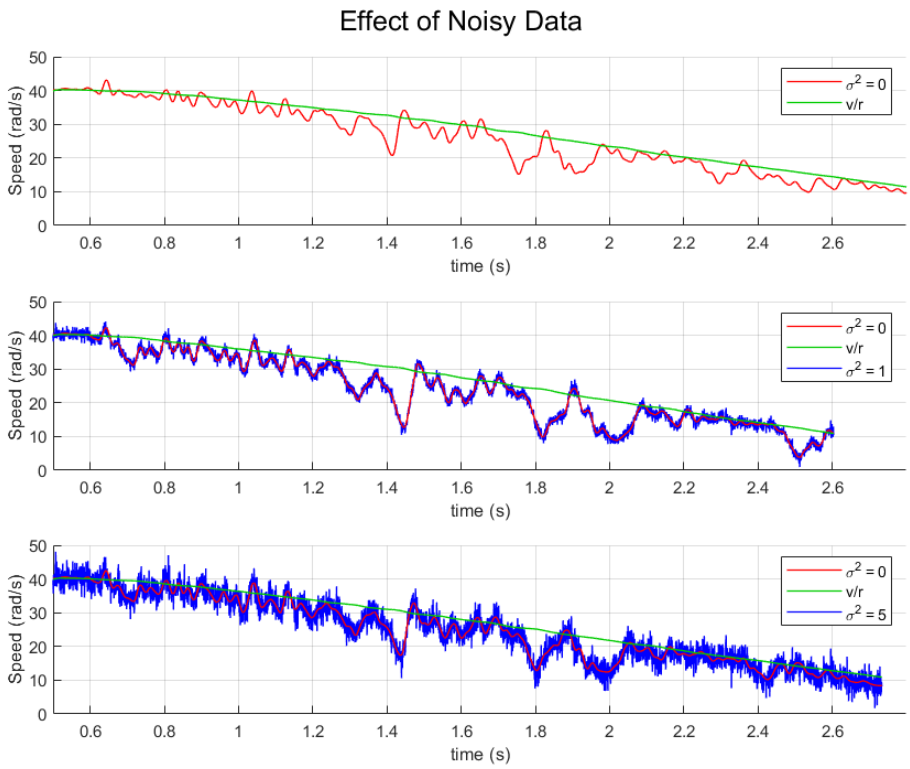


Figure 5.19: Response of the DDQN-TCN algorithm to noisy wheel speed over the Belgian paving

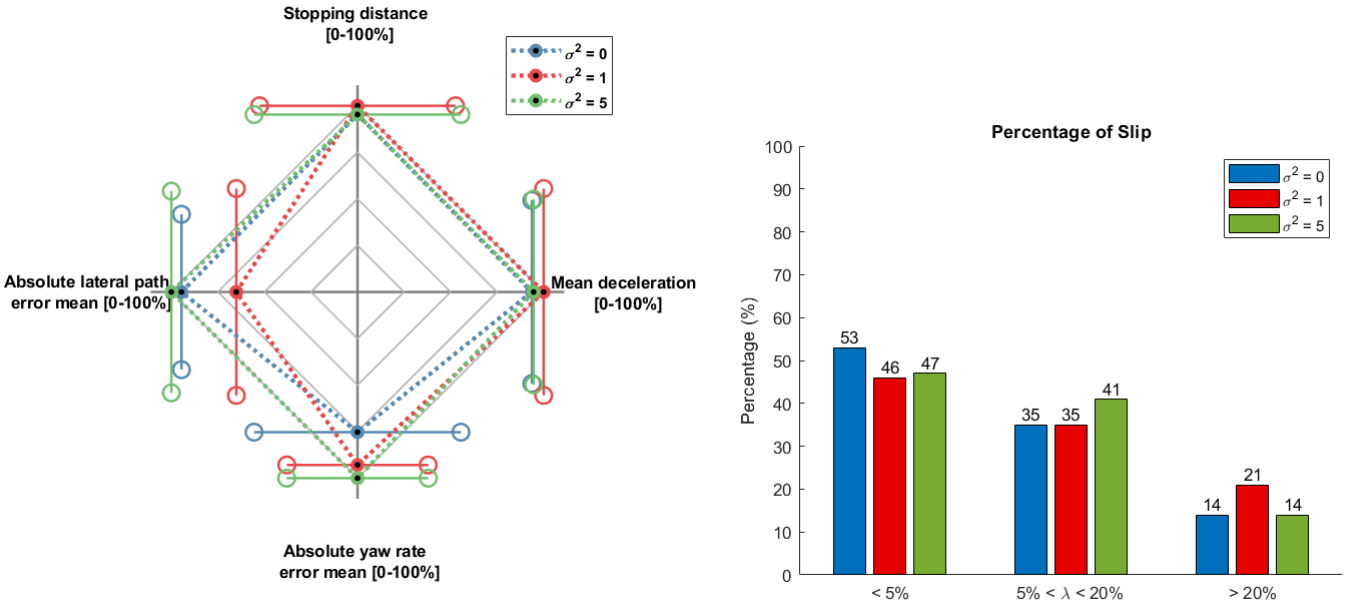


Figure 5.20: Performance of DDQN-TCN algorithm to the addition of noisy measurement over the Belgian paving

Attempts to test the ability of the Bosch algorithm at handling noisy wheel speed data without a filter proved unsuccessful. One possible reason being that the wheel speed is derived to obtain angular acceleration, and due to the noise, it is unable to generate reasonable values leading the algorithm to not brake. Thus, a first order filter (Equation 4.10) is used with β = 0.005. This highlights an advantage of the DDQN-TCN algorithm as testing the Bosch algorithm in the physical world would result in filtering delays above the modulator and actuator delays, whereas the DDQN-TCN algorithm would only experience the latter, resulting in fewer delays and better braking efficiency.

Figure 5.21 emphasises an additional advantage of the DDQN-TCN algorithm, in which it is capable of matching the performance of the filtered Bosch algorithm while achieving better directional stability of the vehicle than the conventional braking system. There is little difference in the performance of the two ABS algorithms, matching the performance of the straight line braking over the Belgian paving without noise (Section 5.2.1.1).

