Enhancing Automotive Stability Control
with Artificial Neural Networks

The intelligent car Project

By

David Andrew Butler


Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Engineering, University of Tasmania

September 2006
Declaration and Authority of Access

This thesis contains the results of research done at the School of Engineering, University of Tasmania, Hobart, Tasmania, Australia between 2003 and 2006.

It contains no material that has been accepted for the award of any other higher degree or graduate diploma in any tertiary institution and, to the best of the author’s knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis contains confidential information and is not to be disclosed or made available for loan or copy without the express written permission from the University of Tasmania (i). Once released the thesis may be available for loan and limited copying in accordance with the Copyright Act 1968.

(i) Enquiries should be directed to the Research and Development Office

___________________________
David Butler
School of Engineering,
University of Tasmania,
Hobart, Tasmania, Australia
September 2006
Acknowledgements

This PhD investigation is the result of much hard work, not only by the author. The greatest outcome, to my admiration, is its contribution to my recent marriage to my new wife Bonnie Butler. After meeting just one year before the PhD started, we were married in January this year. Bonnie’s help in patiently listening to my exhaustive rhetorical conversations, despite very little interest in the subject area, was invaluable. I was particularly amazed at her ability to constantly supply useful points and comments, even though she must have been bored senseless. Credit should also be given for her help correcting spelling and grammar within the thesis, which was a very time consuming task when she was already very busy. Thanks for the support Bon!

I wonder, has a PhD thesis ever been completed without the author thanking his mother? This is no different, with special thanks going to mum for spending a week of her time correcting thesis drafts. My family, and my new family-in-law, must also be thanked for providing continuous love and support.

In regards to the technical content of the thesis, my supervisor Vishy Karri clearly needs to be thanked. His ability to manage a huge number of postgraduates in parallel is amazing and, despite his limited time, his capacity to provide comments on research methodology and thesis structure was very useful. John McColloch is thanked for his data acquisition and LabVIEW programming support, particularly in writing the sensor communication routines and building the wheel speed PIC.

International experience was a significant part of this work too, and Peter Rossmanek is thanked for accommodating me at the Fachhochschule Stralsund, Germany to complete automotive research leading up to the PhD study. Ole Madsen’s help in organising and supporting my study into intelligent control at Aalborg University, Denmark is also highly appreciated.

Final thanks goes to all of the Hydrogen and Allied Renewable Technologies (HART) and Intelligent Car team members. The chance to be involved with each of your projects was a very useful distraction, and I learnt a lot. A highlight was the opportunity to drive the vehicle used in this investigation around the state in the 2006 Targa Tasmania tarmac rally, running on hydrogen of course!
Abstract

Many studies of automotive crash statistics have shown that driver error is a major cause of accident and injury on the roads worldwide. This has lead to the development of many active control systems to aid the driver during panic maneuvers, such as antilock braking systems. Nonetheless, there has been a slow growth in the control methodology of these systems, with wheel speed regulation based on the information derived from a small number of sensors the norm across all past and present systems. To achieve greater performance gains, it is important to control more vehicle parameters and obtain vehicle state information from larger sensor arrays. Problems arise using traditional control methodology, as additional variables create exponential increases in control algorithm complexity, and in computational requirements.

Artificial neural networks (ANN) are presented in literature as an artificial intelligence solution to approaching problems. Significant benefits include, the ability to model highly non-linear and complex systems, capacity to incorporate many model inputs and outputs, low computational requirements and capability for self-learning from observed data. However, previous work has largely been limited to simulation or very narrow practical testing, from which it is difficult to draw useful conclusions.

This thesis addresses these problems by developing two new ANN systems, implemented in broad practical tests. The first uses suspension and wheel speed vibration to intelligently predict road surface conditions, which is a major performance limitation in all current systems. The second models complex vehicle dynamics through a large sensor array and ANN process optimisation to implement intelligent traction control. This method determines the optimal driven wheel speed for maximum acceleration in the driver’s desired direction, in a process that is generic and adaptable to current and future active control systems.

All results are derived from a real test vehicle, which was adapted for this investigation. This included the installation of chassis and engine sensors, data acquisition and control systems, engine management hardware and user interfaces, as well as constructing ANN models and controllers in the NI LabVIEW language. The positive outcomes of this work are a step towards establishing new methods of active vehicle control on a statistical and quantitative basis.
# Table of Contents

- **1 - INTRODUCTION** .............................................................................................................. 1

- **2 - VEHICLE STABILITY BACKGROUND** ........................................................................... 13
  2.1 **SYSTEMS THAT ASSIST DRIVERS** .............................................................................. 14
  2.2 **TYRE DYNAMICS FUNDAMENTALS** ........................................................................ 18
  2.3 **CONTEMPORARY VEHICLE STABILITY CONTROL** ...................................................... 33
  2.4 **RESEARCH METHOD** .................................................................................................. 85
  2.5 **REMARKS** .................................................................................................................. 93

- **3 - ARTIFICIAL NEURAL NETWORKS** ............................................................................ 94
  3.1 **ANN OPERATION** ...................................................................................................... 95
  3.2 **AUTOMOTIVE ANN APPLICATIONS** ....................................................................... 109
  3.3 **REMARKS** .............................................................................................................. 115

- **4 - CHASSIS SENSORS AND DATA LOGGER INSTALLATION** ....................................... 116
  4.1 **TEST VEHICLE** ....................................................................................................... 117
  4.2 **CHASSIS SENSORS** .................................................................................................. 118
  4.3 **ADVANCED DASH LOGGER** .................................................................................... 143
  4.4 **ADVANCED DASH LOGGER INSTALLATION** ............................................................ 144
  4.5 **ADVANCED DASH LOGGER CONFIGURATION** ........................................................ 151
  4.6 **INTERPRETER SOFTWARE** ....................................................................................... 152
  4.7 **REMARKS** .............................................................................................................. 152

- **5 - PAVEMENT FEATURE RECOGNITION DURING STABLE DRIVING CONDITIONS** ................................................................. 154
  5.1 **SURFACE PREDICTION USING ANN** ........................................................................ 155
  5.2 **CHOICE OF ANN MODEL** ........................................................................................ 159
  5.3 **DATA ACQUISITION** .................................................................................................. 168
  5.4 **SOFTWARE** .............................................................................................................. 174
  5.5 **ANN MODEL ANALYSIS** ........................................................................................ 193
  5.6 **REMARKS** .............................................................................................................. 216

- **6 - ENGINE MANAGEMENT SYSTEM INSTALLATION** .................................................... 217
  6.1 **TEST VEHICLE LIMITATIONS** .................................................................................. 218
Automotive transport is arguably one of the most important, yet most dangerous, aspects of life in the developed and developing world. The ability for individuals to travel significant distances in their own time and at their own convenience is a profound liberty. Likewise, the economic advantages of relatively cheap and highly flexible transport are massive. The automobile has provided a freedom of movement that has been dreamed of through the ages, but never attained in history. [1]

This freedom of movement is now taken for granted through much of the world, often to the extent that much of the cost of automotive transport is almost forgotten. At present, cars use non-renewable fuels that deplete the Earth of its resources and cause environmental damage to extract. Further, when these fuels are burnt in a vehicle’s internal combustion engine they produce chemicals that cause damage to the environment, cause human health problems and, through smog and noise pollution, reduce the living conditions of entire cities. The limited availability of fuel is also cause for political unrest, and is often cited as a significant contributing factor to international conflict.

Nevertheless, these are only part of the cost of using automotive transport, which for the private user are relatively abstract problems that relate little to day-to-day life. The obvious cost for the average automotive user comprises of the running cost of the vehicle and the risk and severity of any accident that may happen. The latter cost is often only fully comprehended when an accident or a “near miss” occurs but, when weighed against the benefits of automotive transport, is generally considered acceptable throughout the world.

Since the first fatal automotive crash in 1896, three million people have been killed in traffic accidents in the USA alone. Further, over one million people in total are killed on the roads throughout the world each year, a number that is expected to double by 2020. Each year over 40,000 people die in the USA as a result of automotive accidents, and five million receive minor (MAIS 1) injuries, as shown in Table 1.1. Applying this ratio to the rest of the world suggests more than 120 million automotive injuries each year. For the average person, this represents a 50% chance of being injured in a lifetime; making traffic crashes one of the world’s largest public health problems. [2]
1. Introduction

It can be seen that by sheer numbers, the personal loss for individuals involved in traffic accidents is enormous. From an economic perspective this is equally true. By converting all losses to monetary values, traffic accidents cost the USA government $231 billion in 2000. This is a huge amount of money, and is greater than the GDP of most countries. This is placed into broader scale when the relative safety of driving in the USA is compared to the rest of the world, as is shown in Figure 1.1. [2]

To further break these statistics up, it is necessary to focus on particular countries due to statistical data collection and association difficulties. Literature from the USA will be presented here predominately, due to its widespread publication. This data is considered of greater statistical value than many other countries because of the large number of vehicle on the roads in the USA and the quality and quantity of US Department of Transport record keeping [2].

Firstly, a breakdown of fatalities based on the vehicle type the person killed was occupying is given in Table 1.2. This is followed by Table 1.3, which shows the distribution of the number of vehicles involved in a fatal accident, and also a list of the object that caused the most harm when struck.

<table>
<thead>
<tr>
<th>event</th>
<th>number per year</th>
<th>average travel between events</th>
<th>average time between events</th>
</tr>
</thead>
<tbody>
<tr>
<td>driver killed</td>
<td>26549</td>
<td>172 million km</td>
<td>7300 years</td>
</tr>
<tr>
<td>driver involved in fatal crash</td>
<td>57803</td>
<td>79 million km</td>
<td>3400 years</td>
</tr>
<tr>
<td>MAIS 3 (Serious) driver injury</td>
<td>7800</td>
<td>55 million km</td>
<td>2500 years</td>
</tr>
<tr>
<td>MAIS 3 (Serious) person injury</td>
<td>12000</td>
<td>36 million km</td>
<td>1600 years</td>
</tr>
<tr>
<td>MAIS 1 (Minor) driver injury</td>
<td>5 million</td>
<td>1.5 million km</td>
<td>65 years</td>
</tr>
<tr>
<td>MAIS 1 (Minor) person injury</td>
<td>5 million</td>
<td>1.0 million km</td>
<td>42 years</td>
</tr>
<tr>
<td>involved in crash</td>
<td>16 million</td>
<td>0.3 million km</td>
<td>12 years</td>
</tr>
</tbody>
</table>

*Table 1.1: Average frequencies of various USA crash outcomes [2]*

![Figure 1.1: Fatality rate versus degree of motorisation [2]](image)
It can be seen the road user that suffers from the greatest proportion of fatal crashes is the driver of normal passenger cars, which is indicative of the traffic makeup. Likewise, light truck drivers makeup a significant proportion of deaths, followed by car passengers, pedestrians and light truck passengers. These five categories represent almost 90% of all USA fatalities. While this is indicative of the distribution of automobile/road users, it also highlights the potential benefits that could be realised through even a minor improvement in safety.

The second table shows that approximately half of the fatalities recorded are a result of multiple vehicle crashes, although two car crashes comprise a large majority. In these cases, 90% of the fatalities were cause from striking another vehicle [2]. The other half of USA fatalities are a result of single vehicle accidents, in which case approximately 40% of deaths are a result of a vehicle overturn and at least 50% from collision with a
stationary solid object. The most likely modes of death on the roads, in order of likelihood, are collision with another vehicle, collision with a stationary object and vehicle overturn.

Further, J. Koopman and G. Najm [3] state that in 1998 off-roadway crashes (defined as when the first harmful event occurs off the roadway, and includes collision with stationary objects and some rollovers) comprised of around 11.5% of all reported crashes in the USA. Likewise, rollovers accounted for only 1.8% of USA reported crashes in 1996 [4]. Therefore, even though rollovers occur less frequently than off-roadway crashes they have an unproportionally large contribution to fatal injuries [5].

Speed has an obvious effect on crash survivability, with the effects shown in Table 1.4 and Figure 1.2. It can be seen that as vehicles travel faster, the changes in velocity during accidents is larger, and the survivability of crashes exponentially reduces. The reason why this trend is not carried past $\Delta v > 65$km/hr is simply because of the small number of valid cases used in the study at this speed [6].

<table>
<thead>
<tr>
<th>MAIS level</th>
<th>Injury</th>
<th>Fatal equivalent value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>uninjured</td>
<td>0.0014</td>
</tr>
<tr>
<td>1</td>
<td>minor</td>
<td>0.0087</td>
</tr>
<tr>
<td>2</td>
<td>moderate</td>
<td>0.0417</td>
</tr>
<tr>
<td>3</td>
<td>serious</td>
<td>0.1250</td>
</tr>
<tr>
<td>4</td>
<td>severe</td>
<td>0.2765</td>
</tr>
<tr>
<td>5</td>
<td>critical</td>
<td>0.8463</td>
</tr>
<tr>
<td>6</td>
<td>fatal</td>
<td>1</td>
</tr>
</tbody>
</table>

*Table 1.4: MAIS injury comparison based on economic cost [6]*

Lastly, road surface condition and environmental conditions have an effect on vehicle safety. It can be seen in Table 1.5 and Table 1.6 that most fatal accidents happen in dry roads and in good weather. It can be assumed this is because these form the most common conditions, but may also be because the travel speeds for dry roads and good weather are generally higher than for adverse conditions. For example, even though the likelihood of a crash in snow is higher than on a dry pavement, the overall fatality risk is smaller due to lower travel speeds used in snow conditions.
1. Introduction

Table 1.5: Percent of USA fatal crashes for different road surface conditions [2]

<table>
<thead>
<tr>
<th>road surface condition</th>
<th>percent of fatal crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>dry</td>
<td>83.6%</td>
</tr>
<tr>
<td>wet</td>
<td>12.7%</td>
</tr>
<tr>
<td>ice</td>
<td>1.4%</td>
</tr>
<tr>
<td>snow or slush</td>
<td>1.2%</td>
</tr>
<tr>
<td>sand, dirt, oil</td>
<td>0.1%</td>
</tr>
<tr>
<td>other, unknown</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Table 1.6: Percent of USA fatal crashes for different environmental conditions [2]

<table>
<thead>
<tr>
<th>environmental condition</th>
<th>percent of fatal crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>no adverse environmental conditions</td>
<td>98.1%</td>
</tr>
<tr>
<td>rain</td>
<td>7.5%</td>
</tr>
<tr>
<td>snow</td>
<td>1.6%</td>
</tr>
<tr>
<td>fog</td>
<td>1.4%</td>
</tr>
<tr>
<td>sleet or hail</td>
<td>0.3%</td>
</tr>
<tr>
<td>other, unknown</td>
<td>1.1%</td>
</tr>
</tbody>
</table>

The preceding example assumes that the driver modifies their speed for the prevailing conditions. It is a known fact that the human being represents the weakest link in the ‘driver-vehicle-environment’ system, and that up to 90% of all accidents can be attributed to human error [7]. In addition, of these accidents driver error is directly responsible 19% of the time, and is also associated to environmental conditions 50% and other vehicles 31% of the time [8]. In another study, driver distraction was cited as the cause for 18% of light vehicle crashes [3]. Similar results are also shown in Table 1.7.

Table 1.7: Percent contributions to traffic crashes [2]

It is the driver’s responsibility to ensure that the vehicle is driven at a safe speed for the road condition and visibility. In fact, the driver has a lot to do, and every decision made is a reflection of their ability and their assessment of the risk involved. L. Evans [2] labels these two elements as:

- Driver Performance – what the driver CAN do
- Driver Behaviour – what the driver DOES do

What the driver can do is a function of their driving ability and practice through a range of conditions, and can vary substantially from driver to driver. For instance, it would be expected that a racecar driver would have vastly superior “Driver Performance”
compared to the average driver because they have learnt how to predictively control a vehicle to its limit. “Driver Behaviour” defines how the driver chooses to use the skills they have. To a large degree behaviour is defined by the driver’s perception of the risks involved, and to what degree these risks are acceptable.

Racing car drivers, again, provide useful information about how driver performance and driver behaviour interact. For a long time it has been a widely held view that highly skilled drivers (including racing car drivers) are inherently safe on the road [2]. However, when looking at the statistics of some USA states, it becomes clear that racing car drivers receive traffic fines at much higher rates. They are almost 3 times as likely to be caught speeding, 1.3 times as likely to be issued other moving violations, 1.8 times as likely to receive non-moving violations and 1.6 times as likely to be in a reported crash [2]. Even though these drivers have the skills to drive much more safely than the average, they instead choose to drive more aggressively. It could also be said that racing car drivers accept a slightly higher risk whilst driving and match this to their skills by travelling faster. In fact, this effect is common among drivers, who generally use their increased skills to travel faster and more aggressively at a constant level of risk rather than travel normally with increased safety [2].

Such an observation presumes that the driver has made a calculated judgement of what level of risk is acceptable and what is not, and also if the driver has accurately estimated their own level of skill. However, the driver handles their car chiefly by feel and semi-automatic responses controlled by habit [1], and as such are not in a good position to assess their own skill and safety, as Figure 1.3 and Figure 1.4 show. Here it can be seen that drivers overestimate the skill and safety at which they drive. This is worrying, as the
two effects produce drivers that, on the average, think they drive with more care than they do, and with less skill.

L. Evans [2] supposes that the reason for this outcome is based on “cognitive dissonance”, in which people are likely to interpret additional evidence to support their beliefs. For instance, he states that reported traffic fatalities often confirm perceptions of driving superiority throughout the public, rather than highlighting risk. It is this ability of drivers to ill-interpret observed accidents and near misses that produces these overestimations. One suggested reason for this is:

Human beings are not designed to drive quickly. When man was created and developed, the opportunity did not exist. As a result, we have no innate fear of high speed. Precipices, on the other hand, have always existed and we are naturally afraid of great heights. These two factors are, in fact, the same thing. If you fall out of a window on the top floor of a three-storey building, you will be travelling at a speed of 50km/hr when you hit the ground. Everybody knows that it is dangerous to lean out of windows. The same instinctive protection is lacking in cars.

Vägverket [9]

This deals with how drivers can get themselves into emergency situations. Once a circumstance arises where immediate danger to life, limb and property are highlighted, the driver may become very much aware of the risks involved. It is also in these situations that the driver’s skills can come into play to avoid the accident, or at least mitigate damage.

A. Zanten et al [8] presents data that in critical situations, just before accidents occur, drivers initiated evasive manoeuvres 48% of the time ahead of all accidents, 50% of all collisions and 64% of all road departures. Likewise, M. Dilich et al [10] also state that the USA National Highway Traffic Safety Administration and the Fatality Analysis Reporting System shows that most drivers involved in accidents do not perform any avoidance manoeuvres. A reason that Dilich et al presents for this phenomenon is that:

...when one is confronted with a sudden peril requiring instinctive action...and...in the event that a driver of a motor vehicle suddenly meets with an emergency which naturally would overpower the judgement of a reasonably prudent and careful driver, so that momentarily he is thereby rendered incapable of deliberate and intelligent action...”

M. Dilich et al [10]

This basically means that drivers who do not possess sufficient skills in driving “at the limit” are liable to panic and react in an inappropriate way, which includes taking no action at all. In fact, M. Dilich also quotes “this phenomenon trends to affect cautious
drivers more severely because the accident situation is even further beyond their normal driving experience.” What more, during accidents, drivers also revert to their biological “emergency mechanisms”, which applied in an automotive environment produce undesirable effects as shown below.

When dangers, whether physical or psychological appear imminent, the ‘drives’ which influence behaviour become stronger and behaviour undergoes certain characteristic changes. Responses are more readily elicited. They tend to be more forceful, more extensive and more rapid, while at the same time they tend to be less regular, less organised and less coordinated. However, many of the dangerous situations which human adults meet require not vigorous activity but restrained, deliberate and accurate responses...

When a threatening situation arises which demands hard braking, swerving or both, most drivers lack experience to predictably and successfully handle their vehicles. The uncertain and potentially dangerous outcome of such aggressive handling may restrain drivers from fully utilising the capability of their vehicle’s control systems. Panic braking and swerving at high speed is uncomfortable to most drivers.

M. Dilich et al [10]

With this in mind, it is hard to imagine a situation where anyone is capable of handling their vehicle to its maximum performance limits in an emergency. In fact, E. Gohring [7] presents evidence that an overall improvement of driver reaction time of only half a second prior to an impending road accident would prevent 60% of rear end collisions, 50% of all collisions and 30% of all frontal crashes. Further, Table 1.8 presents other safety increases changes to driver behaviour would produce.

<table>
<thead>
<tr>
<th>Behaviour change</th>
<th>Risk reduction</th>
<th>measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>drive 5km/hr slower on urban road</td>
<td>50%</td>
<td>crashes</td>
</tr>
<tr>
<td>drive 10 km/hr slower than traffic on rural road</td>
<td>55%</td>
<td>crashes</td>
</tr>
<tr>
<td>decrease rural interstate speed from 101 to 92km/hr</td>
<td>34%</td>
<td>fatalities</td>
</tr>
<tr>
<td>drive at zero blood alcohol content compared to 0.13%</td>
<td>90%</td>
<td>crashes</td>
</tr>
<tr>
<td>not use a mobile phone</td>
<td>77%</td>
<td>crashes</td>
</tr>
<tr>
<td>wear a seat belt</td>
<td>42%</td>
<td>fatalities</td>
</tr>
</tbody>
</table>

Table 1.8: Risk reductions from changes in driver behaviour in USA [2]

It can be seen that very large safety advances can be made with seemingly small changes to the way the automobile is used. In particular, it is clear that when a driver is involved in an imminent crash they will find it very difficult to control their vehicle with a high level of deliberate and skilful control. If drivers fully understood the risks of automotive transport, as they naturally do of falling from a great height, they would genuinely seek to lower their actual risk to at least the level of their perceived risk.