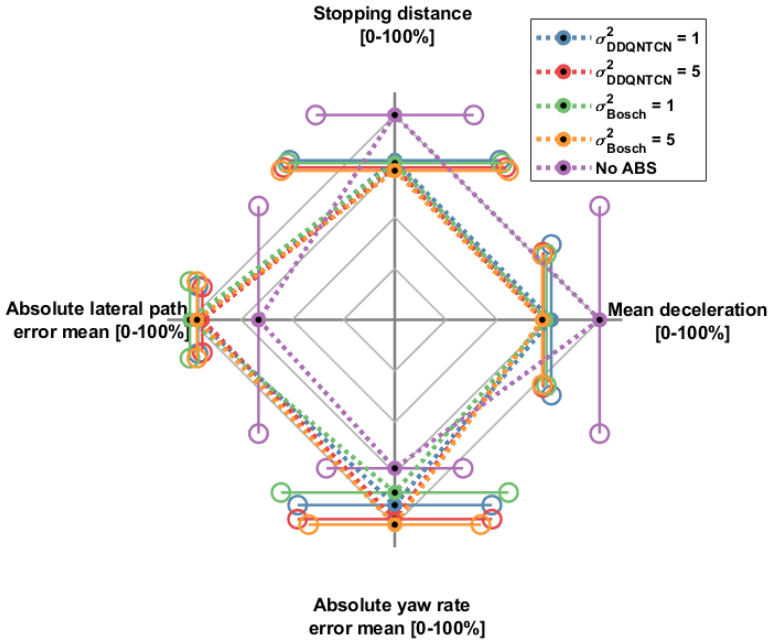


Figure 5.21: Comparison of ABS performance to noisy wheel speed over the Belgian paving

5.2.4.2 Class D Road

The same straight line braking simulation is performed over the class D road, and the effect of introducing no noise, a variance of 1, and a variance 5 to the left tyre of the vehicle is plotted in Figure 5.22.

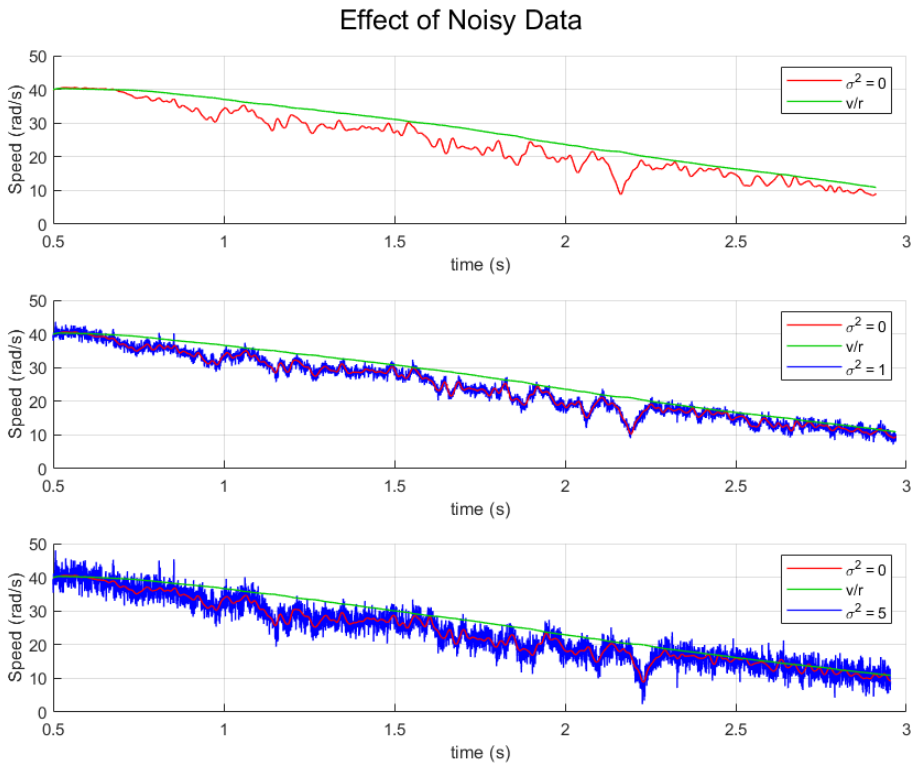


Figure 5.22: Response of the DDQN-TCN algorithm to noisy wheel speed over the class D road

As with the previous simulation over the Belgian paving, the DDQN-TCN algorithm is able to prevent wheel lockup despite an increase in the noise variance. This is confirmed by the comparison in Figure 5.23, and further substantiates the robustness and adaptability of the algorithm to noisy data over rough terrain. Figure 5.24 provides a comparison to a filtered Bosch algorithm and conventional braking system. Once again, the robustness of the DDQN-TCN algorithm is seen by its ability to match the performance of the filtered Bosch algorithm, while achieving better directional stability of the vehicle than the conventional braking system. In the event that the DDQN-TCN algorithm is unable to handle noisy data, noisy measurements can be introduced into its training. This may better the generality of the algorithm.