Accounting for all of these points it becomes clear that drivers need support when their skills are not sufficient to control their vehicle as needed. Furthermore, in an emergency
even the most skilful of driver will not be able to achieve optimal performance due to their biological emergency mechanisms.

Stability controllers, such as anti-lock braking systems (ABS), traction control systems (TCS) and vehicle stability control (VDC), have been developed to provide this support at the limits of vehicle performance. In this way, ‘mistakes’ that the driver may make in an emergency are corrected by each of the systems to varying levels of ability. This provides higher vehicle performance than the driver may be capable if achieving, ensures that the vehicle remains easier to control, and reduces the likelihood and severity of driver panic.

Stability controllers are limited in their application through a number of factors. This includes limited control and scope of vehicle parameters (such as brake pressure) to achieve optimum performance, significant complexities in producing accurate controller algorithms and limited sensor data that leads to controller shortcomings (such as control limitations on dirt roads). These are very significant problems, which is evident in the 17-year development time between ABS and VDC introduction, despite many similarities in system design. There are still many advances that can be made in these respects, but it has been observed that the development time of these controllers grows exponentially with each increase in functionality. While we wait for these advances the world experiences hundreds of thousands of preventable deaths each year, so there is a clear need to bring these technologies to market now. Using traditional control methods this simply is not possible, and new technologies must be developed.

This investigation attempts to develop such a new technology. In particular, Artificial Neural Networks (ANN) will be used to develop a new kind of stability controller and to evaluate the ANN ability in road surface identification, with the latter representing one of the major problems in stability controller implementation.

ANN modelling offers a number of benefits over the traditional methods, and it is the principle goal of this thesis to determine if this potential can lead to actual application. New research will, therefore, be presented to develop and show the performance of these two systems on a real vehicle in real driving situations. Furthermore, the research is constructed so that many of its outcomes can be exported to future work that will encompasses more scope than the resources available for this investigation allow.
As such, this investigation contains a number of elements. Particularly, the ANN models must be developed to a point where they perform as necessary, which requires a thorough knowledge of their behaviour and significant model testing. They also require data from the vehicle for model development, alongside an ability to perform real-time stability control. As such, this investigation also includes the design, installation and programming of a comprehensive data acquisition and control system within the test vehicle, including the development of new ANN models that can be utilised in flexible real-time controllers. Furthermore, this installation also has the goal of providing a flexible test rig for future research within a number of areas, including the hydrogen conversion of the engine that immediately followed this work.

This thesis, therefore, attempts to advance the argument for ANN stability controller applications by evaluating the technologies in a real-life application, with the secondary goal of developing a test vehicle to provide flexible functionality above that required for this research. Furthermore, the work was planned to proceed through a number of discreet stages, which are reflected in the structure of the thesis. The “Vehicle Stability Background” chapter discusses past and present work into stability control applications, as well as presenting background information of various issues that are pertinent to the topic. It also advances the argument for the benefits ANN modelling can offer above other methods, and clearly states the how the investigation plans to achieve them. The following “Artificial Neural Networks” chapter then discusses the concept, design and considerations of ANN implementation, and provides details for the two types of ANN technologies that will be used to compare new and older ANN methods.

The initial development of the test vehicle is then presented in the “Chassis Sensors and Data Logger Installation” chapter, which details the hardware that was used to acquire the data for ANN surface identification. The following “Pavement Feature Recognition During Stable Driving Conditions” chapter then details the specific benefits ANN surface identification can bring, the ANN model programming and development process and the resultant model performance.

The installed hardware is then carried through to subsequent work with significant additions. In particular, the “Engine Management System Installation” and “Data Acquisition and Control System Installation” chapters detail this process, which sees the installation of an aftermarket engine management computer, vehicle based PC, data
acquisition and control PCI card and many other devices to enable comprehensive engine and chassis data acquisition and real-time ANN control of the test vehicle. This then leads in to the “Intelligent Closed Loop Traction Control” chapter, which details the controller development methodology, test track details and ANN controller performances. The final results of all of the work are then given in the “Conclusions and Future Work” chapter alongside a discussion of where this work might lead.
VEHICLE STABILITY BACKGROUND

Many aspects of the overall problem of vehicle safety have been presented, as have the concept of implementing control systems to aid the driver. This chapter carries on from these observations by discussing the actual benefits of driver assistance systems to place them in context. A detailed presentation of the complexities of tyre dynamics is given, based on the argument that the way the tyre grips the road fundamentally determines vehicle handling and performance.

The “state of the art” of vehicle stability is then presented with particular emphasis placed on road surface identification and non-traditional stability controller design. This then leads to a discussion on relevant directions this research should take, and what should be its principle goals. This includes a broad overview of how investigation testing should be carried out, how a positive result may be determined and where it might lead.
2.1 Systems that Assist Drivers

Aside from law enforcement and driver education, there are a number of systems that help the driver fill the gap between actual and perceived risk, as well as supporting them when their judgement fails or they are about to be involved in an accident. These systems have a high capacity to reduce injury and death on the roads by providing support to the element that is the direct cause of the majority of accidents – the driver. The effects can be dramatic with, for example, an estimate from J. Koopmann et al [3] suggesting that the number of crashes in the USA would reduce by up to 29% with the addition of a system that simply tells the driver that a slippery surface is approaching. Furthermore, if the same system also warned the driver of an appropriate speed (based on road curve shape, speed limit and surface friction) then the overall crash rate could be reduced by 41%.

Through these two examples it can be seen that by simply providing the driver with additional information and feedback on the way they operate a vehicle, driver assistance systems have the potential to massively increase vehicle safety, and already have done so. Systems that override driver controls when a mistake is made, however, can build on these benefits even further.

Of particular focus are systems that aid the driver during emergency manoeuvres, such as Antilock Braking Systems (ABS), Traction Control Systems (TCS) and Vehicle Dynamics Controllers (VDC). These systems will be discussed in detail, but have been in public use for some time and generally evaluate the driver’s responses with a view to provide assistance if the driver is not operating the vehicle optimally. By doing this, these systems reduce workload and stress of the driver, while obtaining greater performance and control. In emergency situations this has a twofold increase in safety because, not only does the vehicle performance improve, but the reduction in human workload and stress allows the driver to control the vehicle’s trajectory more precisely. As such, these systems artificially increase the “Driver Performance”, and produce all the associated benefits. In fact, Koopmann estimates that if these systems were incorporated into vehicles over and above the levels in 2002 the number of crashes in the USA would decrease by up to 42.5% generally, and up to 64.3% on freeways [3]. Sgt Peter Bellion of the Accident Investigation Branch, Victoria Police, Australia also identified similar figures in an interview in 2002.
Probably one of the most common type of crashes I will go to is a loss of control incident from oversteering. More than 50% of fatalities or major injury crashes are caused by loss of control. There are systems available now like dynamic stability control that can correct oversteer more quickly than the driver can. If that system was on every vehicle, I would say that a lot of these high speed loss of control incident could be avoided.


The figures given above state the maximum impact each of the safety system additions would give if universally adopted and 100% effective. This is clearly not the case in reality, as the use of seatbelts demonstrates. In 1974 seatbelts became standard for all cars with the USA, however 27% of drivers and passengers continue to not wear seatbelts. Further, this proportion of drivers represented 43% of all fatalities on USA roads in 2001 [2]. Clearly, even though wearing a seatbelt is effective at significantly increasing occupant safety, the lack of universal adoption has substantially effected its ability to reduce fatalities overall. Another interesting observation is that a 1% increase in vehicle speed has been associated with wearing seatbelts [12]. It is suggested that this is because the driver feels safer wearing a seatbelt and, because people tend to alter their driving style to produce constant perceived risk, subconsciously choose to drive at slightly higher speeds. As stated previously, a small increase in speed can significantly increase risk so, by altering driving style because the seatbelt makes the driver feel safer, the driver effectively trades some of that safety for higher speed.

In the case of seatbelts, this trade off easily errs on the side of safety, and throughout much of the world wearing them is a legal requirement. Many other systems, on the other hand, that show great theoretical potential in decreasing automobile risk but do not produce such clear benefits when applied to the ‘human’ element of driving. ABS is a good example.

The goal of ABS is to increase vehicle controllability during panic braking manoeuvres, with particular emphasis on providing steering control. Systems do this by automatically altering the braking force at each wheel to ensure the wheel does not slide excessively on the road, as will be discussed in detail later. Intuitively, such a system should substantially increase driver safety because a high level of support is provided in situations when the driver is least likely to operate the vehicle to its maximum capabilities. However, in the USA studies have shown that ABS has limited overall effect on crash risk, with a reduction of only 3%, and no significant effect on injury risk [12]. This is only part of the picture, nonetheless, and ABS has been shown to substantially influence the likelihood of crashes in certain situations, as follows:
2. Vehicle Stability Background

- (34±15)% lower risk of a pedestrian crash [12]
- (13±4)% lower crash risk on wet roads [12]
- (13±5)% lower crash risk when it is raining [12]
- (32±8)% lower risk of crashing into lead vehicle on wet roads [2]
- (30±14)% increase in risk of being struck in the rear on wet roads [2]
- (44±22)% increase in rollover crash risk [12]

These observations clearly show that ABS is of significant benefit when roads are subject to adverse conditions, such as rain, and the pedestrian safety is massively increased. However, because these accidents represent less than 20% of all fatalities, and considering that studies show that only approximately 45% of drivers actually activate ABS in limit braking manoeuvres [13], the net effect on fatality risk becomes small [2]. Another interesting observation is that the large reduction of risk of colliding with the vehicle in front is equalled by an increase of risk of being hit by the vehicle behind. It is supposed that this is because the braking performance increase of ABS helps the driver to avoid crashing into a lead vehicle, but this increase also places extra performance demands on the following vehicle. It can been seen that ABS alters the likelihood of the types of crashes vehicles may be involved in, but does not necessarily decrease overall crash risk greatly. Furthermore, this effect is exacerbated because of the changes in driver behaviour that ABS equipped cars produce. There is a sufficient evidence to show that drivers of ABS fitted cars travel at higher speeds, and they are more likely to ‘tailgate’ [2, 12]. In fact, it is supposed by L. Evans [2] that the associated increase in travel speed is directly responsible for the large increase in rollover crash rates, which grows sharply with speed. This effect is described below.

More efficient brakes on an automobile will not in themselves make driving the automobile any safer. Better brakes will reduce the absolute size of the minimum stopping zone, it is true, but the driver soon learns this new zone and, since it is his field-zone ratio which remains constant, he allows only the same relative margin between field and zone as before.

J. Gibson [12]

In this case, better braking can be used for increased safety or increased mobility. While the objective is to increase safety for automobile traffic, the natural reaction for many people is to use the extra performance for improved mobility. This observation has lead to the anecdotal solution for increased safety of “mounting a sharp spike from the steering wheel pointed directly at the driver”, which shows one extreme designing
vehicles to improve driver behaviour. In this case, the driver would see the very high and real chance of death in the case of an accident, and would drive much more conservatively as a result. The other extreme, of course, is to improve vehicle performance to the limit and leave the determination of an appropriate safety margin to the uneducated driver. It is supposed, however, that neither of these methods would increase safety, as the first relies on the construction of a wholly unsafe vehicle which leaves no scope for human error, while the last assumes that drivers can adjust their driving style to an acceptable level of risk (which has been statistically shown to be overestimated).

With this in mind, it is hard to imagine a reason why a driver would not prefer to drive a vehicle fitted with ABS, for example, over one without [2]. It clear then that, if the objective is to increase safety, appropriate education as well as advancement of technology should be utilised to adapt more closely with limited human capabilities and behaviour [7]. In this way, drivers can be aided to make conscious and subconscious decisions based on a realistic understanding of risk, and posses a vehicle that can perform emergency manoeuvres if required and protect its occupants in case of a crash. To avoid accidents, this requires thorough understanding and implementation of systems that complement and augment driver behaviour, as well as integrating improved vehicle performance.

Improved vehicle performance to aid the driver is a major facet for improving vehicle safety, and is a major research area in itself. To improve the performance of a vehicle, its dynamics must be well understood, including how they impact on the driver’s ability to control the vehicle. This understanding includes the dynamics of the entire vehicle system, as well as its elements such as tyres and suspension. When this is known it is then possible to develop systems that will increase performance further, including systems that provide the driver with control feedback and/or active assistance. Of particular interest is the tyre, as this is the element that fundamentally determines the dynamics and performance of the vehicle as a whole and provides driver feedback. As such, tyre dynamics will be discussed at length below.
2.2 Tyre Dynamics Fundamentals

In a certain sense the aeroplane is merely a device for exploiting the principles of the aerofoil; similarly, railroad traffic is dedicated to obtaining useful results from the flanged steel wheel; and in this sense an automobile is a device which makes use of the dynamic properties of pneumatic tyres for the general benefit of the community.

M. Olley [1]

It is easy to see that tyres are the fundamental element of the automobile, and that vehicle performance is greatly dependant on our ability to fully utilise them. Tyres are the connecting element between the vehicle and the road, and it is where they meet that ultimately defines vehicle performance and safety [14]. The tyres must provide the reactions against the road to provide braking, accelerating and cornering forces to the vehicle, which in turn affects vehicle dynamics and controls performance [15]. To fully utilise the tyre, and thus the vehicle, it is important to understand how this very complex object works.

By its nature, a pneumatic tyre is not solid, and deforms due to the weight of the vehicle where it meets the road. This deformed area is termed the “Contact Patch” and is shown in Figure 2.1. Here it can be seen that the normal force exerted on the tyre ($F_N$) produces a deformation within the tyre to produce the contact patch. Furthermore, it can be seen that the pressure the tyre exerts on the road is complex, and is a result of the tyre construction and sidewall characteristics. The tyre also deforms due to cornering forces, as photographed in Figure 2.2, as well as braking and accelerating forces, road irregularities and contact angles, which all further add to the complexity of the tyre/road union.

![Figure 2.1: Tyre contact patch pressure distribution [14]](image1)

![Figure 2.2: Tyre interaction with the road [16]](image2)
The pneumatic nature of the tyre causes further complexity, because it produces spring effects, as shown in a simplified manner in Figure 2.3. These effects offer great benefits for use in automobiles by increasing grip, increasing ride comfort and reducing noise. However, the complex spring behaviour that develops due to the construction of the tyre, in combination with the effects of dynamic load, also makes tyre behaviour very difficult to fully understand and to model.

![Figure 2.3: Tyre deformation and spring effects [16]](image)

The sections that follow will outline and discuss a number of important conventions, definitions and relationships to provide an overview of the basic principles behind tyre dynamics, and how they relate to vehicle dynamics.

### 2.2.1 Definitions, Coordinate Systems and Conventions

The Society of Automotive Engineers (SAE) has defined a number of conventions for use in automotive applications, as set out below. Figure 2.4 shows the Cartesian coordinate system used for the tyre, and defines a number of important terms. It can be seen that the inclination angle of the tyre (camber) is independent of the coordinate system. This means that the normal, longitudinal and lateral forces are always referenced to the same plane, regardless of tyre inclination. In the opposite sense, however, the directions of the longitudinal and lateral forces are always respectively parallel and perpendicular to the wheel plane, regardless of steering angle. It can be seen then, that the longitudinal force represents forces derived from braking and engine torque (via torque through the spin axis), while the lateral force consists of cornering forces. It should also be noted that some literature, while retaining this coordinate system, uses
different notation such as: Normal Force = \( F_N \); Longitudinal Force = \( F_L \); Longitudinal Force from Braking = \( F_B \); Longitudinal Force from Engine Torque = \( F_A \); Lateral Force = \( F_S \).

Figure 2.4: Tyre coordinate system and terminology [17]

The above coordinate system can be applied to all wheels on the vehicle (which is assumed throughout this study to be four). Each wheel, then, provides forces through the contact patch to the vehicle mass, the direction of which is referenced to the ground. The vehicle reacts to this combination of forces and accelerates in a particular direction as a result. Using classical mechanics theory, this acceleration can be considered the result of forces and moments applied at the centre of gravity of the vehicle, the coordinate directions of which are shown in Figure 2.5. Similar terminologies are used for the vehicle coordinate system, with side velocity often referred to as lateral velocity. Because the vehicle mass rests on the wheels through the suspension the vehicle, the vehicle system contains elements of “sprung” and “unsprung” mass. In this way, components that move with the tyre are considered unsprung, while parts of the vehicle that rest on the suspension are referred to as sprung. As a result of this, the vehicle coordinate system is reference to the vehicle body, with the yaw (often referred to as \( \beta \)), pitch and roll showing the body position in relation to the fixed road plane. [17]
2. Vehicle Stability Background

In this way, Figure 2.4 and Figure 2.5 define normal, longitudinal and lateral forces for individual wheels and for the entire vehicle. Likewise, the wheel and vehicle headings are defined along the longitudinal axes (x), and side velocity along the lateral axes (y). These are shown in a much clearer manner in Figure 2.6.

Of particular note is the wheel velocity vector \( v_\alpha \). This is the resultant of the tyre heading and side velocities, with the angle \( \alpha \) it forms termed the slip angle. When the wheel is travelling with purely longitudinal velocity this angle equals zero. However, if the wheel is producing lateral forces, then some side velocity is produced, and the slip angle increases. The mechanical process that produces this effect in the tyre is depicted in Figure 2.7. It can be seen that the lateral force causes the tyre to deflect and the contact patch to move. As the tyre rotates, this results in some lateral movement, producing the slip angle. As a result it can be seen that, so long as traction is maintained with the road, as lateral forces increase so does slip angle. [18]
In the same way that lateral force cannot exist in a rotating pneumatic tyre without a slip angle, longitudinal force can only be produced through longitudinal slip. This slip is defined as the difference in the actual rotational speed of the tyre ($\omega$) and its estimated free rolling speed ($\omega_o$), as shown in Eqn 2.1. When either braking or engine torque is applied to the wheel the contact patch deforms, as shown previously in Figure 2.3. This elastic deformation causes the contact patch to move longitudinally, compressing the tread elements on one side only to return to their original shape at the other. In this manner the contact patch slides due to its deformation, producing slip. The tyre cannot produced any longitudinal forces without the presence of slip. [17]

\[
\text{Slip} = \lambda = \left(\frac{\omega}{\omega_o}\right) - 1 \quad [18]
\]

Eqn 2.1

Contact patch deformation defines, to a large extent, the way in which forces are transferred from the tyre to the road. In this respect, any parameter that alters the way the contact patch interacts internally or with the road will significantly affect performance. As has been shown above, camber is one such variable. By tilting the wheel as shown in Figure 2.3 pressure is distributed through the contact patch in a different manner, and tyre performance is significantly altered. In particular, negative camber (where the top of the tyre is closer to the vehicle centreline) produces a lateral force component called “camber thrust” that aids in cornering but reduces possible longitudinal forces.
Steering angle is also an apparent, and obvious parameter that affects force transmission. Steering a wheel changes its heading direction and creates a slip angle, which in turn produces a lateral force. This force is then transmitted to the vehicle, and used by the driver to ensure the vehicle follows its intended path. As a consequence of this, the longitudinal velocity of the vehicle is complemented with elements of lateral and yaw velocity. The resultant linear velocity produces an angle ($\beta$) with the longitudinal axis, and is called the “side slip angle” or “float angle”. [14]

2.2.2 Experimental Relationships

When considering the tyre individually, there are many factors that may affect its performance at any instant. Many of these parameters are non-linear, and are often difficult to measure and model. As a result, it is possible that a small change in one value may produce large, and unforeseen, performance variation. Some of these parameters are listed below [19]:

- Road surface condition ($\mu$)
- Tyre normal load ($F_N$)
- Tyre dynamic effects
- Tyre construction
- Tyre temperature
- Tyre pressure
- Tyre velocity ($V_x$)
- Tyre slip angle ($\alpha$)
- Tyre slip ($\lambda$)
- Tyre camber ($\gamma$)
- Tyre wear
- Tyre wear

This investigation has a very strong focus on measuring and determining what is happening at the contact patch, and as such, it is important to identify prominent performance relationships for future discussion. The section that follows presents a number of key trends to show how some of the above parameters affect performance, and in particular force transmission. Furthermore, because tyre dynamics are dependant on a great many parameters, all graphs and values given should be considered general in nature. In this respect all figures and tables are supplied to provide information on trends, and specific values should be disregarded. In addition, many texts quote either force ($F$) or coefficient of friction ($\mu$) when describing the tractive behaviour of the contact patch. Due to the relationship $F = \mu F_{\text{normal}}$ it can be considered that force and coefficient of static/dynamic friction are proportional [18], and therefore the two terms are used interchangeably throughout this document as appropriate.

Many of the relationships presented here are not possible, or extremely difficult, to measure on a vehicle. This is partly because it is not possible to keep ‘everything else
constant’, and also because some things cannot be measure directly on a road. In this respect, much of this data, although coming from a variety of sources, originate from machines such as the one presented in Figure 2.8. Of particular note is that all measurements, unless otherwise stated, are completed in equilibrium conditions – eliminating the extremely complex problem of dynamic loading.

Comprehensive data on lateral and longitudinal force as a function of slip angle and slip ratio is relatively rare. Few facilities are available to run a comprehensive set of these tests, which are time consuming and costly.

W & D Milliken [17]

Firstly, Figure 2.9 presents information on how well the tyre can grip a tarmacadam road in different environmental conditions, at different levels of tyre wear and at different speeds. Here the coefficient of static friction is used to determine the maximum coefficient of adhesion [14]. A number of trends emerge, and it can be seen that generally, maximum tyre grip decreases significantly with road surface condition left to right and that tyre grip decreases with increased speed. It can also be seen that a worn out tyre actually performs better than a new one in dry conditions, but much worse when the road is wet. This is because when load is applied to the tyre in dry to damp conditions, enough heat is produced at the contact patch to boil away any water that may be present through humidity. A worn tyre presents more tread to the road than a new one (in the same manner as a racing slick) and more grip results so long as this water can be dissipated. However, when there is water present on the road, the heat produced cannot boil all the water away, and must instead clear the bulk of the water away. This is what
2. Vehicle Stability Background

the tyre tread is for, and also why a worn tyre performs poorly in the wet. Nonetheless, it can be seen that the tyre grip at high speeds in 1 or 2mm of water is very poor regardless of tyre wear. This is because hydroplaning has occurred, and the tyre is making little to none contact with the road [20].

Figure 2.10 and Figure 2.11 show the effects of wheel slip on longitudinal force and slip angle on lateral force respectively, and also the effect of different road surface conditions. These two graphs are very important from a conceptual point of view, because they relate many significant parameters together. As stated above, the driver has control over slip angle of the steered wheels though the steering system (which in turn affects the non-steered wheel slip angles). Likewise, the driver has control over wheel slip through the brake and throttle pedals. As such, slip and slip angle can be considered ‘input’ parameters. Longitudinal and lateral forces, conversely, represent the result of the given operating conditions, and can be thought of as ‘output’ parameters. Factors such as road surface, wheel camber, tyre type, etc, can be considered as ‘state’ parameters. Figure 2.10 and Figure 2.11, therefore, show the input/output relationship of individual tyres for different road surface states – which has already been identified as a significant parameter effecting performance.

![Figure 2.10: Longitudinal tyre friction as a function of slip for different road surfaces][21]
In general, both of these figures follow the same basic trend with increasing slip or slip angle, and are often mirrored in negative and positive force directions. At values near zero, it can be seen that the longitudinal force/slip and lateral force/slip angle is essentially linear – and is termed the “linear zone” or the “elastic zone”. The next region, at increased magnitudes of slip or slip angle, is bounded by the point where the curve becomes non-linear and the point that represents the curve peak (producing maximum force). This region is called the “transition zone” and the curve peak defined here as the “critical slip” and “critical slip angle”. Finally, the region at further increased slip or slip angle magnitudes that extends past the critical slip and critical slip angle is denoted as the “frictional zone” or “unstable zone”. [16]

In general, most street driving is completed within the linear zone because very high longitudinal and lateral forces are rarely required. Furthermore, in this zone a simple input/output relationship exists and the driver can operate the vehicle confidently and predictably for the given road surface. As higher forces are required, a wheel may progress into the non-linear transitional zone. In this zone, increased slip and slip angle produce less force than the driver may predict, and driver control becomes less confident. As the slip or slip ratio increases further, the critical value may be reached. At this point the tyre in producing the maximum force in a specific direction, and represents the performance limit of the tyre. It is this value that racing car drivers try to maintain for improved lap times. Past the critical value, the curve exhibits a negative slope within the frictional zone, which means that increasing slip or slip angle will produce a reduction in
transmitted force. Excursions into this zone are wholly undesirable, as performance is reduced and the vehicle handles much less predictably. Furthermore, the slope direction produces a negative feedback loop which is very difficult for the driver to control. For example, under pure longitudinal acceleration the driver may increase throttle position through the linear and transitional regions, and past the critical slip. When the tyre passes the critical slip it can transmit reduced force to the road, and the difference in drive torque is used to accelerate the wheel further. As such, even a small excursion into the frictional region can result in massively increased slip, and resulting performance loss. Furthermore, to regain adequate slip, the applied drive torque must be reduced significantly – again reducing overall performance. [17]

Focussing on the effects of differing road surface condition, it can be seen that tyre performance is greatly influenced by road characteristics. Different surfaces have different slopes in the linear region, transitions regions differ in size and shape, critical points occur at a large range of slips, slip angles and loads, and frictional regions have varying slopes and shapes. Further, it can be seen that on loose surfaces the transitional zone can extend indefinitely, with a critical value never being reached.

![Figure 2.12: Lateral tyre friction as a function of slip angle for different tyres [18]](image-url)

Figure 2.11 can be compared to Figure 2.12, which shows the effect of racecar tyre construction on the lateral force/slip angle curve. It can be seen that different tyres can have a significant effect on performance in this regard. In addition, the street tyre is designed to lower levels of performance, but also to operate efficiently at higher slip
angles. In this way, the street tyre aids the street driver in avoiding excursions into the frictional zone.

Of course, longitudinal and lateral forces rarely operate independently on a automobile, and the above graphs represent idealised cases. In practice, increased longitudinal force is normally achieved through a sacrifice in lateral force, and vice versa. Figure 2.13 and Figure 2.14 shows this relationship by depicting longitudinal force and lateral force against slip angle and slip. They also separate longitudinal force into driven (traction) and braked components. Starting with Figure 2.13, slip ratios of greater than 0.6 and less than –0.6 extend past the critical slip. Furthermore, increased slip angle results in decreased longitudinal force at constant slip, and this effect is non-linear and different in acceleration and braking. Likewise, Figure 2.14 shows that increased slip angle results in increased lateral force at constant slip, and that this effect is again non-linear and different for acceleration and braking. Further, increased slip ratio significantly reduces lateral force transmission.

Figure 2.13: Longitudinal force as a function of slip angle and slip [17]

Figure 2.14: Lateral force as a function of slip angle and slip [17]

Figure 2.15 shows similar results rearranged along a force/slip axis, in which case the trade off between longitudinal (F<sub>x</sub>) and lateral (F<sub>y</sub>) force transmission with increasing slip is clear. In this case, if maximum longitudinal force is to be achieved, then very little lateral force can also be provided. Likewise, if a high slip angle is used, the longitudinal force will be limited but lateral force will increase. This is an important
relationship, as it clearly shows the generic association between driver inputs and vehicle force outputs for any condition.

![Figure 2.15: Long. and lat. forces as functions of slip and slip angle [23]](image)

![Figure 2.16: Resultant force as a function of resultant slip [17]](image)

In reality, however, longitudinal and lateral forces and velocities are conventions that are elements of a single force vector and velocity vector applied at the contact patch. Figure 2.16 shows the relationship between these resultant magnitudes for a particular tyre at constant pressure speed, temperature and normal load. Further, the same general curve that was evident for longitudinal and lateral vector components is also relevant with their resultants, and is an important observation. In this way, the theory of linear, transition, critical and frictional zones can be applied to all tyre loads regardless of vector direction.

Another important parameter in dealing with the amount of force transmission is normal load. As the theory suggests, forces in the road plane are proportional to the normal load and the road coefficient of friction, as defined in Eqn 2.1. This is a useful but simple approximation only, and does not accurately describe the behaviour of a pneumatic tyre under normal load. This is because Newton’s Laws of Friction applies to friction between smooth bodies, whereas the tyre grips the road through mechanical grip and transient molecular adhesion. This effect is shown in Figure 2.17, where it can be seen that the coefficient of friction reduces with increased vertical load. The resultant effect of load of tyre force is then shown in Figure 2.18. This departure from the linear relationship Newtonian mechanics would produce is referred to as “tyre efficiency”, in the manner that a lightly loaded tyre will transmit road forces more efficiently than a high loaded one.
A non-linear relationship also exists between normal force and slip angle, as demonstrated in Figure 2.19. Here, the lateral force/slip angle curve presented in Figure 2.12 is repeated for different normal loads. The peak force, and corresponding friction coefficient, is also shown for each curve – highlighting the effect of tyre efficiency. The important observation here, however, is that the curves are not linearly related, as shown by the varying peak force slip angles. For this tyre, increasing load results in decreased slip angle for maximum force transmission, but this relationship is highly dependant on tyre construction. In fact, tyres often have a relationship that is inverse to the one shown, and it can be seen that tyre non-linearity is compounding and exists on many levels.
2. Vehicle Stability Background

There are many more parameters that effect tyre performance but, due to the volume and difficulties in obtaining reliable data, they will not be covered here. Instead, two examples are shown in Figure 2.20 and Figure 2.21. As demonstrated, the lateral force is greatly dependant on camber, in a multi-dimensional non-linear fashion. The other figure shows the effects of hydroplaning on a very wet road, expanding on the observations of Figure 2.9. In this case, the force/velocity relationship is linear below a specific speed, and then reduces and becomes non-linear. Here it can be seen that hydroplaning has the capacity to reduce tyre grip to zero, is highly velocity dependant, and contains non-linear regions.

![Figure 2.20: Lateral force as a function of camber [18]](image1)

![Figure 2.21: Maximum cornering force on a very wet road as a function of speed [20]](image2)

These observations highlight the non-linearity of tyre dynamics very strongly, the effect of which is particularly severe near the performance limits of the tyre [24]. Further, because there are a vast number of non-linear parameters that affect tyre dynamics, these non-linearities compound to produce a complex multidimensional problem. This makes modelling tyre dynamics very difficult.

It is useful to remember, however, that despite all of the parameters that can change tyre performance, the maximum amount of force the tyre can transmit in any given direction in the road surface plane is finite. This “optimum” tyre performance will then be when all of the tyre variables allow the maximum possible force in the desired direction. Likewise, the performance of the entire vehicle is subject to the resultant force vector produced by each wheel, and can be made “optimum” by optimising each tyre force in the desired direction. These limits to vehicle performance are often depicted as the “Performance Circle”, in which case all vehicle accelerations are recorded and plotted, as shown in Figure 2.22. In this way the acceleration limits of individual vehicles can be
evaluated, and the effects of tyre limits compared to other vehicles. In the case of the two vehicles depicted, it can be seen that the luxury vehicle has better braking performance than cornering, and is severely limited in forward acceleration (probably due to a lack of engine power). Likewise, the performance vehicle is limited in forward acceleration, but shows a much rounder performance circle. Here the car can corner at the same acceleration as it can brake, albeit at much higher accelerations that the luxury vehicle.

Nonetheless, it should be pointed out that all of the data presented above represents steady state effects only, and does not deal with transient effects at all. In fact, very little information about tyre transients is available at all, which is presumably because they are very difficult to observe and determine meaningful relationships for. Some insight is given into this area in Figure 2.23 however, which repeats the longitudinal force/slip curve portrayed in Figure 2.10, but includes the transient effects during the test run. Here, all parameters are kept constant, but it can be seen that the tyre force does not follow the simple curve suggested before. This is because the transient effects as the tyre is progressively braked cause significant and short-lived changes to the internal tyre stresses and contact patch pressure and deformation. This is the result of the elastic deformation of the tyre sidewalls and the contact patch and, as such, the information presented here represents the ‘tip of the iceberg’ insofar as true tyre dynamics are concerned, as D. Dennehy et al comments. [26]

The relationship between the vehicle forces and the behaviour of the tyre in the contact patch is highly complex and not fully understood, yet this relationship is critical to the vehicle dynamics and control performance of a vehicle.

D. Dennehy et al [15]

This highlights a critical, and often underrepresented, fact about the way tyres are modelled. Almost all tyre research and modelling have focused on the performance of a
tyre in steady state conditions, even though this does not represent real-world driving conditions.

![Figure 2.23: Longitudinal force as a function of slip in a dynamic situation [26]](image)

It can be seen in Figure 2.23 that transient effects can account for very large variations in grip at specific slips alone, yet little research has been carried out to determine what these are, or to develop empirical relationships for them. Of further consequence, because little is known about them, tyre transient effects are often ignored in tyre simulation models. This could be particularly problematic where vehicle simulation models are used to provide performance evidence of particular design or active control system performances, and where tyre models are directly used in active control system logic. Of particular importance is that, by ignoring transient tyre effect, current active control systems are limited in their ability to forecast tyre grip. This means that the closed-loop control outputs may be erroneous, introducing excessive “hunting” for the optimal value and reducing vehicle performance. As such, determination and incorporation of transient tyre effects into any experimental study, empirical investigation or process simulation has potential in increasing vehicle performance.

### 2.3 Contemporary Vehicle Stability Control

Above the critical speed the disturbing force $F$ may be made infinitely small, but the car will still swerve, just as surely as an egg standing on end will fall over. Only the skill of a tightrope walker on the part of the driver holds such a car on the road at speeds above the critical.

M. Olley [1]

For the most part, controlling vehicle dynamics and stability is the responsibility of the driver, as shown in Figure 2.24. The driver evaluates the road ahead, the traffic and environmental conditions and, based on feedback of the motion of the vehicle through
experience, executes control actions through the brakes, throttle, gears, clutch and steering. The performance of the vehicle is then defined by these control parameters and state variables such as the road at the contact patches, and wind.

![Driver-vehicle system block diagram](image)

**Figure 2.24: Driver-vehicle system block diagram [20]**

The quality of the “driver transfer function” correlates to the skill and experience of the driver, as well as their ability to utilise the control outputs. This has three limitations. Firstly, if the vehicle enters a state that the driver does not have the skills to adequately control, such as during a panic cornering manoeuvre, the driver transfer function will be flawed, as discussed by M. Olley. Secondly, the number of variables that the driver is able to control is limited by the number of physical controls that can be manipulated by the human body, in this case the pedals and the steering wheel. This is a significant limitation because, for example, if the driver wants to brake one wheel only, this is not possible. Thirdly, the human body is not capable of high frequency deliberate control, limiting the speed and effectiveness of the closed-loop block diagram above. Clearly, the effectiveness of the driver-vehicle system can be improved to a significant extent by improving the driver transfer function in these respects.

Electronic control offers a solution. It can actuate many different controls, such as independently braking each wheel, and it can operate at much higher sampling speeds than the human body is capable of for deliberate control. Further, the control logic can be programmed to evaluate the vehicle dynamics and aid the driver in situations where they require assistance. In this way, the electronic control can eliminate a range of driver shortcomings and significantly increase driver performance and augment undesirable driver behaviour. As discussed previously, and depending on how the technology is utilised, this also has a significant bearing on safety by actively assisting the driver to
avoid accidents. Such systems are called “Active Safety Systems” and operate in the accident avoidance zone shown in Figure 2.25. The goal of these systems is to return the vehicle and driver back to “normal driving state” whenever a loss of control incident (defined as when at least one tyre has exceeded the coefficient of friction critical peak [25]) is imminent or has occurred, in order to avoid an accident. It is important to note that any safety actions that occur in the “normal driving”, “warning” and “collision avoidable” states reduce the probability of a collision [27]. This is in contrast to scenarios when an accident cannot be avoided, where injury mitigation is the goal. In these cases, safety becomes the responsibly of “Passive Safety Systems”, such as seat belts and airbags. [14]

![Figure 2.25: Integrated safety system state diagram [27]](image)

Active safety systems attempt to assist driver commands to ensure that a normal driving state is maintained. This includes providing normal driving assistance, such as adaptive cruise control, navigational aids and traffic and road condition information. It also includes providing additional warning assistance to identify risks and inform the driver, including systems such as lane departure, blind spot and low tyre pressure warnings. Lastly, it includes attempting to completely avoid the accident when a collision is imminent. Although this also covers vehicle initiated and controlled systems, such as rear-end collision avoidance braking, it has significant emphasis on assisting the driver by controlling aspects of the dynamics of the vehicle. From the driver’s perspective, this translates to simple predictable control up to the performance limit of the vehicle. [27]

Electronic control systems, however, have limitations, and their overall performance is governed by the quality of the complete system. This includes the number and quality of input sensors and output actuators, the control logic utilised within the electronic control unit and the sampling and control rates. Such elements are governed by the cost of
development and installation and, as such, many performance gains are not yet realised in ordinary road vehicles. As technology increases, component prices reduce and improved automotive safety becomes economical to the consumer, active systems grow in performance and complexity. In 1978, “Anti-lock Brake Systems (ABS)” were introduced to control wheel braking force during panic manoeuvres, followed by “Traction Control Systems (TCS)” in 1987 to control driven wheel force under acceleration and, most recently, “Vehicle Dynamics Control (VDC)” in 1995 to improve steering control in critical cornering. To date, these commercial systems almost exclusively attempt to improve performance by controlling wheel slip only, as shown in Figure 2.26. Further, they utilise as few sensors as possible, make many assumptions and are very limited in determining driver intention. Clearly, there is massive scope for system improvement as these elements become better understood and implemented.

Future systems will involve more sensors, more control outputs and more advanced control logic. These are not simple developments, as the 17 year gap between ABS and VDC suggests. Inclusion of greater sensor arrays in vehicles and providing the means to control more chassis variables is an expensive process, and prices must become economically viable before they will be adopted by the automotive industry. Likewise, software and algorithm complexity is set to increase with more advanced systems. This aspect is particularly problematic because the inclusion of additional sensors and control outputs exponentially increases algorithm complexity, also exponentially increasing controller memory, processor requirements and software debugging time. [28]

The following sections will present the control and performance aspects of the current mainstream active safety systems ABS, TCS and VDC specifically, as well as briefly presenting other systems and future possibilities.
2.3.1 Anti-lock Braking Systems (ABS)

ABS was one of the first active safety systems to be installed on commercially available vehicles, and was introduced by Bosch in 1978. The system was designed to assist the driver in panic braking by preventing wheel ‘lock up’, thereby increasing vehicle stability and control. ABS works on the principle that during panic braking the driver will most likely be required to alter the vehicle’s course, such as in the case of avoiding an unexpected obstacle or an oncoming vehicle. This means that the tyres must be held at a particular slip ratio that provides a reasonable proportion of lateral force while supplying sufficient longitudinal (braking) force. As discussed above, lateral force can only be achieved through the sacrifice of available longitudinal force, so manoeuvrability is gained at the sacrifice of increased stopping distance. As such, Mathues [29] states that the required slip ratio at each wheel must reflect a compromise between these two elements on a variety of road surfaces and environmental conditions. Figure 2.26 shows the general ABS control range.

In order to maintain safe handling characteristics of vehicles, designers strive to maintain consistent, predictable vehicle response to driver steering inputs in the entire range of operation. Unfortunately, because of the particular shape of the tire … force characteristics, there exists two profoundly distinct kinds of vehicle handling behaviour (the linear and the non-linear).

A. Hac et al [22]

The techniques that manufacturers use to achieve this result are varied, but all are based on the same general arrangement. A number of wheel speed sensors are installed on the vehicles that, as well as giving individual wheel speeds, are used to infer vehicle speed. With this estimate, the longitudinal slip ratios ($\lambda_B$) at each wheel can be calculated within the ABS Electronic Control Unit (ECU) and compared with the desired values. If the slip ratios are too high (inferring that a wheel has locked up or is not providing enough lateral force) the ECU will then reduce the brake pressure to the effected wheel(s) by a specified step. This results in an increase in wheel speed and a reduction in slip, which is re-measured, completing the closed loop control function. Further, ABS controllers intentionally cause tyre slips to cycle across the region of peak friction to gain the required information for estimation of vehicle speed and to gain feedback data [30]. Such a system is shown in Figure 2.27 with the addition of an acceleration sensor. Furthermore, Figure 2.28 shows an example of the operation of the brake control hardware, in which brake pressure can be modulated electronically or controlled by the driver. In this case the driver is given full control over the brakes until the ABS
activates. When this happens most control systems will isolate the driver and hold or decrease brake pressure, with the driver returned to control if increased brake pressure is required. In the diagram below, however, this is taken further and the system is capable of also increasing pressure.

![ABS control diagram (with accelerations sensor)](image1)

Figure 2.27: ABS control diagram (with accelerations sensor) [31]

![Brake modulator operation example](image2)

Figure 2.28: Brake modulator operation example [32]

The effect of the inclusion of more sensory information can be seen through the addition of the acceleration sensor. Saito et al. [31] states that the inclusion of a longitudinally placed acceleration sensor can further increase ABS performance by providing the ECU with data as to whether or not the wheel speed sensors are giving an accurate description of vehicle speed. This is because absolute vehicle speeds can be hard to obtain using just wheel speeds sensors alone. Due to slip at each tyre during braking, the predicted vehicle speed is always slightly lower than the actual speed. Also, when a wheel starts to lock, it will further reduce the predicted speed until the system chooses to ignore it from the vehicle speed calculations. To illustrate this, consider a four-wheel drive (4WD) travelling on an ice-covered road. The front and rear axles are connected so the possibility of all four wheels slowing by equal amounts is much higher than in two-wheel drives (2WD), resulting in highly erroneous vehicle speed estimates and thus slip ratios. The likelihood of simultaneously locking all four wheels is also high and both situations result in severely reduced ABS effectiveness [31]. Therefore, increasing the amount of input data can significantly increase performance. Further, many ABS arrangements may not provide the independent wheel speed measurement or brake modulation to each wheel, as shown in Figure 2.29. It is common, for example, to measure differential speed, rather than two driven wheel speeds. Likewise, brake pressure may be modulated through the front/back brake circuits only. In these
situations ABS performance is degraded because the control ECU must base control decisions in reduced information and with fewer controllable variables. [33]

The fact that ABS measures the vehicle’s wheels speeds only means that the information base it works from is somewhat limited. At low speeds Strickland et al. [34] found that ABS can result in reduced deceleration when compared to the completely locked wheel scenario it is designed to avoid. Factors such as road surface, suspension travel, steering angle and vehicle yaw rate are not included in the control model and so ABS makes a number of control assumptions. These assumptions are based on ‘normal’ driving behaviour, with slip control in the vicinity of 5 to 15% [25]. Deviations from these conditions, such as when driving on an unsealed road, lead to less than optimum brake control as the control logic fails.