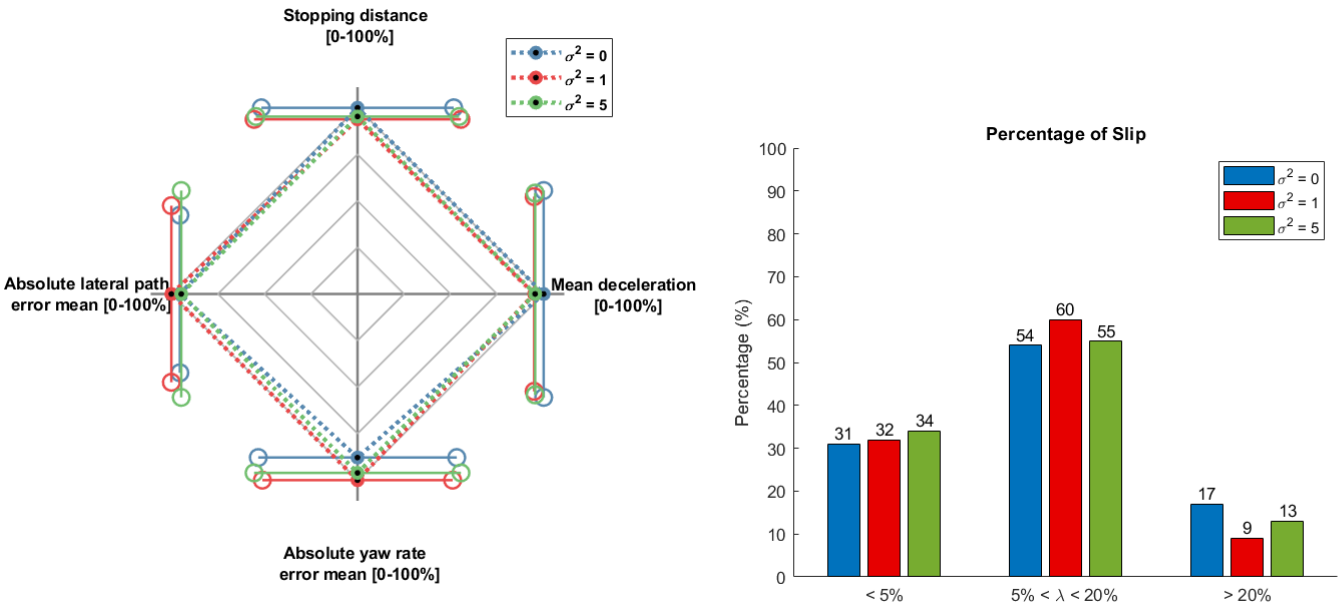


Figure 5.23: Performance of DDQN-TCN algorithm to the addition of noisy measurement over the class D road

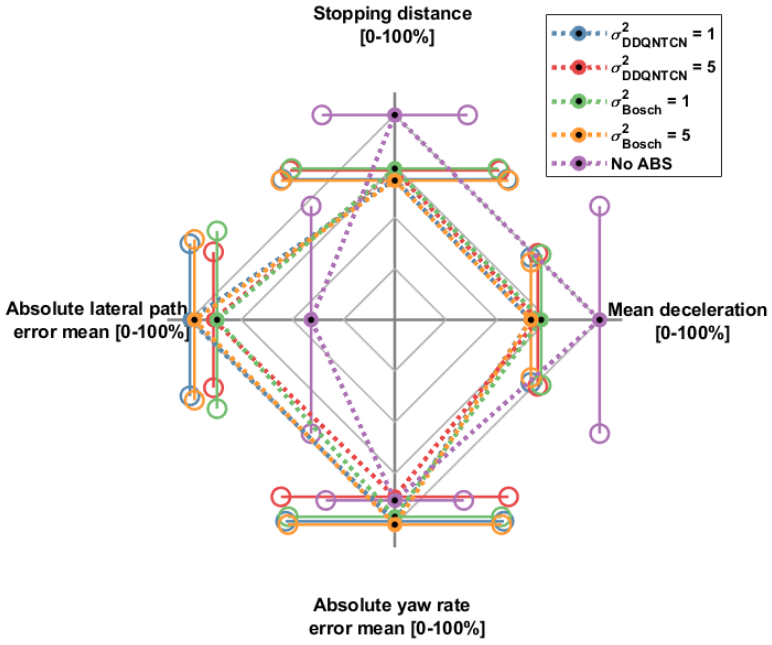


Figure 5.24: Response of the DDQN-TCN algorithm to noisy wheel speed over the class D road

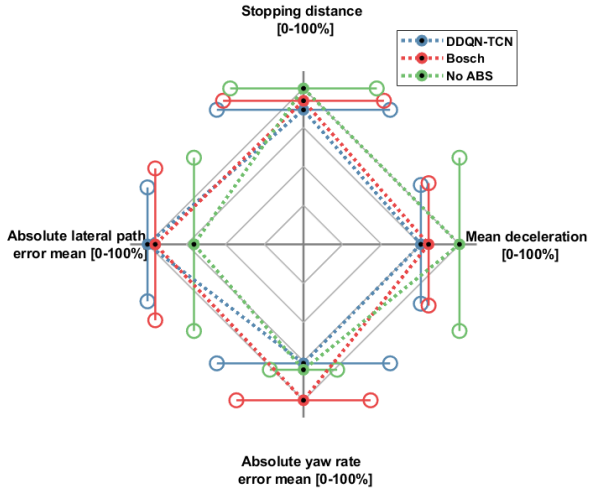
With the robustness and generality of the DDQN-TCN algorithm evaluated and substantiated, a summary of the previous results is provided.

5.2.5 Summary of Results

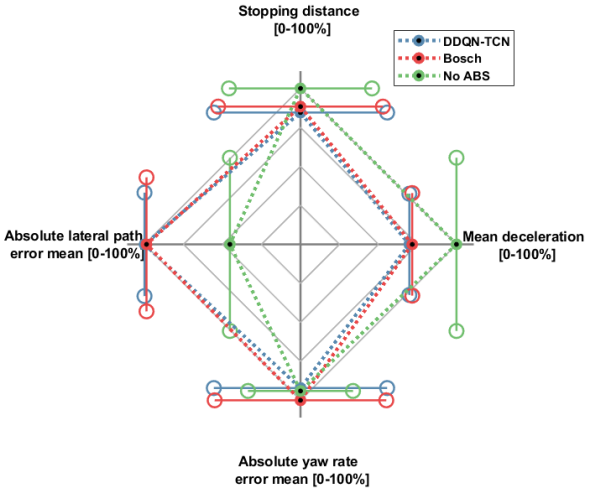
Using the ABS performance evaluation technique introduced by Hamersma (2017), a summary of the results for each braking scenario is provided in Figure 5.25.

The objective of the proposed DDQN-TCN algorithm is to prevent wheel lockup on rough terrain without significantly deteriorating the stopping distance. This objective is met in all simulations performed over rough terrain: straight line braking over the Belgian paving and class D road, and braking-in-turn over a class D road. The DDQN-TCN algorithm ensures the best directional stability of the vehicle. Its generality is further highlighted by being the best braking configuration for straight line braking and braking-in-turn on a high friction flat surface, as well as achieving the best directional stability and yaw rate mean error over the split-mu surface. The robustness and adaptability of the algorithm to noisy data over rough terrain is underlined by its ability to prevent wheel lockup and achieve similar braking performances to a filtered Bosch algorithm, without the use of a filter.