This limitation can be shown through example by referring to Figure 2.10. By approximating an unsealed road to dense sandy soil we can see that the optimum slip for braking is at about 40% in this case, while the optimum slip for a dry or wet road is at about 20%. Since the ABS has little idea of the road surface it may not allow the wheels to slip further than 20% when activated, which in this case reduces the braking force by about a half, increasing stopping distance dramatically. In this case the ABS has a significant negative effect on braking to produce small stability gains, which is an undesirable mode of operation. Further, the cyclic nature of the ABS wheel slip control, from one side of the critical slip to the other, results in reduced braking performance and the introduction of very large transient forces within the tyre.

In summary, the subjective performance of ABS with regard to vehicle safety has been presented above. Particularly, it was found that ABS equipped cars experienced different types of accidents to non-ABS equipped cars, although crash rates remained essentially
constant. While many aspects of safety were improved, the resulting change in driver behaviour reduced the overall effects by reducing safety in others – namely rollover accidents. Aside from the arguments presented above, this can be partially explained by considering the example of a car that has headed off the road towards a tree. A car without ABS may not be able to manoeuvre around the tree and end up crashing into it, while the car with ABS may avoid the tree and continue on into the off-road terrain, with consequent risk of rollover. In this case the ABS equipped vehicle has avoided a serious accident, but by doing so may have converted this impact accident into another type of accident. In this regard, it is useful to consider that the main goal of ABS is to enable inexperienced drivers to control their vehicle predictably and precisely under panic braking, by effectively helping them to emulate the driving style of experienced drivers.

In this way, ABS contributes to controllability and performance while braking to a significant extent. Heavy braking, however, is only one of the functions through which a vehicle may become unstable and difficult to control. Oversteering through excessive throttle, whereby a vehicle might ‘spin out’ and leave the road, is another potential source of accident. In this case, control of the driven wheels of the vehicle can be used to avoid excessive longitudinal slip under acceleration, and thus preserve lateral force transfer to avoid oversteer. Such a system also has significant benefit on slippery roads, whereby it can be very difficult to gain enough traction to accelerate up a hill, or when the coefficient of friction at each wheel is reasonably different (split $\mu$).

### 2.3.2 Traction Control Systems (TCS)

Traction control systems were introduced to help control a vehicle during acceleration manoeuvres, when excessive engine torque may be applied to the wheels resulting in reduced longitudinal and lateral force transfer. The first system was launched in 1987 by Bosch in the interests of optimising both the available longitudinal and lateral forces generated by the vehicle’s tyre. This is a very similar goal to ABS, except under acceleration rather than braking, and the control logic is very similar [35]. The potential benefits of TCS can be summarised as follows [7]:

- Enhanced driving in straight line running and cornering by maintaining the tyre forces within their optimum slip ranges;
- Higher traction forces can be transmitted to the road when moving off from stationary and when accelerating;
Vehicle stability can be compromised in a number of ways during an acceleration manoeuvre. If traction while accelerating is broken in front wheel drive (FWD) vehicles, the front tyres are no longer able to produce significant lateral force and the driver loses steering control. This manifests itself in vehicle under-steering. If traction while accelerating is broken in rear wheel drive (RWD) vehicles it is the rear tyres that cannot provide enough lateral force, giving the vehicle an over-steering attitude and the possibility of spinout as yaw stability is lost, as shown in Figure 2.30. Also, the widespread use of ‘open’ differentials, which can only deliver equal amounts of torque to each driven wheel, has a negative effect on vehicle stability during acceleration on surfaces of varying levels of friction coefficient. If one wheel is travelling on a surface that offers limited traction and starts spinning excessively, the torque that can be transmitted through the other wheel is severely reduced regardless of the level of \( \mu \) it is travelling on.

Just as with ABS, there is a clear need to precisely control wheel slip in the interests of maximising longitudinal and lateral tyre forces for a range of situations, and therefore it relies on data gathered from wheel speed sensors. Wet and slippery roads on a gradient and cornering can create critical situations and place excessive demands on the driver, and incorrect reactions can result. TCS can intervene in such situations and optimise stability to an extent that is beyond the abilities of the driver. In most production vehicles it is this demand for stability that is the overriding function of TCS, and when
activated, the controller attempts to keep the tyre slip in the range shown in Figure 2.26. It is noted, however, that incorrect operation on some roads is a common TCS problem because of the large variation in critical slip for different surfaces, and all traction controllers incorporate methods to deactivate slip control [37]. In any case, the application of TCS should utilise the following information for optimum operation in the interests of producing a control strategy based on directional control, traction and steerability [38]:

- Vehicle speed, to give traction priority at low speed and directional control priority at high speed,
- The speed difference of the non-driven wheels to detect cornering manoeuvres,
- Vehicle acceleration and throttle position to identify situations where TCS operation is sensitive to small changes.

In contrast to ABS, TCS can be activated through a wide range of control options. The simplest method of passive traction control is the use of “Limited Slip Differentials (LSD)”, which offers improved traction by proportioning driven wheel torque across the differential [39]. This aspect of control can be made active by utilising differential locking control, although this type of control is rare. More active control options are available through the brake controller in Figure 2.28, which can be used to reduce excessive slip and also emulate the effects of an LSD. Furthermore, engine torque control can be used to limit slip, with torque limited through throttle valve intervention or by retarding or preventing combustion in specific cylinders using fuel injection and ignition control. Each system can give reasonable results under specific conditions, but can be combined with other systems to build on the advantages of each to provide comprehensive and comfortable control, as shown in Figure 2.31.

To elaborate further, limited slip differentials can proportion the torque developed by the engine between the driven wheels on surfaces of uneven friction between tyres (called “Split µ” surfaces). On an open differential the forces produced at the low µ driven wheel determine the force developed by both tyres, and represent a significant limitation. When one wheel starts to slip excessively or torque is developed unevenly, however, the LSD will effectively lock the differential and avoid this effect. In this case the force developed at the low µ wheel no longer limits the force at the high µ wheel, and greater acceleration can result. This also has the effect that the low µ wheel will not slip a
massive amount, increasing lateral grip. This is also the case for electronically controlled locking differentials. The drawback in these cases, however, is that if the engine produces too much torque, or the surface is sufficiently slippery, both wheels may start to lose traction and spin. Therefore, there is a clear need to limit the torque delivered to the wheels.

Ordinarily the driver reduces drive torque when needed by releasing the accelerator slightly. This requires a level of skill and inexperienced drivers may ‘spin out’ in situations when more experienced drivers could have controlled the vehicle. Clearly, TCS can be used to help the inexperienced driver control the accelerator to improve traction in this regard. Throttle control for TCS can be achieved in two ways, fly-by-wire control or the addition of a second throttle valve. Fly-by-wire removes the standard cable link between the pedal and the butterfly valve, replacing it with a potentiometer at the pedal, which controls a servomotor at the butterfly valve. This method enables the TCS to control the throttle when needed though the engine ECU, but has the potential of failure resulting in an uncontrollable throttle, and so requires many safety considerations. The second throttle control places a secondary, electronically controlled, butterfly valve upstream of the valve controlled mechanically by the driver, as Asami et al. [41] shows. In this case, when the driver opens the throttle too much the second valve can close a little and limit the airflow to the engine, reducing its power. The advantage of this method is that failure of the system can at worst stall the engine, while also enabling the relatively simple addition of TCS to vehicles without an engine ECU or fly-by-wire technology. This system offers a number of benefits and, by reducing driven wheel torque, can improve stability considerably. It also has the advantage that steering control is not affected by the TCS operation, as it is under the operation of an LSD. Its operation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stability</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Traction</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>System</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Complexity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Operational Time</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Response</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 2.31: TCS concept evaluation [40]
is very smooth and noise free, and as such makes the systems very driver friendly. It does, however, suffer from comparatively slow response, which is a significant limitation.

The slow response in the use of the throttle to control engine power can be overcome by replacing the throttle control element with ignition / injection control. By progressively altering the way engine cylinders fire, this system provides almost instantaneous command of reduction of engine power and extremely fast modulation of engine torque when the vehicle requires TCS intervention. This means that it can control wheel slip within much tighter bounds than throttle TCS, but also induces a reasonable amount of vibration through the vehicle and can cause engine damage if operated for too much time. Further, this form of engine control requires significant engine ECU control of ignition and injection, which must be precisely controlled to avoid damage to the engine and the catalytic converter. Likewise, during injection suppression the liquid fuel film on the port walls evaporates quickly, meaning that when injection resumes abnormal combustion may result.

Lastly, as in ABS, wheel slip can be controlled using the braking system for traction control. Instead of optimising slip for all wheels under deceleration, the TCS uses only the brakes on the driven wheels to control slip under acceleration. In this way the brakes can both limit the amount of torque transmitted to the road (by converting the torque supplied by the engine into heat) and proportion it between the wheels (by providing bias across the differential through the brake system). This results in the ability to accurately control the driven wheel slip to provide the best tractive force available, while also providing stability. Operation of the system, however, requires a hydraulic pressure source in addition to the ABS hydraulic modulators to enable the system to increase braking force, as shown in Figure 2.28. This method would appear to provide a highly effective way of controlling traction and vehicle stability, but the main problem that arises with this system is the amount of heat generated within the brakes. The very high loads imposed by the engine can accumulate heat within the brakes quickly, as can the actuation of the brakes at high speed. As such, brake activation must be limited to a time and speed range to avoid brake overheating and subsequent failure, which is a significant limitation. It also can have the undesirable effect of adversely affect steering and vibration levels.
As set out, the wide variety of TCS control strategies produce many performance tradeoffs. However, it is possible, although more expensive, to combine these systems to improve overall performance. This brings increasing complexity into the design of the TCS ECU control algorithms and associated hardware, but the benefits to performance can be large, as shown in Figure 2.32.

![Figure 2.32: Control deviation using different TCS combinations [42]](image)

### 2.3.3 Vehicle Dynamics Control (VDC)

While ABS and TCS seek to assist the driver during braking and accelerating manoeuvres, VDC aims to help the driver during critical steering scenarios. When a vehicle attempts to negotiate a turn the VDC system will control wheel speeds to ensure a predictable driving state is maintained while also attempting to provide maximum practical cornering acceleration. In many respects this is a very similar goal to ABS and TCS, and combined, they represent a comprehensive system that ensures vehicle controllability and high performance during severe manoeuvres. Although VDC is the newest and most complex development, each system should be viewed as providing separate but comparable functions for different manoeuvres.

For ABS, the wheel is the controlled element, with wheel acceleration controlled to keep the slip sufficiently small to preserve some amount of lateral force capability. For VDC, however, the vehicle is the controlled element, with vehicle motion controlled to keep any deviation from its nominal motion as small as possible, and to conform with the environmental conditions through the control of wheel slips to gain the required lateral and longitudinal forces and control yaw angles. It does this by using the braking system and engine control to regulate the individual wheel torques, but also utilises additional data input from a steering angle sensor, a yaw rate (d$\beta$/dt) sensor and a laterally placed accelerometer [14].
2. Vehicle Stability Background

The effect of yaw rate on stability is an important observation, as the sensitivity of yaw moment on vehicle stability, with respect to changes in the steering angle, decreases rapidly as the slip angle of the vehicle increases. In critical cornering manoeuvres it is therefore important to control the yaw rate of the vehicle to correlate to the drive’s desired path, which is determined from the steering angle. At large vehicle slip angles (where the $\mu_y$-slip curve of the tyre is maximum), variation in the steering angle has little effect on the yaw moment, with the result that manoeuvrability is lost in these conditions. VDC attempts to utilise wheel slip control to manage vehicle yaw rate within selected bounds.

The addition of a yaw rate control does not, however, guarantee stability in all conditions. If the control is used on a slippery road the yaw rate of the vehicle may correspond to the driver’s requested turning rate through the steering wheel, but the vehicle may just be spinning in oversteer, and not following the intended course. In this case, while the yaw rate is controlled correctly, the lateral (cornering) acceleration of the vehicle does not correlate to the driver’s intended path, and it spins off the road. The
addition of the lateral acceleration sensor to the yaw rate sensor eliminates this problem and forms the backbone of the VDC system, shown conceptually in Figure 2.34.

The first task of the VDC controller is to determine the driver’s desired (nominal) path. It does this from the data gathered from the onboard sensors, including driver inputs of steering wheel angle, throttle position and brake pressure, but must also account for unknown variables, such as coefficient of friction, which can affect driving attitude and behaviour. Further, steering wheel angle, vehicle velocity and yaw rate can be used to determine the nominal yaw rate during turns (Eqn 2.2), which is limited by the coefficient of friction of the road (Eqn 2.3). [8, 37]

\[ \beta_{No} = \frac{v_x \cdot \delta_w}{(a + c) \left(1 + \frac{v_x^2}{v_{CH}^2}\right)} \]

\[ \text{Eqn 2.2} \]

where: 
- \( a \) = longitudinal distance from front wheels to centre of gravity of vehicle
- \( c \) = longitudinal distance from rear wheels to centre of gravity of vehicle
- \( v_x \) = longitudinal vehicle speed
- \( v_{CH} \) = characteristic vehicle velocity
- \( \delta_w \) = angle of steered wheel

\[ \beta_{No} \leq \mu_L \cdot \frac{g}{v_x} \]

\[ \text{Eqn 2.3} \]

Once the nominal path is determined, the controller compares it with the actual path of the vehicle, as measured through the wheel speed, yaw rate and lateral acceleration sensors. Any deviations are then sent to the dynamics controller to either brake or accelerate the offending wheel(s) to alter lateral and longitudinal forces and the yaw moment, the dynamics of which are discussed in Figure 2.35.
If the braking slip of the left front tyre is increased by a small amount $\Delta \lambda$ from an initial value $\lambda_o$ and if tyre slip angle is $\alpha_o$, then the yaw moment on the car is in a first approximation changed by the following amount:

$$\Delta M_{yw} = -\frac{dF_s}{d\lambda} \Delta \lambda (a \cdot \cos \delta_w - b \cdot \sin \delta_w) + \frac{dF_B}{d\lambda} \Delta \lambda (a \cdot \sin \delta_w + b \cdot \cos \delta_w)$$

Here, changes in the tyre normal force as a result of a change in the tyre longitudinal of lateral force are neglected, as are the changes in the aligning torque on the tyre. Similarly, the lateral and longitudinal forces on the vehicle will be changed by the following amounts:

$$\Delta F_x = -\frac{dF_s}{d\lambda} \Delta \lambda \cdot \sin \delta_w - \frac{dF_B}{d\lambda} \Delta \lambda \cdot \cos \delta_w$$

$$\Delta F_y = -\frac{dF_s}{d\lambda} \Delta \lambda \cdot \cos \delta_w - \frac{dF_B}{d\lambda} \Delta \lambda \cdot \sin \delta_w$$

These relations which can be derived for each wheel of the vehicle are extremely non-linear, since the derivatives of the forces are highly dependent on the operating point ($\lambda_o, \alpha_o$) of the tyre.

The effect of variation in the tyre slip may be explained best by using the figure. This illustration shows the forces $F_R (\lambda=0)$, $F_R (\lambda)$, $F_B (\lambda)$ and $F_S (\lambda)$. $F_R$ is the resultant tyre force that is obtained by the vectorial sum of the longitudinal and lateral tyre forces. $F_R (\lambda=0)$ is the resultant tyre force acting on the free-rolling tyre and is equal to the lateral force on the tyre that results from the slip angle $\alpha_o$.

If the tyre slip is increased to the value $\lambda_o$, then the lateral force on the tyre is reduced to the value $F_S (\lambda_o)$. At the same time a brake force $F_B (\lambda_o)$ is generated. $F_R (\lambda_o)$ is now the resultant tyre force. At the limit of adhesion between the tyre and the road the absolute values of $F_R (\lambda=0)$ and $F_R (\lambda_o)$ are approximately equal. Clearly, increasing the tyre slip then means rotating the resultant tyre force and therefore changing the yaw moment, the lateral force and the longitudinal force on the vehicle.

**Figure 2.35: Control of yaw moment and tyre forces with slip [37]**

Braking or accelerating any given tyre can be used to control the vehicle slip angle, effectively helping to steer the automobile. By controlling individual wheel slip values, it is possible to significantly aid the driver during cornering manoeuvres. Unfortunately, this can come at a cost of unwanted deceleration or acceleration of the vehicle. It also can cause a lateral deviation from the nominal path, as the ability of the tyres to transmit lateral forces changes with the controlled longitudinal slip. The VDC system must control individual wheel slips to achieve a compromise between these effects, with the overall aims of:
2. Vehicle Stability Background

- Keeping the driver in control by providing vehicle response similar to normal driving conditions;
- Intervening on a ‘smart’ basis and only when needed; and
- Emulating the expert driver to assist the average driver in realising the performance potential of the vehicle.

The operation of VDC (which has the same operation as Bosch’s Electronic Stability Program – ESP) will be described through two examples, as illustrated by Bauer et al. [14].

In the first, two vehicles (one with ESP, the other without) initially travel on a high \( \mu \) road at high speed and enter a tight corner, as shown in Figure 2.36 and Figure 2.37. It can be seen that the vehicle without ESP soon becomes unstable (oversteer) and departs from its intended course. The vehicle with ESP remains on its intended course by selectively braking individual wheels to increase yaw moment, and thus helping the car to steer through the corner.

![Figure 2.36: Vehicle operation on a tight corner without VDC [14]](image1)

![Figure 2.37: Vehicle operation on a tight corner with VDC [14]](image2)
The second example, shown in Figure 2.38, depicts the potential benefits of ESP when a vehicle is accelerating at its physical limit around a corner of constant radius. On a high \( \mu \) (\( \mu_{\text{static}} = 1.0 \)) surface and at a corner radius of 100m, the vehicle without ESP reaches its stability limit at 95 km/hr. The result is significant understeer as the slip angle increases rapidly and the driver experiences difficulty keeping the vehicle on course. As the vehicle speed increases further to 97 km/hr, the rear end brakes away and all stability is lost as the vehicle leaves the course. The vehicle with ESP also reaches its stability limit at 95 km/hr, but at this point the ESP reduces engine torque so the limit cannot be exceeded. It also controls wheel torques to help steer the vehicle through the corner, reducing the sensitivity of driver steering inputs. This control results in small deviations from the nominal path that can easily be corrected by the driver.

Therefore, VDC can be a significantly aid when driving at the vehicle’s stability limit by controlling individual wheel torques. Further, because VDC reduces oversteer, it has also been observed that vehicle resistance to rollover increases [43]. These advantages show what can be accomplished in vehicle dynamics modelling and control above ABS and TCS by the addition of just three sensors (yaw rate, lateral acceleration and steering angle), and also represent ‘state of the art’ commercially available active control systems.

Future systems can be expected to improve safety and performance even further. As new and existing sensors and data sources are developed to a level where they can be economically installed in mass-produced vehicles, control models will become more accurate and incorporate greater possibilities. These advances include absolute vehicle velocity sensors, chassis attitude sensors (which measure vehicle roll, pitch and yaw),
road surface sensors, traffic monitoring via vision, laser and radar, and driver monitoring, using systems such at BAC and fatigue observers [23, 27]. It also includes Global Positioning System (GPS) information and wireless communication and broadcasting of various scales [44, 45, 46]. In this way, new controllers will be able to gain greater information from their surroundings, including communications with roadside devices and other vehicles, and far greater functionality will result.

On the other hand, control actuator advances also have great scope in improving vehicle safety and performance. Of particular note here is that the slip control method used by all ABS, TCS and VDC systems represents only one of numerous possible controllable parameters. Camber control, for instance, results in significant tyre grip increases beyond what is achievable using simple slip regulation. Other similar possible active actuators include variable spring rate and ride height [25, 43, 47], variable anti-roll bar stiffness, variable damping [47, 48, 49] and variable LSD and differential control [33, 38, 50, 51, 52, 53, 54]. Looking from a wider perspective other anticipated advances in active control, such as “brake by wire”, “throttle by wire” and “steer by wire”, offer greater flexibility from a controller point of view. Here, the mechanical linkages between the driver and the vehicle are replaced with electronic actuators, which give various vehicle controllers scope for fast, accurate and improved control of these parameters and capacity for greater integration between systems [27, 45, 46, 55, 56].

Of further interest is the scope for integrating separate automotive systems for reduced cost and greater functionality, as well as exploring emerging controller architectures and modeling techniques. In addition, performance increases can be realised by deriving additional control possibilities from existing actuators and developing new methods to gain greater information from existing sensors. In this way, greater vehicle performance and safety can be increased with very little additional cost. Road surface prediction is one such parameter that has broad scope in improving stability system performance and, as demonstrated below, has been explored using a wide variety of methods.

2.3.4 Road Surface Identification
As demonstrated previously, the coefficient of friction between the tyre and the road surface is an important variable in determining maximum tyre force and, therefore, stability controller decisions [57]. Determining road properties is also a very important aspect of realising “Intelligent Vehicle Systems (IVS)”, which aims to integrate the
entire vehicle for improved performance, as shown in Figure 2.39. It is a parameter that is difficult to measure directly, although many methods have been employed in special test facilities, such as using optical sensors, strain sensors and acoustic sensors. Such technology, however, is very expensive and is difficult to utilise in production vehicles [58]. As a result, many research efforts have been made in this regard, and are either based on the friction process itself, or the parameters affecting it [59]. In particular, dynamic tyre models have been utilised to determine road surface information, but as Bian et al suggests, this is of limited direct benefit.

Utilising pure theoretical modelling to determine adhesion characteristics is limited due to the complexities of the tyre, and therefore the complexities of the model. As such, many tyre models also utilise empirical data to overcome some of these problems. For example, the empirical “Magic Formula” based model presented by Bian et al [58] requires experimental data on the peak value of the coefficient of friction and optimal slip ratio for different road surfaces. In this way, the tyre model can compare the current tyre condition with stored tyre data to determine road surface type.

Current ABS, TCS and VDC systems estimate tyre/road adhesion in a similar manner, generally utilising approximate empirical dynamic models of the tyre [57]. Nonetheless, these models, like the one presented by A. Hac et al [22], can only determine road
surface coefficient of friction during unstable conditions. This is because in the tyre linear region the tyre effects that are measured for the model are predominantly the result of tyre elastic properties, not of road surface type. Therefore, the controller can only recognise situations when the vehicle is at, or near, the limit of adhesion when determining road surface coefficient of friction. The generalities in these models mean that they often produce erroneous results during quick transient conditions.

Clearly, there is great scope for developing tyre/road friction estimation methodologies, but none of the proposed techniques have shown potential for implementation in mass-produced vehicles (as of 2001) [60]. Nonetheless, there has been a significant amount of work into the area, of which V. Ivanov et al [19] presents a summary as discussed below.

Ivanov et al breaks the areas of study into virtual and sensor based procedures, which are sub-classified into:

<table>
<thead>
<tr>
<th>Sensor Based Procedures</th>
<th>Virtual Procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>• On-board direct grip measurement</td>
<td>• Dynamics simulation method</td>
</tr>
<tr>
<td>• On-board indirect grip measurement</td>
<td>• Statistical method</td>
</tr>
<tr>
<td>• Off-board (on road) measurements</td>
<td>• Fuzzy logic method</td>
</tr>
</tbody>
</table>

Some aspects of the Sensor Based Procedures have been touched on previously, and attempt to either directly measure surface features, or parameters that are functions of surface coefficient of friction. On-board measurement of tyre/road friction has been attempted with infrared and microwave radiation, the Doppler effect and optical analysis, which all are usually chassis mounted devices that scan the road surface. Using these techniques the parameters of the road microprofile can be measured, and used to estimate coefficient of friction. Furthermore, some systems integrate these parameters together to develop greater information. For example, a system is presented that uses an ultrasound device to determine the “optical road density” and a laser to determine surface geometry, and combines the results to take into account the micro and macroprofiles for friction coefficient determination. These systems have substantial disadvantages that lie in their sensitivities to external influences and their inherent high cost. [19]

Other methods of coefficient of friction determination rely on measuring or estimating the forces and moments that are applied directly to the wheel, or indirectly through the suspension, drive-train and braking elements. This includes using piezoelectric...
transducers, which are directly mounted into the wheel’s hub. Such systems, however, have a number of drawbacks. Firstly, the measurement methodologies that must be used are prohibitively expensive. Secondly, these systems also require periodic calibration, which is inappropriate for commercial automotive applications. Finally, measurement of drive/brake torque requires an estimation process derived from the brake pressure or engine data. As such, these approaches are only appropriate for vehicle testing and research tasks. [19]

Another possibility in sensor based grip measurement comes from Intelligent Transport Systems (ITS) applications. ITS offers the ability to communicate with individual vehicle systems directly, and can provide road surface information from road mounted sensors. Elements such as road profile and type, and road temperature moisture can be measured directly by stationary apparatus and incorporated into vehicle controllers. Further, this is only one of the many benefits of ITS which, if adopted, will provide significant automotive and traffic advances. However, such a system is very unlikely to be universally adopted on all roads in the short or medium term, as it involves significantly more development and massive capital investment by both vehicle manufactures and road authorities.

The Virtual Procedures, on the other hand, attempt to derive effective road surface identification from affordable vehicle mounted sensors. The dynamics simulation method covers the examples given previously in this section (both theoretical and empirical), which Ivanov et al define as methods to derive the tyre friction coefficient from electronic control unit embedded vehicle models. The goal is to determine specific forces at the contact patch or to determine the tyre/road coefficient of friction, using hardware measured parameters such as rotational velocity of wheels, yaw rate, steering angle and vehicle acceleration.

These systems have many drawbacks including; relying on road surface empirical database approximation; containing limited adaptability; producing road surface data only when near the critical slip; generating outputs of only a few different surface types and; implementing steady state modelling to predict parameters in highly dynamic (even chaotic) systems. [19]
In these approaches the weak point lies in the necessity of database with approximation coefficients for different road surfaces. In addition, the road identification system should be self-tuning during control process. Results of some investigations show that the adaptation process for empirical models is a complicated optimization problem and reveals a stable outcome only within pre-extreme area of $\mu_s$(slip) curves. Other imperfections of these systems lie in the limited set of identifiable road surfaces. The systems operate with generalized concept of dry, wet and ice covered road, as a rule.

Despite a variety of physical models, it is comparatively difficult to find a compromise between the model accuracy and model adaptability to the vehicle control system. One more important disadvantage of the majority of physical models consists in their alignment to steady-state conditions first. But tire/road contact should be dated to “chaotic” systems.

V. Ivanob et al [19]

Vibration analysis is also classified under the dynamics simulation method category and has a number of applications, the most obvious of which is using suspension vibration to determine road surface roughness [61]. This principle, however, extends to other parameters and, as T. Umeno et al [62] demonstrates, has particular benefits to surface recognition that are only recently becoming clear. In particular, this work states that by monitoring the frequency characterises of wheel speed vibration, information can be derived that is indicative of the coefficient of friction, as shown in Figure 2.40. Here, resonance and frequency band strength differences (as shown through a Power Spectral Density – PSD graph) are used to numerically estimate the slope of the longitudinal coefficient of friction/slip curve at the current slip ($\alpha$) for different surfaces, which can be used to determine surface type as depicted in Figure 2.41. This has the particular advantage over other dynamics simulation methods as road surface can be identified in stable conditions.

![Figure 2.40: Power spectrum density of wheel speed vibration for different surfaces [62]](image)

![Figure 2.41: PSD estimation of coeff. friction/slip curve slope, $\alpha$ [62]](image)

Power spectral density has particular advantage here, because it describes how the power of the time series is distributed with frequency, as a frequency-domain plot of power per Hz versus frequency. In other words, it shows at which frequencies time series data variations are either strong or weak, and is considered a useful tool for identifying
oscillatory signal attributes. Further, PSD is normally determined from the Fourier Transform of the time history signal. The signal is broken down into an integral transform that re-expresses the signal in terms of a number of sinusoidal basis functions. This essentially decomposes the time history signal into its component frequencies and amplitudes, and maps it from the time domain into the frequency domain. As a result, the transform represents the time history as a series of sinusoidal waves of different frequencies, often requiring many high frequency waves to reproduce complex or sharp time history curves. The PSD is then determined by calculating the square of the magnitude of the continuous Fourier Transform of the signal. [63, 64, 65, 66, 67]

In this case, Umeno et al [62] constructed PSD information by measuring individual vehicle wheel speeds using magnetic sensors. The sensors counted 48 serrations on each hub, which was then converted to angular velocity using a 32bit microprocessor. The resultant PSD information was then used to estimate $\alpha$, which was supposed as unique to each surface type. The research showed that wheel speed vibration information could be used to determine the difference between dry asphalt and ice-covered roads in stable conditions, as well as determine if hydroplaning had occurred. Of particular significance was that such a technology required only wheel speed sensors, which are already in widespread use.

Nonetheless, dynamics simulations methods generally inherit a number of limitations, as discussed above. To overcome some of these disadvantages, many advanced stability controllers also use statistical methods to determine tyre grip characteristics. In these cases, a large amount of tyre data is measured in extensive provisional studies of specific tyre types. This database is then used in control systems utilising statistical models, such as correlation and regression. A simple example of this can be shown with the aid of Figure 2.42, where a specific coefficient of friction as a function of longitudinal slip curve is given. Here, the statistical model analyses the measured vehicle and tyre parameters, and fits a curve for a specific surface based on what it has observed. This has the particular advantage that most tyres have a discernable correlation between $\mu_1$ and $\mu_2$, which allows for some maximum coefficient of friction forecasting when the tyre is in the slip region. Again, this means that the statistical model, like almost all dynamics simulation models, cannot determine road condition in the linear zone because this relationship is only valid in the transition zone. Furthermore, as can be seen in Figure 2.43, the measured parameters required for the statistical models have a very large
spread on real road pavements. As a result of this, model reliability suffers to a significant extent. [19]

Further, L. Jun et al [68] presents a similar method for road surface identification for ABS. The research here is described as using a number of pre-stored road surface curves to identify the road surface type by comparing measured and assumed wheel angular deceleration. This process is depicted in Figure 2.44.

Here the measured slip is used to determine estimates of the longitudinal coefficient of friction using three general friction/slip curves, which are then multiplied by an estimate of the normal force on each tyre to determine longitudinal force. This data is then used to predict at what rate each of the wheels should be decelerating at for each of the three surfaces. This information is then compared to the measured wheel angular decelerations, and the type of road with least absolute difference is assumed. While such a system has some use in determining road surface, it can be observed that this approach is limited in application. Of particular note is that the statistical data for each of the surfaces is significantly simplified. This is further compounded by the need to estimate
normal tyre force and braking torque to determine the rate of predicted tyre deceleration for each surface. Such a technique requires many estimates to be made, many of which could be highly erroneous.

More acceptable results can be obtained by combining statistical dependencies with dynamics simulation models. As Ivanov et al explains, fuzzy logic methods have application in this regard, and generally agree closely with statistical data of realistic road surfaces. Particularly, fuzzy control has a history in estimating indirectly specified parameters (such as determining and vehicle velocity for wheel slip calculation), but also has been found to be useful in determining road properties.

Fuzzy controllers have a number of benefits over other systems, in that they do not require a detailed mathematical model of the control system, or an understanding of its dynamic nature. Instead, they encode heuristic knowledge, and operate using a set of ‘if –then’ decision rules to control the system. Furthermore, these are often encoded in normal language, which makes the models easier to understand and to program, but results in a need to tune the controller iteratively. [69]

For example, G. Mauer et al [69] presents a fuzzy logic controller that can address the problem of road condition identification based on the comparison of brake pressure and detected slip ratio. If the controller identifies that the tyre slip ratio is larger than anticipated (based on current brake pressure and an assumption of road condition) a new road condition is assumed with lower coefficient of friction. This test then repeats until the assumed road surface correlates to the dynamics of the system. The road condition identifier, used in this investigation in an ABS application, is capable of identifying four separate road conditions; dry, wet, ice covered or blocked (high slip) wheel. Initially the controller assumes a dry road condition. If the slip ratio exceeds a preset limit not encountered during normal operation on this surface (U=14% in this case) wheel blockage is assumed. A series of tests are then run though the identifier to try and gauge the road condition. Expected slip and actual slip are compared for a given brake pressure and then used to identify the road first icy, then wet and then whether or not the road surface has returned to dry in the meantime. Only one of the four conditions can be true, and as such the fuzzy observer is capable of categorising the road surface into three discrete road surface types, or as blocked. This information can then be sent to the ABS controller for improved operation.
Fuzzy logic, nonetheless, is not the only means of combining dynamics simulation and statistical information for road surface identification. The fact that Ivanov et al [19] presents it as a separate Virtual Procedure category does, however, highlight it as a major research area with the presumed exclusion of all else. This is not the case, and research that attempts to derive surface information using Artificial Neural Networks (ANN) has been acknowledged for some time. In these instances ANN simulation models of dynamic processes are derived using statistical data, and then used to determine specific surface features.

Artificial neural network models have been in use in the manufacturing industry for some time, and their application to automotive systems has been a significant research area for approximately the last decade. By mimicking the architecture of the biological brain at a neural level, they enjoy significant advantages to conventional modelling techniques. They have particular advantage over conventional models because they have the capability to model the strong non-linear behaviour of systems and are very resistant to measurement noise [70]. Further, ANN model construction requires only the historical data of the modelled system, not the detailed understanding of the process dynamics that conventional models require. In this way, the ANN model is regarded as being able to “train” itself based on observation. This has significant benefit because the modelled system does not have to be fully understood by the programmer, enabling modelling of complex systems with minimal effort. Such a feature means that ANNs are able to learn and model process behaviour where a priori knowledge of the associated scientific principles is not available, or extremely difficult to obtain [71]. This has the advantage of significantly reduced development time and cost [72]. Furthermore, ANN models are considered robust under a wider variety of operating conditions and require far less computing power than conventional systems once trained. This is further elaborated in the following statement.

Because of the topology of the systems and the manner in which the information is stored and manipulated, the [artificial] neural networks are often capable of doing things that humans or animals do well, but that conventional models do poorly. Moreover, artificial neural networks have the ability (1) to recognise patterns even when the information involving these patterns is noisy or incomplete, (2) to do matching in high-dimensional spaces, and (3) to effectively interpolate and extrapolate from learned data.

Artificial neural networks are useful on several counts. Since they are adaptive, they can take data and learn from it. Thus they conjecture solutions from the data presented to them, often making quite subtle relationships. Artificial neural networks can reduce development time by learning underlying relationships even if those relationships are difficult to find and to describe. They can also solve problems that lack existing
solutions. Since artificial neural networks can generalise problems, they can precisely process data that only broadly resembles data they were trained on originally. Similarly, they can manage imperfect or incomplete data, providing a measure of fault tolerance. Being nonlinear, artificial neural networks can capture complex interactions among the input variables in a system.

H. Kim and P. Ro [73]

Of note is that because ANNs are developed through process learning, they do continue to make some “mistakes” [74]. Such a problem is of particular concern because ANN models are considered “black boxes”, which means that it is extremely difficult to determine the method with which the process has been modelled internally. This is in contrast to conventional models, whereby this is often transparent, and as such ANN models are considered to have the potential to operate unpredictably in some conditions. Nonetheless, for many applications this problem is considered small when compared to the potential benefits of ANN implementation, and ANN modelling is an area of significant growth. ANN modelling will be discussed in further detail in later sections.

Research presented by W. Pasterkamp et al [59] uses the ANN method to estimate the coefficient of friction, slip angle, longitudinal, lateral and normal forces and engine/brake torque for a single wheel based on inputs from steering angle and suspension potentiometers, four strain gauges within the wheel assembly and a load cell on the steering linkage. Using a conventional method of modelling this process would produce either a highly simplified or highly computational intensive model, because many factors must be taken into account. These parameters include tyre characteristics, camber, trail, toe, static angle of inclination, damper settings, anti roll bar stiffness, linkages and joint flex and vehicle body attitude. As such, conventional modelling represents a significant problem because a detailed model would be too slow for useful real-time identification of surface type and tyre forces, while a simplified model would contain a high degree of error. Instead, using an ANN model is identified as providing a possible solution to this problem. In this work, a comprehensive full vehicle multi-body model was created using conventional methodology, and force, moment and slip angle data then derived from it for a specific vehicle operating under specific conditions and coefficient of friction. This data was then used to construct an ANN model of the input/output relationships. Such a model could then be used to emulate the comprehensive conventional model with significantly reduced computational intensity, and thus greatly improved ability in real time parameter estimation, the results of which are shown in Figure 2.45. Pasterkamp et
al [59] then exported this simulation model to an actual vehicle and depicted the results in Figure 2.46. This then lead to the conclusion:

Simulation experiments and experiments with a test vehicle have shown the possibility to estimate side slip angle and friction coefficient directly from measured entities using [ANN]... For actual implementation, it has been shown that artificial neural networks can perform this estimation adequately.

Nonetheless, the above case remains a “model of a model”, and increased error propagation can be expected within the ANN reproduction, as is evident in Figure 2.46. It is noted, however, that the data derived from the conventional model for ANN model construction could have been derived experimentally instead. While this process would have required expensive wheel dynamometers mounted to the test vehicle, it would have been possible to avoid a significant level of this error propagation by utilising this experimental information, rather than the simulation data used here.

This concept of using experimentally based ANN models for the identification and classification of road surfaces is presented by T. Shiotsuka et al [75]. This research differs greatly from the work performed by Pasterkamp et al [59], and seeks to use measured suspension acceleration vibration within ANN models to identify different road surfaces in dry conditions. Here, a test vehicle is fitted with an acceleration sensor mounted to a lower suspension linkage, and is driven on a number of different surfaces. These surfaces include: 1) New asphalt road; 2) Concrete road; 3) Worn asphalt road; 4) Asphalt road with periodic concaves; 5) Brick road; 6) Stone road; and 7) Very rough artificially constructed road. Data was logged on each surface at a sampling interval of
0.003 seconds (approximately 330Hz) while driving the vehicle at 40km/hr, with the measurements shown in Figure 2.47.

Figure 2.47: Suspension acceleration on different roads [75]

In all, the acceleration measurement for each surface is limited to 1024 data points (approximately 3.1 seconds), and is then used to construct an ANN model that predicts future suspension acceleration. Specifically, the ANN model is designed to take as input the ten previous acceleration measurements as a summed time history, and to then predict what the acceleration is expected to be 0.003 seconds into the future. Different ANN models are, thus, constructed for each of the seven surfaces, and the error between the predicted and actual accelerations computed for each one. A new road surface can then be classified into one of the seven categories by testing it against each of the seven ANN models, with the model with the lowest average error considered correct. It is noted, however, that the number of samples (q) used to obtain the average error has a significant effect on model accuracy, with a range of 200 (0.6 seconds) to 800 (2.4 seconds) samples investigated. These results are shown in Figure 2.48, and it can be seen that when more samples are used to determine average error, more accuracy can be obtained. Nonetheless, this comes at the cost of model sensitivity, with higher sample sizes reducing the speed at which changes in road surface can be identified.

<table>
<thead>
<tr>
<th>Number of Samples (q)</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>1234</td>
<td>1345</td>
<td>2345</td>
<td>3456</td>
</tr>
<tr>
<td>Error</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
</tr>
<tr>
<td>%</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
</tr>
<tr>
<td>Type-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>1234</td>
<td>1345</td>
<td>2345</td>
<td>3456</td>
</tr>
<tr>
<td>Error</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
</tr>
<tr>
<td>%</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
</tr>
<tr>
<td>Type-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>1234</td>
<td>1345</td>
<td>2345</td>
<td>3456</td>
</tr>
<tr>
<td>Error</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
</tr>
<tr>
<td>%</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
</tr>
<tr>
<td>Type-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>1234</td>
<td>1345</td>
<td>2345</td>
<td>3456</td>
</tr>
<tr>
<td>Error</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
<td>89.15</td>
</tr>
<tr>
<td>%</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
<td>58.12</td>
</tr>
</tbody>
</table>

Figure 2.48: ANN surface classification accuracy [75]
Shiotsuka et al [75] also presents a variation to the above model, in which Power Spectrum Density (PSD) is used in place of the time history of the suspension acceleration vibration for ANN model development. This was selected using similar justification as Umeno et al [62] described for road surface identification using wheel speed PSD curves discussed previously. In this case, however, the PSD curves of suspension acceleration using the 1024 measurement points for each of the surfaces were calculated (as shown in Figure 2.49) and used as a direct input into a single ANN model. This required some PSD curve simplification, with between 20 and 40 ANN inputs used to depict the curve shape. This was done by using as input “the values at the points with logarithmically equal frequency intervals from 1Hz to 100Hz” [75]. Additional experimental data was also derived by repeating the acceleration measurement for each road type up to four times each, which were then used to create additional PSD curves. This method was utilised because ANN models need more volume of data for accurate results than the single PSD curves for each surface could provide.

Furthermore, the PSD ANN model utilised different outputs compared to the time history ANN model. In this case, the ANN model was constructed with seven outputs representing each of the surfaces, each with a range of 0 to 1. Each of the outputs could be labelled “Surface Type 1”, “Surface Type 2”, ..., “Surface Type 7” respectively,
whereby any output of “1” positively predicted the road surface that corresponded to the specific output. Likewise, an output of “0” and any of the outputs could be interpreted as a negative prediction, stating that the current surface did not correspond to the particular output. Accurate ANN model surface prediction should produce a value of 1 for a single corresponding output, and a value of 0 for the other six.

The ANN model outputs, however, are designed to output an analogue range between 0 and 1, not a binary output as might be considered. This introduces an ability for the ANN model to estimate the accuracy of the output based on the following description:

*Classification is regarded a successful when the output value of the corresponding [model output] in the largest in all [outputs] and at the same time larger than the critical value J. Classification is regarded as having failed when the output value of one of the [outputs] becomes the largest in all [outputs] and at the same time larger than J. Classification is considered impossible when the output values of all [outputs] are smaller than J.*

T. Shiosuka et al [75]

The ANN model outputs can be used to not only determine the road surface, but also provide a means of determining the predicated accuracy of the model. By excluding data that is deemed by the ANN model to have higher than acceptable predicted error (based on J) the overall model was observed to improve in accuracy. The best performance in the study (of 97.1% predictive accuracy) was achieved with 20 PSD points used as ANN model inputs and J=0.5.

This result accuracy can be compared to the best average accuracy of the time history ANN model of 94.6% (based on average error % of Figure 2.48, where q=800 and data no.=1). Therefore, both methods are “very useful for…recognition of road conditions” [75], although it is noted that the increased accuracy of the PSD model comes with the expense of the increased computational complexity required to obtain the PSD curves. It is also noted that the result of 20 ANN model inputs providing better accuracy over high numbers is most likely a consequence of insufficient measurement data for ANN construction, rather than an inherent need to limit points when describing the PSD curve.

From these observations, ANN networks offer potential for determining relationships at, and within, the tyre that would either require excessive computation based using conventional methods or cannot be modelled without great effort. This offers many potential benefits, which appear to have not been investigated to any significant degree. By combining ideas presented for conventional modelling with advances already made
using ANN models, there are many avenues to investigate, each with their own unexplored possibilities. As such, ANN modelling for the determination of road surface parameters is considered an area of study that is both relatively unknown, and has potential in offering significant technology advances. As a consequence of this, road surface identification using artificial neural networks will form a considerable part of this investigation.