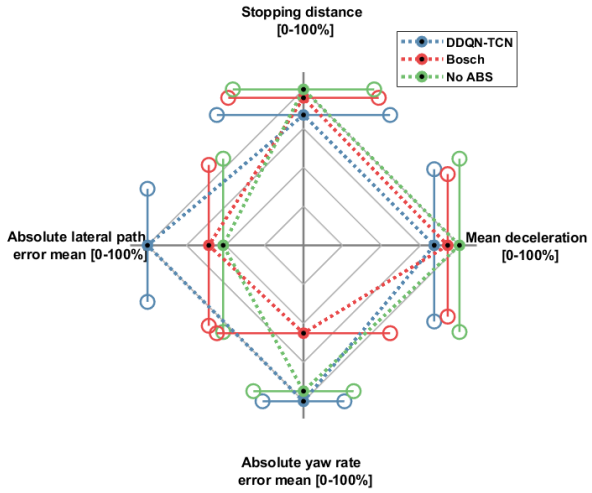
A limitation of the DDQN-TCN algorithm is identified when braking over the low friction flat surface, where wheel lockup occurs. Under these conditions, the best braking configuration is to brake without ABS. Two solutions are suggested in Section 5.2.1.4 to overcome this limitation. The first solution is training the model over a low friction surface. This presents many future opportunities for the controller, such as training over multiple surfaces. The second solution is to improve the braking efficiency of the controller by making use of a more modern ABS modulator. This would result in decreased brake delays that exist between the ABS modulator and the actuators. An investigation into this recommendation is conducted using the single tyre braking model, in which a decrease in brake delay results in an increase in brake efficiency. To visualise a larger difference in performance of the various braking configurations, a higher braking speed is recommended.



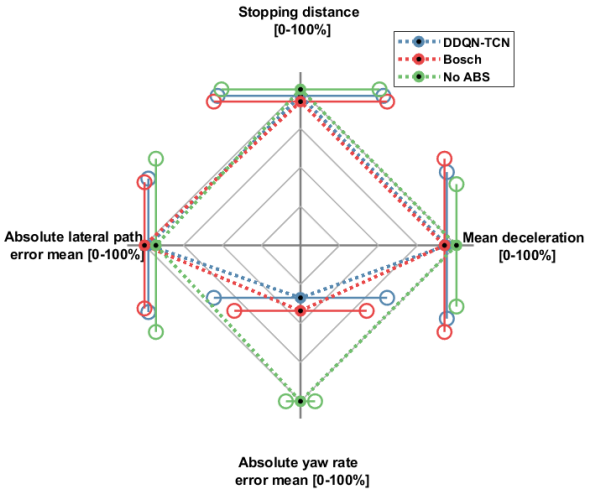
(a) Belgian paving



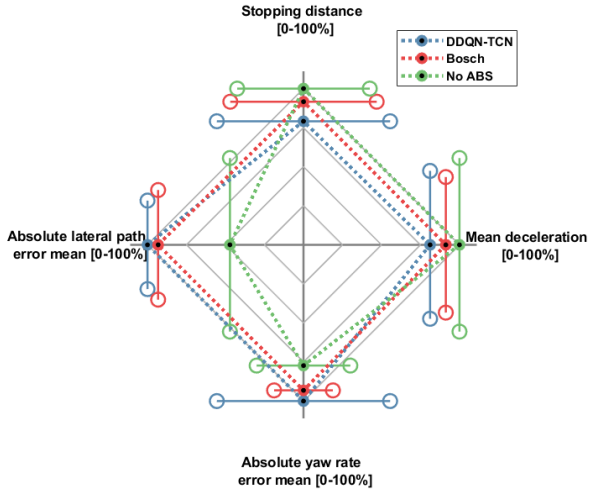
(b) Class D road



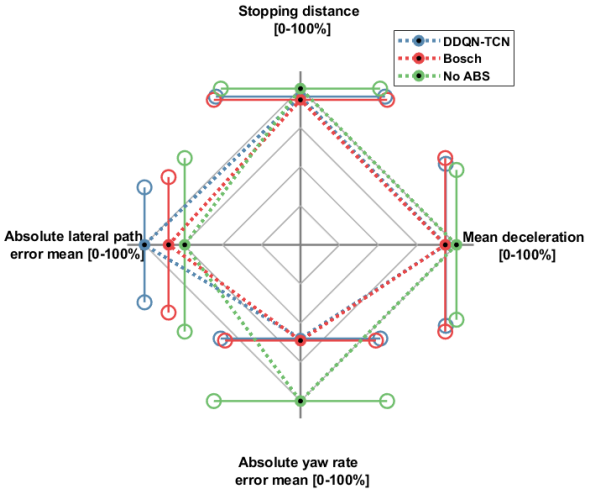
(c) Flat road, μ = 1



(d) Flat road, μ = 0.4

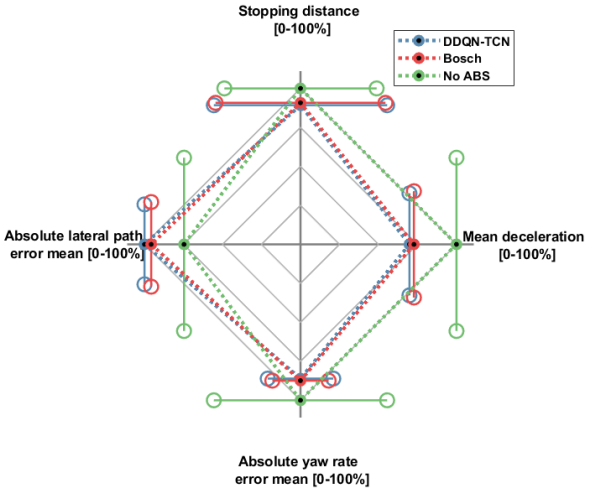


(e) Braking-in-turn: Flat road, μ = 1, r = 70m

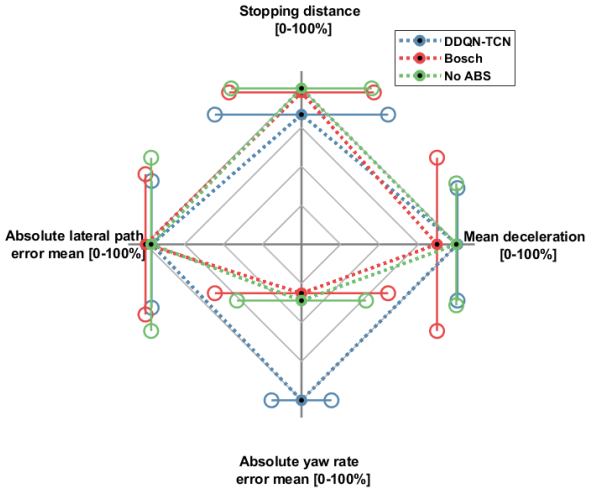


(f) Braking-in-turn: Flat road, μ = 0.4, r = 120m

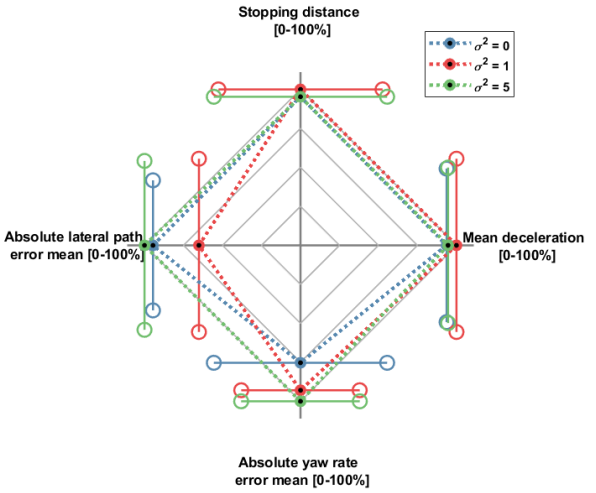
Figure 5.25: Comparison of different simulation manoeuvres



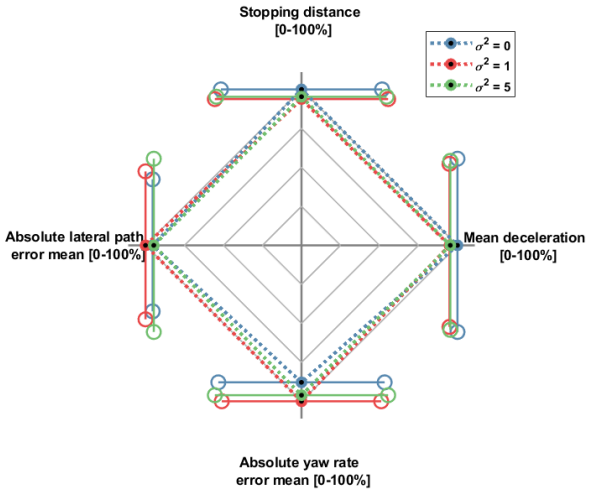
(g) Braking-in-turn: Class D, r = 70m



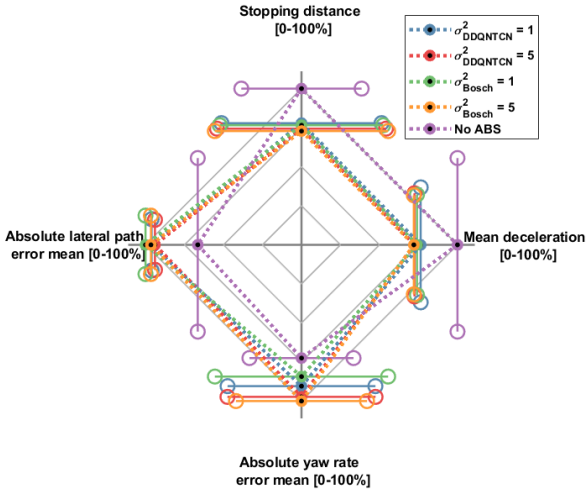
(h) Braking on split-mu surface



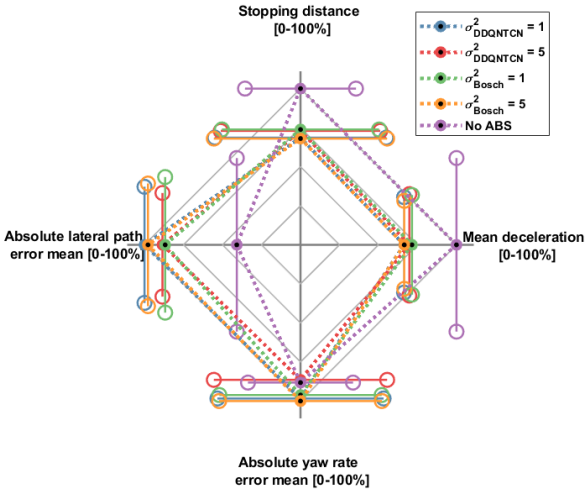
(i) DDQN-TCN noise sensitivity: Belgian paving



(j) DDQN-TCN noise sensitivity: class D road



(k) Noise sensitivity comparison: Belgian paving



(l) Noise sensitivity comparison: class D road

Figure 5.25: Comparison of different simulation manoeuvres

5.3 Conclusion

The DDQN-TCN algorithm is able to prevent wheel lockup over rough terrain without considerably diminishing the vehicle's stopping distance. The robustness of the algorithm to noisy data is emphasised by its ability to prevent wheel lockup and achieve similar braking performances to a filtered Bosch algorithm, without the use of a filter.

A limitation of the DDQN-TCN algorithm is identified when braking over a low friction flat surface, where wheel lockup occurs, and two possible solutions are proposed. The first solution involves training the model over a low friction surface and/or multiple surfaces. The second solution involves increasing the braking efficiency of the full vehicle model by using a more modern ABS modulator with faster response times. A summary of the results for each braking scenario is provided in Figure 5.25.