Surface identification only forms one element of any stability control system. By determining surface features, a stability controller can operate with greater efficiency but, as explained below, there is also significant inroads that can be made in controller design.

2.3.5 Intelligent Stability Control

Active control of a vehicle must measure a large variety of parameters for efficient control, and make appropriate assumptions and estimations when parameter measurement and integration into a control system is either too impractical or too expensive. Due to the complexities of the vehicle system the difficulties involved in modelling such systems to a high level with traditional controllers are many. In fact, current stability controllers generally only measure a small number of parameters in the control logic, and then makes decisions based on this data alone. The limited number of sensory inputs has been shown to provide enough data for proficient automotive control in most conditions, where the general assumptions made hold with reasonable accuracy, but can fall down in abnormal situations, resulting in flawed control. Clearly, measuring more parameters and/or integrating the data into improved controllers have potential to enhance stability controller performance significantly.

The way in which the forces from each tyre interact with the vehicle as a whole must also be considered in detail to ensure predictable drivability. For example, if a single wheel is braked to regain traction, the braking action will produce a turning moment on the vehicle that may be to the detriment of stability. Similarly, transient responses through the suspension system to scenarios such as driving over bumps in the road and erratic driver control can complicate matters. These factors, and others such as vehicle inertia and wind loading, produce complexities in stability controller design that can be addressed to a greater level in future systems for improve performance.
At present, determination of efficient stability control (ABS, TCS and VDC) relies heavily on process simplification and, as such, many assumptions must be made. When conditions arise that negate these assumptions the control logic stands a high chance of not working optimally, and can even act so erroneously it intensifies the problem. It can be seen there are numerous situations when its operation does not represent optimum control. In fact, many systems include a deactivation function when driving in abnormal conditions because its operation can be so erroneous [37]. Consider a vehicle that is loaded to create a high centre of gravity, which is then forced to brake heavily into a corner. In such a case, traditional ABS would seek to provide much higher cornering forces at the braked wheels compared to what is achievable if the driver locks the wheels. However, since the vehicle has a high centre of gravity, the ABS operation actually causes a rollover incident because of the increase lateral acceleration. Such an incident appears statistically common, and often has a higher risk of injury to the occupants than if the ABS was not activated. By ignoring important elements of vehicle dynamics the ABS controller acted erroneously. As such, significant improvements in stability controller performance can be expected by increasing the sensor data to the stability controller, improving stability controller design, technology and integration and providing more controller output actuators. As such, the goals of future control systems will be to:

- Provide optimum performance in all conditions;
- Evaluate driver requests and alter vehicle parameters to suit;
- Determine operation goals such as performance, safety, fuel economy or comfort;
- Accommodate for physical vehicle alterations, such as weight distribution changes and tyre wear; and
- Alter as many control parameters as possible to provide maximum tyre adhesion levels, including active suspension, real time damping, active camber change, automatic load distribution and rear wheel steering [11, 25, 76].

Clearly this means the inclusion of more comprehensive mathematical models, and associated sensory data, using traditional techniques. The complexity of the vehicle system means that the inclusion of increased data would also require extensive investigation and algorithm development to model the effects on the vehicle dynamics. This is in addition to the difficulties that are generally encountered when mathematically modelling vehicle systems that are inherently non-linear during normal driving, and
extremely non-linear when the vehicle is pushed to its performance limit [24]. Of course, overcoming these problems using traditional methods also means a significant increase in the size of the necessary control algorithm computations, which exponentially grow in complexity with the inclusion of additional parameters, as illustrated by Bannatyne [28].

An alternative to the extensive algorithm development required to meet these future goals exists in using intelligent systems. These systems, which are numerous and widely varied, offer a huge array of potential benefits. Some attempt to solve existing problems with new methodologies, some are adaptive, some can be programmed using easily understandable heuristic knowledge, some can learn the dynamics of the processes on their own and others can reduce computational complexity. In addition, the ability to replace the complex mathematical models used in current systems with models based on observation can significantly reduce model complexity, allowing for the addition of extra sensory inputs and control outputs. It is this ability to incorporate additional data into the control algorithm with minimal programming and computing resources that makes the use of non-conventional techniques desirable. Combined with the intense global competition for reducing the time it takes to bring vehicles to the market [78], the possibilities that these systems offer the automotive industry are great, and include:

- Reduced development time and cost,
- Reduced controller computation times,
- Reduced controller hardware costs,
- Use of increased sensor data in controllers,
- Gaining additional information from existing sensors,
- Ability to control more parameters,
- Reduced controller assumptions for robust control, and
- More accurate modelling of non-linear systems.

There have been a wide range of studies that have attempted to provide these benefits to the automotive industry, and some will be presented here.

The first study discussed was conducted by W. Krantz et al [79], and investigates three methods for the estimation of vehicle slip angle (side slip) and yaw rate from measured of tyre forces. Here, the research forecasts “progress in the development of sensor
systems for online determination of tyre forces” and attempts to use this new information to estimate vehicle slip angle and yaw rate in simulated conditions. Although no reference is made to the sensors that may be utilised, the first research method principally supposes that once individual tyre forces are known it is a simple process to evaluate tyre slip and slip angle for a given surface. Conversion into vehicle slip angle and yaw rate is then considered a geometry issue. To do this, an “inverse tyre model” is presented, which requires a significant understanding of both the tyre properties and that of the road.

The second method uses “direct integration” of the vehicle accelerations in the road plane to estimate vehicle slip angle and yaw rate. Here, the measured tyre forces, the vehicle mass and moment of inertia, vehicle geometry, aerodynamic drag, steering angle and tyre camber are used in a simple two track models to determine vehicle linear and angular accelerations and slip angle. When tyre slip is determined to be large, the estimated (or measured) vehicle accelerations are integrated. This information is then combined with the measured tyre forces and yaw rate and vehicle slip angle determined.

The third method presented by Krantz et al [79] consists of a “closed-loop observer model”. Here it is supposed that an open loop model of the vehicle system would deteriorate in accuracy to a significant extent because of the effect of non-linearities within it. Instead, a closed loop system is used to feed back measured signals to the model that are not used as system inputs. In this way, deviation of parameter estimates from true system behaviour can be compensated for, and vehicle slip angle and yaw rate estimated. In practice it was found that this system followed the dynamic behaviour of the vehicle more closely than the open loop model, but large errors were introduced from a number of sources. These sources included model error, wrongly selected tyre parameters, changes in coefficient of friction, incorrect alignment between estimated and measured tyre forces and variations in feedback gain.

Finally, Krantz et al [79] observes that these models depend on a high degree of knowledge of the tyre parameters and the coefficient of friction. Through this observation it is then argued that these systems cannot provide any substantial benefit for vehicle state estimation, when compared to conventional system layouts. Furthermore, this research appears to tackle the problem back to front, in using very hard to measure
parameters (such as tyre forces, tyre properties and coefficient of friction) to estimate yaw rate (which is relatively easy to measure) and vehicle slip angle.

Following a similar line of investigation, A. Hac et al [22] introduces an algorithm to estimate vehicle slip angle and yaw rate, this time using steering wheel angle, wheel speed and lateral acceleration sensors. The algorithm comprises of three separate models, which integrate through an observer model to determine yaw rate and vehicle slip angle. The first model attempts to estimate these yaws rate by modelling the kinematic relationship between the vehicle wheel speeds and the steering angle, in both linear and non-linear conditions (although it is noted that accuracy will be greatly diminished during severe manoeuvres). The second model attempts to do the same, except this time using a speed dependant dynamic model of the vehicle in the yaw plane using vehicle acceleration and steering angle as an input. In this case, a closed loop observer is used to gain greater accuracy by feeding back the mismatch between actual vehicle parameters and those of the model. This method is similar to the one used by Krantz et al [79] above. Finally, it is observed that coefficient of friction has a strong effect on the accuracy of these models, so a coefficient of friction observer is also developed. Here, the observer determines surface friction when the vehicle is at the limit of adhesion (coefficient of friction = current lateral acceleration / maximum lateral acceleration on dry surface), and assumes this is constant when travelling in stable conditions and during quick transients.

In addition, A. Hac et al [22] observes that these models rely on many assumptions and, as such, calculations are made to determine confidence levels in the results. In particular, confidence levels are reduced when conditions such as large slip of undriven wheels (i.e. during heavy braking), ABS activation, low speeds on slippery surfaces, quick transient manoeuvres and when the model estimate exceeds a predetermined level. In this way, increases in accuracy can be obtained by observing the limitations of the models, although it is noted that this increase comes at the cost of restricted operating conditions. The results of the three models, with appropriate confidence level adjustment, is used as input to an observer model for vehicle slip angle and yaw rate determination, as shown in Figure 2.50.
2. Vehicle Stability Background

Figure 2.50: Algorithm flow chart for vehicle slip angle and yaw rate estimation [22]

The non-linear observer represents a simplified model of the dynamics of the vehicle in the yaw plane, using as input the estimated coefficient of friction, vehicle speed, steering angle, lateral acceleration and the preliminary estimate of yaw rate. The latter three inputs are also used as feedback signals within the observer to avoid model divergence from the actual vehicle due to external disturbances and model error, with the gains determined by the confidence levels. The observer then calculates the vehicle yaw rate and slip angle as output.

The research successfully combines purely kinematic vehicle modelling with an estimation method based on a dynamic model of vehicle motion in the yaw plane to produce a relatively robust estimate of vehicle slip angle and yaw rate. However, it is limited in functionality, as stated below, with the estimation method losing accuracy when it is most needed.

It was found that during normal driving, without rapid steering changes and severe braking, when the vehicle remains on or close to the linear range of operation, the preliminary estimates were generally good... The estimation process becomes more difficult during limit handling maneuvers on slippery roads, especially when heavy braking is applied or when the vehicle is allowed to develop large slip angles.

A. Hac et al [22]

Parameter estimation is only one aspect of intelligent stability control, however. The work presented by W. Selby et al [80] discusses the advantages of “Integrated Chassis Controllers”, in which ABS, TCS, VDC, active suspension and active front and rear wheel steering can be integrated for improved performance. This global design methodology is termed in the research as “Intelligent Vehicle Motion Control (IVMC)”, and is described as interfacing theoretical generic controllers with existing chassis
subsystems. The concept is shown in practice by integrating individual wheel braking with active front and rear steering control within a computer simulated model.

Firstly, the IVMC conducts closed loop control of the longitudinal, lateral and yaw motion for normal driving by controlling the individual wheel torques and front and rear steering using traditional modelling techniques. Here, the control values are found by solving a non-linear model of the vehicle using input/output linearisation and a sliding model algorithm, which is in turn developed through conventional tyre models. The goal of this controller is to ensure the vehicle dynamics are optimised, and that the vehicle follows the driver’s intentions. To do this, it interprets the driver demands as desired forces in the three linear and rotational vehicle motions, which are controlled by the driver through the steering wheel, brake pedal and throttle.

The IVMC then integrates a second mode of control, namely yaw control for emergency manoeuvres. This control differs from the first in that the yaw control operates independently of driver demands, and attempts to ensure yaw motion stays within defined bounds to maintain drivability. By combining both controllers in the integrated IVMC it is then argued that such a system will provide vehicle manufacturers with speedy and reliable control solutions, and improve vehicle dynamic performance. However, it is noted in the research that this system requires all parameters to be accurately measured or estimated. This includes high quality estimation of the coefficient of friction, which is at the core of the tyre model accuracy.

From the examples above, conventional techniques can be used to intelligently determine parameter estimates for use in future stability controllers, and apply new control methods. However, it has also been noted that these conventional methods have significant drawbacks as a consequence of the non-linearities of the vehicle and tyres, and difficulties in determining the coefficient of friction. As such, other control strategies have also been developed, including fuzzy logic.

In the work performed by F. Assadian [60], for instance, a traditional controller is coupled with fuzzy logic “correction” for different road surfaces under simulated ABS operation. In this way, an H∞ controller is used to control a “brake by wire” model using state-space equations, which essentially involves discretely linearising the non-linear system. The system consists of a single wheel model that is used in straight-line motion
only. As such, the simulation is only concerned with the brake/slip relationship at an individual wheel, which simplifies the process considerably.

In a simulation the traditional dynamic controller computes the slip of an individual wheel and, if the slip is greater than the optimal value, the controller reduces brake torque by a determined amount. However, the controller is predominately concerned with modulating the brake pressure correctly to avoid excessive slip oscillation, and does not compute the optimal aim slip amount (which it assumes is a constant value). To solve this problem fuzzy logic laws are implemented to relate deceleration to optimal slip to achieve reduced stopping distance. Using this method, deceleration is observed by the fuzzy logic controller at the current slip, and fuzzy rules are used to indirectly determine which surface acceleration/slip curve fits best. Optimum slip for different surfaces can be estimated using fuzzy rules and input back into the brake controller for improved operation. The results can be seen in Figure 2.51, in which vehicle speed is plotted against distance travelled after brake activation.

![Figure 2.51: ABS controller performance on ice at 50km/hr](image)

As with the ABS fuzzy controller, TCS can benefit from fuzzy logic in a similar way. By replacing complex mathematical models with heuristic decision rules, control strategies can be accomplished with limited understanding of the systems and less effort. This can be shown through the work performed by Cheok et al. [21]. In this situation fuzzy TCS was installed into a 4WD vehicle to directly control the throttle, brake and transmission, where transmission intervention was included into the system to enable gear up-shifts to reduce engine torque coupled with throttle relaxation. The study also included the incorporation of engine and transmission speed, steering angle, three dimensional acceleration, yaw rate and wheel speed sensors.
Fuzzy logic controllers were developed first to control the brake pressure at each wheel, with transmission speed, throttle position and calculated slip at each wheel used as model inputs. Each of these variables were then designated quantities of either large (L) or small (S) and, since the hydraulic modulators fitted to the vehicle allowed only brake pressure modulation levels of 0, 50 or 100%, the control logic for the model outputs utilised S, M & L labels respectively to represent output magnitudes. Each of these four fuzzy control tables in Figure 2.52 show that small transmission speed, small throttle position and large slip produce a medium (50%) brake pressure control output.

The second fuzzy control stage is used to determine appropriate transmission up-shift and throttle relaxation based on the previous brake pressure control. This done to reduce the load on the brake discs, and hence increase the period of time they can operate.

In this case, the control logic used S for small values of torque reduction and throttle position reduction (relaxation), and L for large values, with the actuation signals of each determined by experimentation. The overall fuzzy control logic is also shown in Figure 2.52. Furthermore, the effects of the fuzzy TCS performance is compared to the case of no traction control in Figure 2.53 and Figure 2.54. Here the fuzzy TCS has had a positive effect on increasing yaw stability and that the steering and steering feedback is more stable. It is noted too, that the research states that the fuzzy controller performed well in a variety of conditions and on different surfaces.
K. Buckholtz [81, 82] presents another form of fuzzy control for yaw rate limitation, one of the main aspects of VDC. In contrast to the work by F. Assadian [60] and Cheok et al. [21] above however, this investigation seeks to use a fuzzy rule set to determine appropriate individual wheel braking couples to generate corrective yaw moments. As such, the simulated problem is not concerned with surface variation or brake control elements, but rather in determining the appropriate slip at each wheel to efficiently control yaw rate. The fuzzy controller acts as a higher-level supervisory module, with a number of traditional slave controllers then used to modulate individual brake pressures to achieve the goal wheel slips.

The first step of the model consists of defining the error between the actual yaw rate of the vehicle, and the desired vehicle yaw rate (which is presumably calculated from driver inputs and vehicle condition). This equation is shown in the first publication [81], and is repeated below in Eqn 2.4. Buckholtz suggests an improvement in a later article [82], in which the error term includes scaled yaw acceleration error too. This equation is shown in Eqn 2.5, which Buckholtz describes as providing “sideslip angle limitation” and helps to avoid unnecessary operation of the yaw controller. This error term is compared to a threshold value \( \Omega_{th} \), and if it is greater than this it determines that control is required.

\[
e_\Omega = |\Omega| - |\Omega_d| \quad [81] \tag{Eqn 2.4}
\]

\[
e_\Omega = (\Omega_d - \Omega) + (\Omega_d' - \Omega') \quad [82] \tag{Eqn 2.5}
\]

\[
ds = \Delta e_\Omega + ye_\Omega \quad [81] \tag{Eqn 2.6}
\]
where: $e_\Omega = \text{yaw rate error}$  
$\Omega = \text{vehicle yaw rate}$  
$\Omega_d = \text{desired yaw rate}$  
$\Omega' = \text{vehicle yaw acceleration}$  
$\Omega_d' = \text{desired yaw acceleration}$  
$k_\Omega = \text{a design constant}$  
$d_s = \text{required yaw correction}$  
$\gamma = \text{a design constant}$

The amount of change needed to correct the yaw rate error is defined through Eqn 2.6, which Buckholtz identifies through past work. The fuzzy controller uses this single input to determine the goal slips for each wheel to correct the yaw error, with the process shown in Figure 2.55. These decisions are based on the observations in Table 2.1, in which changes to individual tyre forces are listed as producing a pro and contra cornering yaw moments.

![Figure 2.55: Fuzzy supervisory yaw controller [81]](image)

![Table 2.1: Tyre forces for yaw moment correction [81]](table)

Here it can be seen that if the longitudinal braking force ($F_x$) is increased in the on the inside rear tyre then a pro cornering moment will be created that would counter an understeering condition. Likewise, increasing braking force in the $F_x$ direction for the outside front tyre will produce a contra cornering moment suitable for correcting oversteer. However, when considering the interplay of forces in tyre dynamics, purely increasing longitudinal (braking) force will alter the lateral force delivered by a tyre. If the tyre is operating in the stable region (which this research assumes) then an increase in braking force $F_x$ will produce a decrease in lateral force $F_y$. As such, the benefits of braking the inside front and outside rear wheels for yaw correction are unclear, and depends on the dynamics of the tyre. The fuzzy controller, thus, brakes the outside front wheel (by increasing aim slip to the brake controllers) for all oversteering conditions ($\gamma > 0$) and, when large yaw moment is required, is assisted by braking the inside front. Likewise, when understeering is detected ($\gamma < 0$) the fuzzy controller primarily brakes the inside rear tyre, which is assisted by braking the outside rear tyre. This process is shown in Figure 2.56 and Table 2.2, which depict the controller fuzzification and
defuzzification membership functions for the required yaw correction ($d_s$) and individual wheel aim slips ($\lambda$), and the associated fuzzy logic controller rules.

**Figure 2.56: Membership function fuzzification and defuzzification [81]**

The controller performance relies on the rules placed into the logic rule table, which in this case is determined through simulation experimentation and the expert knowledge of the designer. While this controller operation is based on simplified control rules and does not take into account parameters such as surface type and higher order vehicle dynamics, it does prove that fuzzy methods can be used to simplify controller design for otherwise difficult control problems.

Aspects of stability control have been attempted using artificial neural networks, which can offer a number of significant potential benefits, as discussed above. H. Sasaki et al [83] presents an ANN method for estimating vehicle slip angles based on measured yaw rate, lateral acceleration, steering angle and vehicle speed. Here, Sasaki et al argues that vehicle slip angle is crucial information for efficient yaw control, but that it is difficult to utilise within a controller due to the very high cost of the sensors required to measure it directly. Furthermore, it was observed that the widespread traditional modelling methods are very limited in estimating this variable because they rely on time integration of the yaw rate, lateral acceleration and vehicle velocity, and as such are highly susceptible to accumulation of noise and measurement errors.

The research then supposes that ANN models can be used to overcome these errors. As such, a vehicle was fitted with yaw rate, lateral acceleration, steering angle, vehicle speed and vehicle slip angle sensors, and was driven at a range of speeds and steering manoeuvres on a concrete pavement with all data logged. An ANN model was trained with yaw rate, lateral acceleration, steering angle and vehicle speed time history inputs (e.g. $A_y(k)$, $A_y(k-1)$, $A_y(k-2)$, $A_y = $ lateral acceleration, $k = $ current sample time interval)
and vehicle slip angle output. Running the ANN model on the test vehicle and comparing its prediction of vehicle slip angle to the measured value was used to test the model accuracy. The ANN model was shown to have only 5% error in determining maximum vehicle slip angles in the range of the ANN training data, with the results shown in Figure 2.57.

![Figure 2.57: Vehicle slip angle ANN estimation](image)

This represents significant potential in ANN modelling for vehicle slip angle estimation, especially because the ANN model only uses sensors that already exist in current VDC systems. Nonetheless, it is noted that this investigation is limited to only one type of surface, is susceptible to error if the vehicle is altered in any way and requires the expensive vehicle slip sensor for ANN training.

ANN models can be used for control of rear wheel steering, as described by T. Shiotsuka et al [84]. In this case the authors observe that for most four wheel steer applications, the modelling conditions, such as car mass, friction coefficient and tyre characteristics, are assumed constant irrespective of driving circumstances and vehicle motion. To this end, a need for adaptive active systems is highlighted and, in particular, control models that take into account the non-linearity of tyre friction are noted as having not been researched at all.

Two ANN models are presented to attempt to solve some of these problems. The first determines the dynamics of the vehicle using a simple traditional vehicle dynamics model, but utilises ANN models to determine the non-linear relationship between slip
angle and cornering force for each rear wheel and adjust steering controller gains. Two
ANN models are first trained based on data derived from a traditional non-linear tyre
model, with coefficient of friction, vehicle speed and slip angle used as inputs and
cornering stiffness (which is proportional to tyre lateral force) as output. The cornering
stiffness predictions for each rear wheel and vehicle speed are used as inputs to another
ANN model, which determines appropriate gains for direct control of the rear wheel
steering mechanism. This ANN model is also trained using simulation data at a range of
vehicle speeds and cornering stiffnesses. The rear wheel steering controller utilises ANN
models to adaptively adjust controller gain for improved performance in the non-linear
region. Furthermore, when tested within a simulation, the adaptive gain ANN controller
was shown to have superior performance to traditional fixed gain rear wheel steering
controllers. It is noted, however, that the ANN model requires coefficient of friction and
tyre slip angles and input, which are both difficult to measure and thus use within a real
world controller.