Chapter 6: Conclusion and Recommendations

6.1 Conclusion

This study set out to develop a model-free intelligent ABS control method capable of detecting and reducing high slip conditions on rough terrain. In an effort to overcome unmodeled dynamics and parametric uncertainties that lead to unsatisfactory performances of an ABS over rough terrain, the controller needs to prevent wheel lockup and ensure good directional stability of the vehicle without significantly deteriorating the stopping distance.

The Double Deep Q-Learning Network (DDQN) algorithm with the Temporal Convolutional Network (TCN) (labelled DDQN-TCN) is proposed as the intelligent control algorithm. A validated full-vehicle simulation model, based off the 1997 Land Rover Defender, is used with the FTire tyre model in ADAMS. This model was trained over the measured Belgian paving, with stochasticity used in the measurements to ensure the generality and robustness of the controller.

Applying the performance evaluation technique introduced by Hamersma (2017) and an initial braking speed of 55km/h, eight braking scenarios were simulated according to the SAE (2014) requirements. These scenarios can be separated into three main types: straight line braking, braking-in-turn, and braking on split-mu surface. Three braking configurations were also considered: DDQN-TCN algorithm, Bosch algorithm tuned for off-road braking, and conventional braking system (no ABS). The effect of introducing varying noise to the wheel speed measurements was also investigated.

The DDQN-TCN algorithm is able to prevent wheel lockup without significantly diminishing the vehicle's stopping distance in all simulations over rough terrain. This includes straight line braking over the Belgian paving and class D road, and braking-in-turn over a class D road. This is supported by the results from the simulations over the high friction flat surface and split-mu surface, highlighting its generality and robustness. The robustness of the algorithm to noisy data over rough terrain is underlined by its ability to prevent wheel lockup and achieve similar braking performances to that of a modified Bosch algorithm, without the use of a filter.

A limitation of the algorithm is identified on the low friction flat surface, where wheel lockup occurs

following the inability of the algorithm to cycle the wheel at the required speed. Two possible solutions are proposed: 1) training the model over a low friction surface and/or multiple surfaces, and 2) decreasing the brake delays present in the simulation model through a modern ABS modulator. These two solutions enable the algorithm to generalise over multiple surfaces, and allow for faster response times over low friction surfaces.

6.2 Recommendations

Multiple aspects which may improve the performance and evaluation of the DDQN-TCN algorithm have been identified:

1. **Experimental Validation**

   The Vehicle Dynamics Group (VDG) at the University of Pretoria have a 1997 year model Land Rover Defender 110 TDi, which is equipped with an OEM fitted ABS. Further evaluation of the control algorithm can be done experimentally over the Belgian paving at the Gerotek Testing Facilities, as done previously by (Hamersma, 2017) and (Penny, 2016).

2. **Overcome Catastrophic Forgetting**

   To ensure more effective training and a potentially better algorithm, an alternative method such as experience replay through REMIND, is recommended to overcome the catastrophic forgetting.

3. **Modern ABS Modulator**

   Current brake delays in the ABS modulator amount to over half a second (Penny, 2016). By utilizing a modern modulator, these delays can be significantly reduced, leading to faster response times and increased braking efficiency as highlighted in Section 5.2.1.4. However, it is important to maintain a balance as a faster response time may lead to finer control, resulting in an optimal slip controller without lateral force generation of the vehicle.

4. **Training over Multiple Surfaces**

   The current model is trained over the Belgian paving and its generality is tested over smooth and rough terrain. It is possible to train the model over multiple terrains to further improve its performance. Additionally it can be trained with a road classifier.

5. **Additional Performance Affecting Parameters**

   Four parameters that affect the performance of an ABS are considered: terrain, type of

controller, tyre model, and suspension setup. Additional parameters can also be explored such as tyre carcass oscillation and tyre force run-in-effect.

6. **Revisit the Reward Function**

   In the initial study performed, the reward function limited by the parameters available from a simplified quarter car model. Increasing the complexity to the full vehicle simulation model results in additional parameters being accessible; these include the yaw rate and/or path following error which can be used to improve the performance of the algorithm. It is important to avoid redundancy.

Bibliography

Adcox, J., Ayalew, B., Rhyne, T., Cron, S., and Knauff, M. (2012). Interaction of Anti-lock Braking Systems with Tire Torsional Dynamics. *Tire Science and Technology*, 40(3):171–185.

Adcox, J., Ayalew, B., Rhyne, T., Cron, S., and Knauff, M. (2013). Experimental Investigation of Tire Torsional Dynamics on the Performance of an Anti-lock Braking System. *Proceedings of the ASME Design Engineering Technical Conference, 1*.

Aly, A. A., Zeidan, E.-S., Hamed, A., Salem, F., et al. (2011). An Antilock-Braking Systems (ABS) Control: A Technical Review. *Intelligent Control and Automation*, 2:186–195.

Bai, S., Kolter, J. Z., and Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. Available at https://arxiv.org/abs/1803.01271.

Bakker, E., Pacejka, H. B., and Lidner, L. (1989). A New Tire Model with an Application in Vehicle Dynamics Studies. Technical report, SAE Technical Paper 890087.

Bauer, H. and Girling, P. (1999). *Driving-Safety Systems (2nd Edition)*. Stuttgart, Bosch.

Bhandari, R., Patil, S., and Singh, R. K. (2012). Surface Prediction and Control Algorithms for Anti-lock Brake System. *Transportation research part C: emerging technologies*, 21(1):181–195.

Blundell, M. and Harty, D. (2004). *Multibody Systems Approach to Vehicle Dynamics*. Elsevier, UK.

Botha, T. R. (2011). High Speed Autonomous Off-road Vehicle Steering. Master's thesis, University of Pretoria. Available at https://repository.up.ac.za/handle/2263/29665.

Botha, T. R. and Els, P. S. (2015). Digital Image Correlation Techniques for Measuring Tyre–Road Interface Parameters: Part 2–longitudinal tyre slip ratio measurement. *Journal of Terramechanics*, 61:101–112.

Breuer, B. and Bill, K. H. (2008). *Brake Technology Handbook*. Warrandale, Pa. SAE International.

Canudas de Wit, C., Olsson, H., Astrom, K. J., and Lischinsky, P. (1995). A New Model for Control of Systems with Friction. *IEEE Transactions on automatic control*, 40(3):419–425.