Nonetheless, T. Shiotsuka et al [84] presents a second ANN controller that seeks to
overcome these problems. This controller consists of two ANN models only, and is
trained from experimental data and tested on an actual vehicle. To construct the ANN
models, time histories of vehicle slip angle (β), yaw rate (r), vehicle speed (V), front
steering angle (δf) and rear steering angle (δr) are measured from the test vehicle for
different driving motions and speeds at a sample interval of 0.02 seconds (50Hz). The
“System Neural Network” is trained with this time history data to predict vehicle slip
angle and yaw rate 0.02 seconds into the future from all measured parameters. This
ANN model is constructed to model the dynamics of the entire vehicle, including the
non-linearity of tyre friction and overall dynamics. The accuracy of the ANN model is
tested against newly measured data and the results of the traditional vehicle model used
in the first ANN investigation. Here, the ANN model agrees well with the experimental
results for all tested conditions, whereas the traditional model is only valid for high
speeds. Such a result is regarded in the research to mean “car modelling with the system
ANN can represent the non-linear characteristics of both cornering force of tyre and car
structure very well in all cases.”

Simply modelling the vehicle dynamics does not provide enough scope to control the
rear steering angle. To this end, a second ANN model was developed, and referred to as
the “Neural Network Controller”. This ANN model uses yaw rate, vehicle speed and
2. Vehicle Stability Background

front steering angle as inputs and predicts the appropriate value for the rear steering angle 0.02 seconds into the future (i.e. for the next control step). Training the ANN controller using this method is a difficult task, and will be discussed in the ANN control theory in later sections. Nonetheless, the process can be seen in Figure 2.58, in which the controller error information (as determined from the system ANN) is backpropagated through the control ANN for training.

![Figure 2.58: Error backpropagation for ANN controller training](image)

The ANN controller training method works in the following way:

1. Initial system state data is input into the ANN controller and the required rear steer angle is predicted for system control;
2. The same system state data is input in the system ANN, except with the inclusion of the ANN controller prediction of rear steer angle;
3. System ANN predicts the resultant change in yaw rate and slip angle for the rear steer angle input, with the results fed back to the ANN models as new system states;
4. The error from the desired system state (vehicle slip angle $\beta=0$) and the modelled system state is determined at “Evaluation”;
5. This error is converted to an estimation of error at the rear steer control prediction; and
6. Rear steer error is used to train the ANN controller, and the process is repeated iteratively until this error remains small.

The ANN controller is capable of learning what is required to control the system ANN to minimise vehicle slip angle, which is the goal of the rear steer system, and is thus of an “Inverse Controller” type. Once the ANN controller was properly trained, it could be
installed into the test vehicle to directly control the aim rear steer angle. Driving experiments were conducted to determine the performance of the controller in a range of conditions, and it was found that this model closely followed the $\beta=0$ goal for all conditions. In fact, its performance was found to significantly excel all other systems that were examined. However, such a system has a number of drawbacks for simple operation. Firstly, by training one ANN model with another the scope for error propagation increases. Secondly, the training methodology is quite tedious because the process of converting vehicle slip angle error to rear steer angle error is difficult. Finally, the model is not robust to changes within the vehicle, wherefore changes to the vehicle require total model re-training.

Similar ANN modelling and control techniques can be utilised in other areas of stability control. A literature review conducted by M. El-Gindy et al [24] suggests that the suspension non-linearity that results from built-in bump stops, impacts and dry friction can be emulated within an experimentally trained ANN model, termed a “Process Network”. It is supposed that this model could be used to construct a control ANN in a very similar manner to that used by T. Shiotsuka et al [84]. This control ANN is referred to as an “Inverse Controller Network” and is utilised as shown in Figure 2.59.

![Figure 2.59: Adaptive ANN control of suspension dampening](image)

Here, the input to the inverse controller ANN is the desired output of the suspension system and a number of state variables. Once properly trained, the inverse controller ANN can predict the ideal control input to the suspension, which is passed to both the vehicle suspension and the process ANN. This information flow is in contrast to the previous example, where the ANN controller output was passed to the process (system) ANN for training only, and to the vehicle only for actual control. This is done because...
running the process ANN and the actual vehicle control in parallel produces extra functionality within the controller to adapt to external disturbances. Since the process ANN model should accurately emulate the actual vehicle, any error between the two is either the result of inaccurate ANN training or an external disturbance on the system. As such, this error can be fed back to the inverse controller ANN to adapt for these changes, and even update the inverse and process ANN learning.

![Figure 2.60: Measured and ANN predicted suspension loads [85]](image)

Although M. El-Gindy et al [24] does not provide any evidence of such a control system in practice, some of its abilities can be gleaned from the work performed by M. Burnett et al [85]. In this case, it is observed that the traditional linear suspension modelling techniques cannot adequately model many ride quality attributes, which are often dependant on the non-linear behaviour and high frequency characteristics of elastomeric and fluid filled components. As such, ANN modelling is attempted to predict the force produced by the suspension spring/damper using three inputs of suspension displacement, suspension velocity and suspension acceleration. The ANN training data is derived from a hydraulic damper rig than measures force and displacement, and associated velocity and acceleration is calculated to supply time history information. The ANN training data falls within the –0.6 to 0.6m/s range of suspension speeds, and the modelling results are then shown in Figure 2.60. Here, the model appears to predict suspension force very well in the –0.6 to 0.6m/s range, but accuracy soon diminished at high speeds. This is a consequence of ANN modelling, whereby good data fitting can be expected during operation within the bounds of model training but a degree of error is expected outside this range. Even so, the ANN model is still robust enough to forecast (or ‘guess’) results when conditions arise that are outside of its experience, which is a desirable feature.
ANN models have also been shown to have useful potential in tyre modelling, as H. Kim et al [73] demonstrates. As discussed previously, many aspects of tyre dynamics are non-linear, and the study particularly identifies camber as highly non-linear with regard to lateral force generation in the low camber region. It also explains that many tyre properties are not thoroughly recognised, and that improved tyre modelling is integral to improved stability control. As such, the study attempts to improve the capabilities of non-linear tyre modelling by introducing ANN models to the application. In particular, the study presents an ANN model that predicts the lateral force produced at a single tyre by monitoring vertical load, slip angle and camber angle as model inputs. Model training was accomplished using measured data, although it was noted that significant training computation was required to process the quantity of data. As such, the single ANN model was replaced with six separate ANN models, with each model used to predict lateral tyre force for different ranges of normal load. As a result, each ANN model was only required to model discrete regions of the tyre dynamics, and the level of training necessary for each ANN could be significantly reduced. This reduced training time significantly, and increased model accuracy to around 4% maximum error. Some of these results can be seen in Figure 2.61, in which the estimates of a conventional model are also shown. The figure demonstrates that the ANN model is far superior to the conventional model, but it is observed again that this model is specific to a particular surface and relies on model inputs that cannot be directly measured without expensive instrumentation.

![Figure 2.61: Measured, ANN predicted and conventional model predicted tyre lateral force [73]](image)

M. Gindy et al [24] provides further investigation into the potential uses of ANN modelling. In addition to the ANN suspension controller that was discussed earlier, this
work includes comments on the practicalities of ANN modelling for the entire vehicle, and provides the controller example shown in Figure 2.62. This is presented as a semi-trailer vehicle dynamics controller, but can be considered generic with some modification of the adaptive controller functionality. To this end, the vehicle produces two types of useful feedback outputs. The first are sensible to the driver, and their feedback allows the driver to control the vehicle accurately, at the passive level. The other feedback signals may not be able to be fully perceived by the driver and, as such, additional feedback and control can be implemented at an electronic level (controlled level) for improved performance. Therefore, the functionality of the entire vehicle ANN controller is to provide control assistance to the driver based on increased ability to acquire data and actuate different parameters.

![Figure 2.62: Suggested control model of driver/vehicle system [24]](image)

Although no research is conducted in this reference, it is argued that future research into vehicle encompassing ANN controllers has the potential of overcoming many of the limitations of traditional techniques. This is commented below:

The advantage [of ANN] is that the modelling of a vehicle using complicated physical laws can be avoided, yet at the same time accuracy of the emulation is maintained. When modelling a vehicle, researchers always try to simplify the complex sub-systems, but this procedure usually results in several assumption being made – and hence in inaccurate predictions. This situation can be avoided if the ANN can be used, as the input and output data required for training the ANN should be similar to what the vehicle will be facing in a real vehicle operation…

How to measure the appropriate signals? How to design the controller? And how to adapt the controller to handle the various operating conditions that may arise? Research into this area is required to answer these questions.

M. Gindy et al [24]

To date, several studies have shown that non-linear automotive systems can be successfully controlled using ANN control structures [86]. In addition, many studies
have shown that major improvements in automotive control accuracy can be realised using ANN models in closed-loop control application [70]. This is in contrast to the traditional and fuzzy logic controllers presented above, which show limited benefits over current systems.

The intelligent integration of new traditional modelling techniques into existing controllers, for instance, has a clear application in improving current stability controllers, but require an extensive amount of investigation. This is due to the exponential increase in mathematical modelling complexity with small increases in functionality, as the 17 year gap between ABS and VDC highlights. Such difficulties clearly limit the growth in functionality of stability controllers into the future, and any system that can achieve similar results within a smaller timeframe will have significant benefit. This gap is filled in part by the use of fuzzy logic controllers, as demonstrated above. Here, the requirements to mathematically model a vehicle system are replaced with a logical set of rules, which can be programmed based on simulation, experimental results and expert knowledge. This simplifies the process of gaining increased functionality for stability control, and can bring many benefits by allowing difficult control tasks to be accomplished in short time scales to a reasonable degree of performance. However, it does produce some problems. Not the least, fuzzy rules provide little scope for optimally controlling complex and multi-dimensional non-linear systems. This is because, by simplifying the control process to a level where fuzzy rules can be utilised effectively, the controller must operate in a simplified manner. As a result, the control logic cannot predict future states, cannot account for unanticipated disturbances, and must be totally re-evaluated if any vehicle modifications are performed.

ANN applications, on the other hand, offer great potential to the automotive industry, much of which has not been explored in great depth. ANN models offer the ability to model complex and multi-dimensional non-linear systems with reasonable accuracy; they remove the requirement to develop tedious and expensive mathematical representations complex systems; and they have potential in actually decreasing computational effort within controllers. These abilities then lead to scope for greatly increased stability controller functionality including greater sensory information, better utilisation of chassis control actuators, increased controller robustness, better control when tyres are in the non-linear transition region, reduced development time and cost, cheaper installation and adaptive control. As can be seen in the examples above, that
each of these goals are achievable at present – within bounds. The challenge of identifying the real potential in ANN modelling for stability control, thus, lies in the problem of how to broaden these boundaries. In this way, ANN control will have proven automotive application when the ANN control boundaries are similar to the operational boundaries of a generic vehicle. These boundaries include:

- Assuming unmeasurable/expensive parameters will be available to actual controllers and for ANN training;
- Using process simulation only to evaluate models;
- Using other models to train ANN models;
- Limiting the investigation/application to only one surface type;
- The susceptibility of constructed models to changes within the vehicle;
- Difficulties in obtaining data for adaptive learning; and
- Complexity of comparing performance to idea solutions.

Of particular note in this investigation is that almost all of the research on stability controller ANN models has been with the goal of “attempting to prove ANN models have potential use”. While this is useful to provide grounding for future work, this philosophy has persisted for over a decade of exploration. In fact, no studies have been found that attempt to construct and test stability controller ANN models that have a direct practical application in theory and in practice. If an ANN stability controller could be constructed and utilised in a robust manner within a vehicle it would provide compelling proof of the possibilities of ANN control, over and above the work already completed. As such, this is the principle objective of this investigation, and also covers the ANN surface identification concept that was highlighted earlier.

### 2.4 Research Method

The goal of this investigation is to build ANN tools that are founded in previous work, with the anticipated outcome of constructing functional ANN stability controller algorithms. This is a large scope, however, and must be defined in more detail.

Information on both actual value of $\mu_s$ coefficient and maximum value of friction coefficient $\mu_{max}$ for current driving conditions would be appropriated at the working process of advanced active safety systems.

V. Ivanob et al [19]
As discussed in previous sections, road surface characteristics (particularly coefficient of friction) are important input variables to stability controllers for proficient operation. These are also a group of parameters that are very difficult to estimate, and very expensive to measure. The potential ability of ANN to model and predict road surface features appears to be high, and is clearly worthy of additional research. Furthermore, if a robust ANN surface predictor can be developed to a level appropriate for widespread implementation, it will have a direct capacity within existing stability controllers to greatly increase performance. The development of such a surface predictor forms one of the goals of this investigation, within the bounds of available infrastructure, equipment and development time.

Prediction of surface features only form part of the potential of ANN modeling to stability control. As shown previously, ANN controllers can be developed to significant effect, and with major potential for the future. Furthermore, particularly little work has been done in this area using actual vehicles, so many of the benefits of ANN have not yet been realised in practice. In this respect, a stability controller utilising ANN modeling will be developed within this research, with the goal of implementing the controller on a real vehicle to test performance.

The goal of developing an ANN stability controller, in itself, is particularly broad in scope. In particular, developing and evaluating an ANN stability controller that incorporates ABS, TCS and VDC functionality is a difficult task, and is limited by available infrastructure, equipment and development time. Traditional stability controllers, for instance, are subject to extensive development that cannot be replicated here. It would be unwise for this investigation to try and imitate and compare systems that consume many millions of dollars in development in the commercial sector. In fact, if an ANN model were developed that could operate with similar performance to aspects of traditional control, the sheer discrepancy between commercial resources and the resources of this investigation would strongly indicate an advantage. Furthermore, if a broad ANN stability controller was developed, the level of experimental investigation required to effectively compare it to conventional VDC systems would require infrastructure, equipment, expertise and time that are simply lacking in the scope of this research. As such, the scope of the investigation must be reduced to a level where aspects of stability control can be realised in practical application, and compared to existing technologies within the limitations of the available resources. It is noted that
any methodology that is utilised in this way must be generic in nature, so proof in operation at this level provides compelling evidence for operational ability in more complex systems.

To this end, the process of driving a vehicle can be considered one of controlling vehicle linear and angular accelerations. The generic function of stability controllers are, therefore, to ensure that linear accelerations can be realised to their maximum in the driver’s desired direction while controlling angular accelerations to ensuring drivability. In this regard, any ANN control function within this scope that can compare with, or exceed, the functionality of traditional controllers will prove the ANN application in broader stability control.

Of particular difficulty is the ability of the stability controllers to determine how the maximum acceleration of the vehicle can be achieve in the desired direction. While this has been presented as being able to be determined from surface type, this is only one factor. The maximum force any tyre can transmit to the road is a function of many, often non-linear parameters in addition to road surface type. With knowledge of road composition the tyre slip that produces the maximum force can be estimated using empirical models, but this may give rise to significant controller error. Furthermore, as active steering, active damping, active spring, active anti-roll and active camber systems are adopted, this one-dimensional control output may extend into many dimensions. The problem of determining which slip produces optimum force will grow in complexity to which combination of tyre slip, tyre slip angle, tyre camber and tyre load produce optimum force. The complexity that these increases in functionality will induce in traditional controller algorithms will be significant, but also represents one of the greatest strengths of ANN. If an ANN tyre model can be used to predict the optimal combinations of these parameters for maximum force, it will have obvious and clear benefits.

Previous work has attempted to use ANN methods to model individual tyres. This work is considered to be limited in application because it requires the measurement or simulation of parameters that are difficult and expensive to determine in practice. It is also restricted in functionality because if all forces can be estimated at each wheel control, algorithms are still needed to utilise this information to determine appropriate control combinations. This presents new complications due to the potentially large
number of controlled variables and the non-linear nature of most of the controlled variables.

A different approach, investigated here, is to construct ANN models of the entire vehicle. The constructed models can be used in a goal-orientated approach, where the control output is calculated directly from driver linear and angular accelerations demands. The ANN models of the entire vehicle can be used to determine controlled variable combinations that will produce the required acceleration characteristics of the vehicle. While it is noted that such a method has significant potential benefits for controlling vehicle ride characteristics within the stable region, it has the capacity to maximise vehicle acceleration in emergency manoeuvres and ensure drivability. This is a generic stability controller goal and, as such, any ANN controller that fits within this framework can be shown to have generic application. Furthermore, when considering the difficulties of complex ANN controller design and testing presented above, this provides a basis from which to simplify the investigation without jeopardizing its practical application within the broad topic of stability control. The argument for this simplification process is given below, with reference made to the two following statements.

Longitudinal tractive effort could be achieved by providing at each wheel a driving torque consistent with the driver’s intent and to the maximum value dictated by the available tire patch friction.

S. Mohan et al [76]

Loss of vehicle control implies that the car has exceeded the coefficient of friction at the front, the rear, or both.

D. McLellen et al [25]

The longitudinal force produced at a tyre is a very important aspect when considering vehicle stability. Performance will increase up to the critical slip value, and then the tyre will become unstable with associated performance decrease. This is in contrast to lateral force, which always decreases with increased longitudinal slip. In such a manner it is possible to define the point at which maximum longitudinal force is developed as the threshold of instability. With any increase of slip past this value, the performance will always be less than what can be achieved within the stable region.

If the driver desires maximum acceleration through a combination of lateral and longitudinal forces, there is a specific slip at which this will occur. The determination of operating conditions that produce maximum force in the needed direction at each tyre is,
therefore, at the heart of stability controller algorithm development, and this concept can be applied to the vehicle as a whole. As such, the measured accelerations of the vehicle can be used to determine vehicle forces within an ANN model, and provide the training data that has been missing from previous studies. In this manner, the input – output relationship of the vehicle can be fully defined for ANN training, which also makes adaptive control possible.

The principle functionality goal of the ANN stability controller presented here is to determine the maximum acceleration in the longitudinal direction of the driven wheels that can be generated by the vehicle in a range of conditions. This information can be used to control vehicle actuators to realise this maximum. As such, this method excludes the additional requirements of stability controllers to control angular accelerations (such as yaw), which will not be considered in this investigation.

This single control goal simplifies the investigation considerably, and provides functionality to control a range of actuators to achieve maximum accelerations. However, controlling many actuators would involve a significant amount of work and expense to install and to train the ANN models. They would also be very difficult to test and evaluate with available resources, and would provide relatively small conceptual advances. As a result of this, the only controlled variable that will be used within this investigation is “percentage engine cut”, in which the ignition and injection pulses are controlled to produce a sliding scale of power delivered by the engine. This effectively provides slip regulation to the driven wheels under throttle and, because it operates through an open differential, the associated torque reduction to each wheel is approximately equal.

This second reduction in scope simplifies the control process to a stage that can be tackled appropriately, within the bounds and resources of this investigation. By effectively constructing an “Intelligent Traction Controller”, it is possible to develop systems that, with more resources, should be capable of wide ranging stability control. Although the remainder of the study will refer to the ANN controller being developed as a “traction controller”, the underlying control principles are more akin to the operation principles of modern VDC systems. This is because the functionality of attempting to maximise acceleration in the driver’s desired direction is more a function of VDC than TCS, which only attempts to avoid slip transitions into the unstable region. Furthermore,
the sensory information that will be utilised within the investigation is more comparable to VDC systems.

Finally, this research has the goal of developing an integrated and adaptable system for chassis measurement and control. This system allows research based on the outcomes of this investigation, and in significantly different technologies, to be carried out into the future. For instance, there are plans to integrate GPS speed zone detection and SMS anti-theft devices within the vehicle as part of the University of Tasmania “Intelligent Car” program. Furthermore, the test vehicle will be converted to run on hydrogen and as a hydrogen/petrol hybrid at the conclusion of this investigation, so the development of system that will support these is also of a priority. In particular, these systems must be highly flexible and each system that is installed should take into account possible future functionality.

2.4.1 Test Equipment

This research forms the latest chapter in the “Intelligent Car” series of projects, and was conducted on a real vehicle under real driving conditions. As such a new test vehicle (a 2002 Toyota Corolla) was acquired and fitted with a comprehensive sensor array, data logger, user interface, radio telemetry, PC mounted data acquisition and control device and a new engine management computer. Some of this equipment was cannibalised from the earlier Intelligent Car, which was a small Formula SAE racecar, which was instrumented to allow chassis data logging for off-line data analysis into ANN modelling aspects from which some of this study is based [40, 71, 87, 88, 89, 90].

As such, the hardware installation falls into three separate tasks. The first includes the transportation of the chassis sensors and data logger to the new test vehicle. This required finding ways to mount the sensors in appropriate positions and ensure adequate operation, designing new wiring looms that incorporate signal wire shielding, installation and configuration of the data logger unit and development of an appropriate and flexible user interface. This stage included the addition of a radio telemetry system for use in future research.

The second stage of installation required the test vehicle factory fitted engine management computer to be replaced with a new, fully programmable one. This step was required for the future hydrogen conversion, but also allowed for a great degree of
2. Vehicle Stability Background

engine control. In particular, the new engine computer contained the functionality to reduce engine power progressively (using ignition and injection cut) via a single external analogue control signal, tunable for traditional traction control. This gave the two fold benefit to this investigation of allowing simple closed-loop electronic control of the engine power and providing a traction controller from which ANN model performance could be compared.