Canudas-de Wit, C., Tsiotras, P., Velenis, E., Basset, M., and Gissinger, G. (2003). Dynamic Friction Models for Road/Tire Longitudinal Interaction. *Vehicle System Dynamics*, 39(3):189–226.

Christodoulou, P. (2017). Deep-Reinforcement-Learning-Algorithms-with-PyTorch. Available at: https://github.com/p-christ/Deep-Reinforcement-Learning-Algorithms-with-PyTorch.

Christodoulou, P. (2019). Soft Actor-Critic for discrete action settings. Available at: https://arxiv.org/abs/1910.07207.

Cook, M. (2017). TCP vs. UDP: What's the difference? Available at: https://www.lifesize.com/en/blog/tcp-vs-udp/ [Accessed 19 August 2021].

Cronje, P. H. and Els, P. S. (2010). Improving Off-road Vehicle Handling using an Active Anti-roll Bar. *Journal of Terramechanics*, 47(3):179–189.

Drechsler, M. F., Fiorentin, T., and Göllinger, H. (2021). Actor-Critic Traction Control Based on Reinforcement Learning with Open-Loop Training. *Modelling and Simulation in Engineering*, 2021.

Els, P. S., Botha, T., Hamersma, H., Becker, C., Savitski, D., Heidrich, L., and Höpping, K. (2013). The Effect of Controllable Suspension Settings on the ABS Braking Performance of an Off-road Vehicle on Rough Terrain. In *Proc. of the 7th ISTVS Regional Americas Conference at Tampa, Florida*.

Els, P. S., Theron, N. J., Uys, P. E., and Thoresson, M. J. (2007). The Ride Comfort vs. Handling Compromise for Off-road Vehicles. *Journal of Terramechanics*, 44(4):303–317.

Fu, Y., Li, C., Yu, F. R., Luan, T. H., and Zhang, Y. (2020). A Decision-making Strategy for Vehicle Autonomous Braking in Emergency via Deep Reinforcement Learning. *IEEE transactions on vehicular technology*, 69(6):5876–5888.

Garrott, W. R. and Mazzae, E. N. (1999). An Overview of the National Highway Traffic Safety Administration's light vehicle Antilock Brake Systems Research Program. Technical report, SAE Technical Paper 1999-01-1286.

Gillespie, T. D. (1992). *Fundamentals of Vehicle Dynamics*, volume 400. Society of automotive engineers Warrendale, PA.

Gipser, M. (1999). FTire, A new fast tire model for ride comfort simulations. In *International ADAMS User's Conference Berlin. Available at https://www.cosin.eu/wp-content/uploads/ftire_eng_1.pdf*.

Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018). Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1861–1870. PMLR.

Hamersma, H. A. (2013). Longitudinal Vehicle Dynamics Control for Improved Vehicle Safety. Master's thesis, University of Pretoria. Available at https://repository.up.ac.za/handle/2263/40795.

Hamersma, H. A. (2017). *ABS Braking on Rough Terrain*. PhD thesis, University of Pretoria. Available at: https://repository.up.ac.za/handle/2263/61607.

Hamersma, H. A. and Els, P. S. (2014). Improving the Braking Performance of a Vehicle with ABS and a Semi-Active Suspension System on a Rough Road. *Journal of Terramechanics*, 56:91–101.

Hayes, T. L., Kafle, K., Shrestha, R., Acharya, M., and Kanan, C. (2020). Remind Your Neural Network to Prevent Catastrophic Forgetting. In *European Conference on Computer Vision*, pages 466–483. Springer, Cham.

Hester, T., Quinlan, M., and Stone, P. (2012). A Real-time Model-based Reinforcement Learning Architecture for Robot Control. In *2012 IEEE International Conference on Robotics and Automation*, pages 85–90. IEEE.

ISO (2016). ISO 8608:2016 Mechanical Vibration - Road Surface Profiles - Reporting of Measured Data. *International Organisationg of Standardisation. Geneva, Switzerland*, Available at: https://www.sis.se/api/document/preview/921131/.

John, S. and Pedro, J. O. (2013). Neural Network-based Adaptive Feedback Linearization Control of Antilock Braking System. *International Journal of Artificial Intelligence*, 10(S13):21–40.

Keshmiri, R. and Shahri, A. M. (2007). Intelligent ABS Fuzzy Controller for Diverse Road Surfaces. *World Academy of Science, Engineering and Technology*, 2(2):62–67.

Kikuuwe, R., Takesue, N., Sano, A., Mochiyama, H., and Fujimoto, H. (2006). Admittance and Impedance Representations of Friction Based on Implicit Euler Integration. *IEEE Transactions on Robotics*, 22(6):1176–1188.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al. (2017). Overcoming Catastrophic Forgetting in Neural Networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526.

Lee, C.-C. (1990). Fuzzy logic in control systems: Fuzzy Logic Controller. I. *IEEE Transactions on systems, man, and cybernetics*, 20(2):404–418.

Liang, W., Medanic, J., and Ruhl, R. (2008). Analytical Dynamic Tire Model. *Vehicle System Dynamics*, 46(3):197–227.

Lin, C.-M. and Hsu, C.-F. (2003). Neural-Network Hybrid Control for Antilock Braking Systems. *IEEE Transactions on Neural Networks*, 14(2):351–359.

Michelin (2021). Michelin LTX A/T2 TIRES. Available at: https://www.michelinman.com/auto/tires/michelin-ltx-a-t2 [Accessed 18 August 2021].

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D., and Kavukcuoglu, K. (2016). Asynchronous Methods for Deep Reinforcement Learning. In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1928–1937, New York,USA. PMLR.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. *Available at: https://arxiv.org/abs/1312.5602*.

OpenAI (2018). Part 2: Kinds of RL Algorithms. Available at: https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html [Accessed 1 November 2020].

Pacejka, H. (2005). *Tire and Vehicle Dynamics*. Elsevier, UK.

Pan, S. J. and Yang, Q. (2009). A Survey on Transfer Learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359.

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., and Wermter, S. (2019). Continual Lifelong Learning with Neural Networks: A Review. *Neural Networks*, 113:54–71.

Penny, W. C. W. (2016). Anti-lock Braking System Performance on Rough Terrain. Master's thesis, University of Pretoria. Available at https://repository.up.ac.za/handle/2263/56099.

Penny, W. C. W. and Els, P. S. (2016). The Test and Simulation of ABS on Rough, Non-deformable Terrains. *Journal of Terramechanics*, 67:1–10.

Poursamad, A. (2009). Adaptive Feedback Linearization Control of Antilock Braking Systems using Neural Networks. *Mechatronics*, 19(5):767–773.

Radac, M.-B. and Precup, R.-E. (2018). Data-Driven Model-Free Slip Control of Anti-lock Braking Systems using Reinforcement Q-learning. *Neurocomputing*, 275:317–329.

Reul, M. and Winner, H. (2009). Enhanced Braking Performance by Integrated ABS and Semi-active Damping Control. In *Proc. 21st Enhanced Safety Vehicles (ESV) Conf*, pages 1–14.

RTMC (2017). *National Road Safety Strategy 2016-2030*. Road Traffic Management Corporation. Available at: http://www.rtmc.co.za/images/rtmc/docs/nrss/National%20Road%20Safety%20Strategy1.pdf.

SAE (2014). *Antilock Brake System Review*. SAE International, USA. DOI:https://doi.org/10.4271/J2246_201404.

Sardarmehni, T. and Heydari, A. (2015). Optimal Switching in Anti-lock Brake Systems of ground vehicles based on Approximate Dynamic Programming. In *Dynamic Systems and Control Conference*, volume 3: Vehicle Dynamics Control. American Society of Mechanical Engineers.

Shao, J., Zheng, L., Li, Y. N., Wei, J. S., and Luo, M. G. (2007). The Integrated Control of Anti-lock Braking System and Active Suspension in Vehicle. *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*, 4:519–523.

Stallmann, M. J. and Els, P. S. (2014). Parameterization and Modelling of Large Off-road Tyres for Ride Analyses: Part 2–parameterization and validation of tyre models. *Journal of Terramechanics*, 55:85–94.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction. 2nd Edition*. The MIT press. Cambridge, Ma.

Thoresson, M. J. (2007). *Efficient Gradient-based Optimisation of Suspension Characteristics for an Off-road Vehicle*. PhD thesis, University of Pretoria. Available at: https://repository.up.ac.za/bitstream/handle/2263/26984/Complete.pdf?sequence=8&isAllowed=y.

Thoresson, M. J., Botha, T., and Els, P. S. (2014). The Relationship between Vehicle Yaw Acceleration Response and Steering Velocity for Steering Control. *International journal of vehicle design*, 64(2-4):195–213.

Uys, P. E., Els, P. S., and Thoresson, M. (2007). Suspension Settings for Optimal Ride Comfort of Off-road Vehicles Travelling on Roads with Different Roughness and Speeds. *Journal of Terramechanics*, 44(2):163–175.

Van der Merwe, N. A., Els, P. S., and Žuraulis, V. (2018). ABS Braking on Rough Terrain. *Journal of Terramechanics*, 80:49–57.

Van der Westhuizen, S. F. and Els, P. S. (2015). Comparison of different gas models to calculate the spring force of a hydropneumatic suspension. *Journal of Terramechanics*, 57:41–59.

Van Hasselt, H., Guez, A., and Silver, D. (2016). Deep Reinforcement Learning with Double Q-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence, 30(1)*.

Viswanathan, M. (2013). White Noise: Simulation and Analysis using MATLAB. Available at: https://www.gaussianwaves.com/2013/11/simulation-and-analysis-of-white-noise-in-matlab/ [Accessed 8 October 2021].

WABCO (2003). Hydraulic add-on ABS System. Available at: http://gershon.ucoz.com/WABCO/815_430.pdf [Accessed 28 September 2021].

Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine learning*, 8(3-4):279–292.

WHO (2015). *Global Status Report on Road Safety 2015*. World Health Organization.

Yu, H., Qi, Z., Duan, J., Taheri, S., and Ma, Y. (2015). Multiple Model Adaptive Backstepping Control for Antilock Braking System based on LuGre Dynamic Tyre Model. *International Journal of Vehicle Design*, 69(1-4):168–184.

Zhao, D., Wang, B., and Liu, D. (2013). A Supervised Actor–Critic Approach for Adaptive Cruise Control. *Soft Computing*, 17(11):2089–2099.

Appendices

Appendix A: Hyperparameters

Table A.1: Hyperparameters used for the DQN agents

Hyperparameters	**Value**
learning_rate	0,01
batch_size	256
buffer_size	10000
epsilon	1
epsilon_decay_rate_denominator	1
discount_rate	0,99
tau	0,01
alpha_prioritised_replay	0,6
beta_prioritised_replay	0,1
incremental_td_error	1,00E-08
update_every_n_steps	1
linear_hidden_units	[128,256]
final_layer_activation	None
batch_norm	FALSE
gradient_clipping_norm	0,7
learning_iterations	1
clip_rewards	FALSE

Table A.2: Hyperparameters used for the Actor Critic Agents

Hyperparameters	**Value**
learning_rate	0,005
batch_size	256
epsilon	1
epsilon_decay_rate_denominator	1
discount_rate	0,99
exploration_worker_difference	2
clip_rewards	FALSE
update_every_n_steps	4
linear_hidden_units	[128,256]
final_layer_activation	[Softmax, None]
normalise_rewards	TRUE
automatically_tune_entropy_hyperparameter	TRUE
entropy_term_weight	None
add_extra_noise	FALSE
do_evaluation_iterations	TRUE
learning_updates_per_learning_session	1

Table A.3: Hyperparameters used for Actor

Hyperparameters	**Value**
learning_rate	0,0003
linear_hidden_units	[128,256]
final_layer_activation	Softmax
tau	0,005
batch_norm	FALSE
gradient_clipping_norm	5
initialiser	Xavier

Table A.4: Hyperparameters used for Critic

Hyperparameters	**Value**
learning_rate	0,0003
linear_hidden_units	[128,256]
final_layer_activation	None
tau	0,005
batch_norm	FALSE
buffer_size	1000000
gradient_clipping_norm	5
initialiser	Xavier