The final stage of installation required the addition of a fully featured data acquisition and control device, similar to that used by W. Bartlett et al [13]. This was required because none of the other systems installed in the test vehicle had the functionality to allow for ANN control, but also because such a device would allow a very large range of possibilities for future research topics. This device was PC mounted, so required the installation of all the equipment required to convert a desktop PC to an automotive application. Furthermore, the process of integrating this PC with the other systems already installed within the vehicle required a range of data communications to be used, including analogue and digital signal wires, serial comms and the development of a CAN backbone.

Software development forms a significant part of the investigation, with all of the programming completed within the LabVIEW [91] environment. This programming includes the development of ANN models and training algorithms, ANN evaluation method, controller architectures and communications with input and output devices. LabVIEW was used in this case because it is developed in unison with the NI data acquisition and control device (DAQ) used, and because it contains all of the required functionality, is easy to program using graphical techniques and high quality support was available.

2.4.2 System Appraisal
Two, effectively different, ANN models are to be developed within this investigation. The first determines surface features from common measured variables. The second determines the maximum achievable vehicle acceleration in the driver’s desired direction, and to execute control so that this level of acceleration is realised when needed. To effectively assess these methods requires a thorough appraisal process, that attempts to base results on clear understanding of the ideal solutions in addition to observed performance.
For the surface identification problem, the appraisal process is highly dependant on what surface characteristics are investigated. Dynamic coefficient of friction, for instance, is very hard to measure and utilise within ANN models. Furthermore, it would also be very difficult to appraise such a model in a real driving situation because of the large degree of variation. As such, the surface identification algorithm presented here attempts to identify roads into categories that can be determined from human observation, as wet or dry for instance. In this application it is very easy to determine if the ANN models are behaving as expected, and the appraisal process should be simple. On the other hand, maximum achievable acceleration cannot be directly calculated, nor can it be observed simply. Optimum slip is a concept that is difficult to determine, and so it is not possible to directly appraise the operation of the Intelligent Traction Controller based on predetermined performance variables. Instead, statistical processes are introduced to appraise performance of the ANN controller based on observations of best performance.

In this respect, statistical data is compiled based on a number of simple manoeuvres that are repeated many times. The maximum acceleration that results from specific conditions can be statistically determined; in much the same way as a racecar traction controller would be tuned. This can be used to estimate the optimum slip for the driven wheels, and compared to the ANN controller prediction. Furthermore, the statistical data can be used to develop a traditional traction controller, which is comparable to the ANN controller. In this way, two things can be shown. Firstly, that the Intelligent Traction Controller can predict the slip which results in the maximum acceleration by comparing performance to statistical observation. Secondly, the ANN controller can be shown to have the same, or better, performance as a traditional controller. To this end, the first appraisal method would highlight the general abilities of the ANN model to determine optimum performance, and the second would show that the Intelligent Traction Controller is capable within the role. If this is the case, the results can be argued to philosophically extend to stability control generally.

2.4.3 Thesis Structure
The development of an ANN surface identifier and an ANN stability controller form a major aspect of this investigation. However, the development of appropriate hardware and software comprises a significant proportion of the research, as does the development of ANN models. As such the thesis is constructed as a logical progression through the
different stages of research, and is in approximate chronological order. In particular, the thesis contains chapters on the present ANN models, data logger installation, ANN surface identification research, engine management computer installation, PC and real-time data acquisition and control device installation and ANN stability controller research. All Appendices are included within the attached data DVD to provide additional information and comprehensive research results, and are generally not referred to within the thesis body. Furthermore, the developed software, ANN models, logged data, and other information are included for reference within the DVD. Finally, the accuracy of the references have been checked, and all installation and research work was carried out by the author unless explicitly stated otherwise.

2.5 Remarks
This section shows the current state in stability control technology. The broad accident statistics within the introduction chapter are narrowed to the observed effects within the community through the introduction of various types of active safety systems. In particular, the statistics from the USA on ABS are the most complete, and show that the stability control can produce mixed safety results. However, it is clear that stability control increases vehicle performance and that, coupled with relevant understanding and implementation reducing driver risk taking, can produce significant safety increases.

The discussion of tyre and vehicle dynamics provided a background to the fundamental aspects of stability control, namely gaining the required performance from the tyre/road interface. The complexities of fully understanding this problem were made clear through discussion and presentation of a number of tyre grip relationships. This background enabled a thorough discussion of “state of the art” stability controllers to be presented, as well as an indication of where this technology might lead. In particular, the roles of road surface identification and utilising intelligent systems with stability controllers were singled out as relevant avenues of study, and their backgrounds were covered in depth.

The chapter then concludes with a discussion of the research method to be employed, and highlights the fundamental assumptions that must be made to reduce the investigation to one within the scope of a single PhD project. In addition, there is a short discussion of the relevant problems with data acquisition and control, system programming and final evaluation of the results.
ARTIFICIAL NEURAL NETWORKS

The concept of Artificial Neural Networks has been presented in the previous chapter, in addition to a number of examples of their operation. However, a discussion of the specific design and mathematical implementation of ANN models is critical to obtaining an understanding of the modelling process. As such, this chapter portrays the general philosophy of ANN modelling, and provides specific ANN model algorithms.

The parallels of ANN modelling to the operation of the biological brain will first be discussed. This will be followed by the general principles of ANN operation, before the presentation of two specific ANN designs. In particular, Feedforward Backpropagation ANNs are presented as among the most common forms of ANN implementation, followed by Optimised Layer by Layer ANNs that represent a much newer accelerated learning type of feedforward ANN.

The considerations that must be taken into account to ensure effective ANN training and error minimisation are presented. Finally, the current state of ANN modelling and control within the automotive industry is broadly presented to provide a basis for further discussion. This is complemented by a brief introduction to the state of play of purpose built ANN hardware, which forms a pivotal role in realising the full potential of ANN modelling and control.
3. Artificial Neural Networks

3.1 ANN Operation

Artificial Neural Networks (ANN) attempt to mimic the operation of the brain at a neural level. As such, they exhibit some similar features, including the ability to learn and model process behaviour where *a priori* knowledge of the associated scientific principles is not available, or extremely difficult to obtain. This means that an ANN model can be programmed from observation of a system, without the need to develop the complex mathematical representations that would otherwise be necessary to characterise the inner workings of the system. Furthermore, they inherently associate items that they are taught, and physically group similar items together within their structure. This “generalisation” ability enables ANNs to operate with incomplete, noisy or partially incorrect data, to estimate results when presented new problems, and to act at slowly degrading performance levels during input sensor failures. In particular, ANNs can provide the following advantages [70, 86]:

- Adaptive learning – model can adapt itself based on training and observation;
- Self-organising – model organises its internal parameters while learning;
- Non-linear mapping – can learn and model complex and non-linear processes;
- Many input/outputs – models can be easily built with many inputs and outputs;
- Robust – can recognise relationships even when training is noisy or incomplete;
- Fault tolerance – input and network faults only lead to reduced performance; and
- Real-time operation – potential of very fast computation with parallel computing.

The way ANNs learn process behaviour, however, produces new problems. Like the biological brain, ANNs do not always behave as anticipated and are rarely exceptionally accurate. The generalisations that the ANN must make when learning produce general rules, which have limited ability in producing exact solutions [93]. Specific inputs produce general outputs that, even though they can model complex non-linear relationships, are not perfectly accurate. In fact, many ANN models exhibit absolute accuracies in the range of 90% only. Furthermore, and again just like the biological brain, ANN models are a “black box” system [85]. This means that it is not possible to determine how the ANN will behave simply by looking at it, although performance indications can be evaluated by observing its reactions to specific inputs. In this way, while it is desired that a certain input produce a particular output, how the network
achieves this output is left to a self-organising process. The accuracy with which it does this depends on the structure of the network including:

<table>
<thead>
<tr>
<th>Network properties</th>
<th>Neuron properties</th>
<th>System learning dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>• network topology</td>
<td>• summation function</td>
<td>• neural weight initialisation scheme</td>
</tr>
<tr>
<td>• types of connections</td>
<td>• activation function</td>
<td>• activation error calculation formula</td>
</tr>
<tr>
<td>• number of connections</td>
<td>• output function</td>
<td>• learning rule</td>
</tr>
<tr>
<td>• neural weight range</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.1.1 Artificial Neurons

Biological neural networks offer natural proof of the potential of ANNs. The neural structure of the brain provides an extremely powerful tool to recognise complex patterns and to generalise these past observations into actions of the future. In short, it can learn complex relationships from experience. By attempting to mimic these properties, ANN modelling has great potential, and is only limited by our understanding of the biological process and our ability to implement it.

The most basic element of the brain is the neuron, as shown in Figure 3.1. Based on our limited understanding of the operation of the brain, it is this cell that provides us with our abilities to think, remember and apply previous experience in our actions. Its operation is largely unknown but, basically, it receives inputs from a number of other neurons, combines them in some way, performs a non-linear operation to the result, and then outputs this to other neurons. The dendrites are hair-like extensions of the soma that
receive electrochemical data through the synapse of other neurons and act as input
channels. The soma then processes these signals over time and produces an output
within the axon, which is subsequently sent to other neurons through the synapses.

In the human brain there are approximately 100 billion of these cells, each
interconnected with anywhere up to 200,000 other neurons. The power of the human
mind is thought to come from the sheer number of these components and the connections
between them, including genetic programming and learning. Artificial neural network
research seeks to harness this process to produce intensely parallel computer algorithms
that act on past learning for pattern recognition, and not on complex programming and
modelling. [74, 94, 95]

However, the biological brain is extremely complex and there are many functions that it
performs that are currently unknown. As such, our understanding of how to implement
these processes within an artificial application is highly simplified, and is the subject of
intense biological and mathematical investigation. Furthermore, current computing
power cannot hope to emulate these processes in a timely manner due to the apparent
complex functions and the vast numbers of neurons and interconnections. The simplified
structure of artificial neuron in Figure 3.2 reflects this.

![Figure 3.2: Artificial neuron structure](image)

Previous neuron outputs $x_1, x_2, x_3 ... x_i$ provide data input to this neuron (neuron $j$) after
being sufficiently weighted, simulating the role of the dendrites. The neuron then
performs a summation function to gain a single value, which is then passed to the
activation function. The activation function applies a non-linear function to the result
before passing it to the output function, which performs some form of signal
conditioning. These three functions play the role of the soma and axon. The output of
the neuron is made available for use by other neurons in a similar manner to the operation of the synapses. [74, 96, 97]

The input function can thus be written as:

$$net_j = \sum_i x_i w_{ji}$$  \hspace{1cm} \text{Eqn 3.1}

where:  
\begin{align*}
net_j &= \text{summation function result for neuron } j \\
x_i &= \text{output from neuron } i \\
w_{ji} &= \text{weight factor applied to } x_i \text{ at neuron } j
\end{align*}

The goal of the activation function is to perform a non-linear operation to the summation result. This operation determines the function of the neuron, and different arrangements can alter the characteristics of the ANN. Typical activation functions include step functions, linear functions, ramping functions, sigmoidal functions and hyperbolic tangent functions [98]. The most common of these, however is the sigmoidal function, which has the form:

$$f(net_j) = \frac{1}{1 + \exp(-net_j)}$$  \hspace{1cm} \text{Eqn 3.2}

$$f'(net_j) = f(net_j)[1 - f(net_j)]$$  \hspace{1cm} \text{Eqn 3.3}

Finally, the purpose of the output function is to condition the activation function result before it is passed to other neurons. Generally however, this is not required and the output function does not normally perform any operation.

### 3.1.2 Network Structure

The simplified artificial neuron is implemented into the neural network, with a typical neuron accepting many inputs from other neurons and passing the calculation result to many more. This artificial neural network structure is highly simplified, and generally consists of a few dozen processing neurons only, which are usually organised into a simple structure. This small network size is required in practice because the combination of simplified neuron functions, simplified learning algorithms and computing limitations culminates to limit ANN performance when many neurons are used.

Particularly, it is the arrangement of the neurons and their connections that define the network architecture. Artificial neurons are usually arranged into layers – with neurons
in each layer grouped together because they behave in a similar manner (defined by their activation function and pattern of weighted connections). This ensures that the way neurons send and receive data within each layer is the same.

Although there are useful networks that contain only one layer, or even one neuron, most applications require more complex structures. The most common group of ANN types are “Feedforward”, in which inputs are accepted and passed through the ANN structure in a single direction to obtain the output. In this case, the neuron interconnections always feed in one direction and do not connect neurons together in a loop (i.e. so the output of one neuron is not passed in a neuron that was used as input to the first, as may be encountered in a biological brain). These types of networks are normally structured into layers of neurons, and normally contain at least three types of layers – input, hidden and output. The role of the input layer is to simply pass forward the presented input pattern to the neurons in the subsequent layer. It generally performs no computation on the data presented to it and is there simply to collect data from the outside world. Following the input layer are the hidden layers. There can be any number of hidden layers, although only one or two are normal, which perform the bulk of the ANN internal processing. The output layer neurons then correlate all of the data sent from the other neurons and send the result through the ANN output channels. [74, 96]

3.1.3 Multi-layer Feedforward Backpropagation (FFBP) ANN

The ability of an ANN to model process behaviour depends to a large extent on the network architecture. There are many proven architectures, in a range of applications, and new ones are continuously being developed. Among the most common is the multi-layer feedforward backpropagation (FFBP) ANN, which has been used by the manufacturing industry for some time for systems control. This architecture is well known and well documented [96, 99], and is able to approximate any continuous function to any degree of accuracy if large numbers of neurons are used [100]. It is also very simple to construct, is robust and often provides reasonable accuracy - but can take a comparatively long time to train through error backpropagation. [97, 99, 101, 102, 103]

BP ANNs are structured as shown on Figure 3.3, with an input layer, one or more hidden layers and an output layer, with \(i\), \(j\) & \(k\) processing units (called neurons, or nodes) in each respectively. The role of the input neurons is to simply pass forward the model input data to the hidden layer, and so the number of input neurons is equal to the number
of model inputs. The hidden and output layers then perform all of the network processing, with each neuron performing an input (summation) function of the weighted inputs, an activation function and an output function, shown in Figure 3.2.

Looking at the architecture, it is evident that changing the neural weights will alter the ANN characteristics. Iteratively modifying these neural weights until a state of minimum error is achieved gives the ANN the ability to ‘learn’ a process, called Network Training. As the name ‘backpropagation’ suggests, this is done by comparing model predicted outputs (at the output layer) with the training data, and propagating this error value back through the network to the input layer, updating weight magnitudes on the way.

Therefore, the error values for the output layer neurons are given by:

\[ \delta_k = (t_k - a_k) \cdot f'(net_k) \]

Eqn 3.4

where:
- \( \delta_k \) = error value for neuron \( k \)
- \( t_k \) = target training value for neuron \( k \)
- \( a_k \) = output value of neuron \( k \)

And the error values for the hidden layer neurons are determined using:

\[ \delta_j = \left[ \sum_k \delta_k W_{kj} \right] \cdot f'(net_j) \]

Eqn 3.5

where:
- \( W_{kj} \) = weight factor neuron \( k \) from neuron \( j \)

These error values can then be used to calculate the required weighting factor adjustments for the next training iteration, as shown:
3. Artificial Neural Networks

\[
\Delta W_{ji}^h = \eta \cdot \delta_j \cdot a_i + \alpha \cdot \Delta W_{ji}^{h-1}
\]

Eqn 3.6

where:
\( \eta \) = learning rate, \( 0 < \eta < 1 \)
\( \alpha \) = momentum constant, \( 0 < \alpha < 1 \)
\( \Delta W_{ji}^h \) = weight adjustment at iteration \( h \)
\( \Delta W_{ji}^{h-1} \) = previous iteration weight adjustment

Through this process the network weights can be continuously adjusted during the training process, with the aim of converging on the arrangement that will give minimum RMS error for the given ANN architecture, as shown by Frost [96].

Typical learning values used in training are \( 0.05 < \eta < 0.9 \) and \( 0.05 < \alpha < 1 \), with higher values improving convergence speed, and lower values enabling network ‘fine tuning’. Values should be chosen to produce reasonable convergence speed while maintaining the ability to converge to a specific solution, avoiding false minima [90].

3.1.3.1 Algorithm
As an aid to understanding the training procedure for the FFBP ANN, the following algorithm is supplied [104]:

1. Initialise the weights of the network with random values.
2. Use the ANN in a feedforward manner, by exposing it to certain process inputs with known process outputs.
3. Compare the known process outputs with the ANN predicted outputs to calculate the error. Backpropagate this error back to the hidden layer(s).
4. Adjust the weights of the network based on these errors at individual neurons. The level of this adjustment is affected by the learning rate and momentum constants.
5. Repeat step 2 to step 4 with all of the process input and known output patterns (i.e. for the training data) and compute the RMS error of the entire training set.
6. Cease training if the RMS error is within a tolerable range, otherwise repeat step 2 to step 5.

3.1.4 Optimised Layer by Layer (OLL) ANN
The Optimised Layer by Layer ANN was introduced by Ergezinger et al [105] in 1995, and has been the subject of significant investigation at the University of Tasmania by
Kiatcharoenpol [104]. As such, these two references are used heavily throughout this section.

Like the FFBP, this form of ANN is of the feedforward type, but is considered to yield “results in both accuracy and convergence rates which are orders of magnitude superior compared to backpropagation learning” [105]. The underlying difference between the two ANN types is the way in which the weights are updated during training. While the FFBP simply passes the error terms backward to iteratively update weights, the OLL observes the dynamics of the hidden and output layers and attempts to solve the weights exactly. In this way, the learning algorithm reduces the problem of optimising the interconnection weights of each layer to a linear one. This linearisation can then be used to calculate the exact weights required by the ANN to model the system, but does contain some linearisation error that leads to a need to complete some iterations of the learning process. As such, the iteration process includes a special penalty term that is utilised as part of a cost function to allow new iterations to be evaluated against past results. Unlike the FFBP, however, these iteration parameters do not need to be tuned by the user. [105]

The structure of the OLL ANN is similar to the structure of the FFBP ANN, as can be seen in Figure 3.4. However, there are three significant differences. The first is that the OLL structure only allows one hidden layer, which is a restriction based on the training method used. The second is the “bias” neurons that are used as inputs to the hidden and output layers ($x_0$ and $z_0$ respectively). These neurons can provide some extra modeling ability within the ANN, and are also commonly used in FFBP. The third, and most important, is that the activation functions within the output layer are not of the sigmoid
type, but are instead linear ($f(\text{net}) = \text{net}$). This is, again, a consequence of the requirements of the training function, and splits the ANN into a non-linear hidden layer, and a linear output layer. [104]

Since the ANN consists of non-linear and linear layers, the training process must be split in this manner too. As such, the weights for the output layer can be optimised exactly because they do not require further linearisation, then the hidden layer weights can be optimised by linearising the non-linear sigmoid functions. This linearisation of a non-linear process then leads to some error within the hidden layer weight calculation, which leads to the requirement of some iterative learning to reduce the cost function given in Eqn 3.7.

\[
E(R,S) = \frac{1}{P} \sum_{p=1}^{P} \frac{1}{2} (t^p - y^p)^2
\]

\text{Eqn 3.7}

where:
- $E = \text{cost function}$
- $p = \text{current training data pattern}$
- $t^p = \text{target output of the training data } p$
- $y^p = \text{network output of the training data } p$
- $P = \text{number of training data}$
- $R = \text{hidden weights matrix}$
- $S = \text{output weights matrix}$

The structure of the ANN then leads to the requirement to adjust weights one layer at a time, as the name “Optimised Layer by Layer” suggests. In particular, the output weights are adjusted first, and these new values used to adjust the hidden weights. The first step is to determine the optimum output layer weights ($S^{\text{opt}}$) based on the linear activation functions. This is achieved by calculating the gradient of the cost function with respect to output layer weights, and setting this to zero as shown in Eqn 3.8.

\[
\frac{dE}{ds} = \frac{1}{P} \sum_{p=1}^{P} (s^T z^p - d^p) z^p = 0
\]

\text{Eqn 3.8}

where:
- $s = \text{individual output layer weight}$
- $s^T z^p = \text{equivalent to the network output } y$
- $z^p = \text{hidden neuron scalar output for training data } p$
- $d^p = \text{target output at the training data } p$

This results in a set of linear equations that can be used to find the optimum output layer weights. By considering the general linear matrix representation in Eqn 3.9 the optimum weights can be determined by performing the calculations in Eqn 3.10 to Eqn 3.12.
Further, the matrix \( A \) will have a size of \((J+1) \times (J+1)\) dimensions, matrix \( b \) of \( K \times (J+1) \) and matrix \( S^{opt} \) of \( K \times (J+1) \). 

\[
A.S = b \quad \text{Eqn 3.9} \\
S^{opt} = A^{-1} h \\ 
A = \text{matrix } [a_{hj}] ; a_{hj} = \sum_{p=1}^{P} z^p_k z^p_j \quad h, j = 0..J \quad \text{Eqn 3.11} \\
b = \text{matrix } [b_{kj}] ; b_{kj} = \sum_{p=1}^{P} t^p_k z^p_j \quad j = 0..J; \ k = 1..K \quad \text{Eqn 3.12}
\]

The next stage of ANN training is to determine the optimal hidden layer weights (\( R^{opt} \)). This is done using a similar method as for the output layer, but the non-linear sigmoid functions first need to be transformed into a set of linear equations using a Taylor series expansion. The first step of this process is to replace the output layer with simple linear functions, with new linear weights between the hidden layer and output layer defined in Eqn 3.13.

\[
s_{lin} = f'(net_j) s_j \quad j = 1..J \\
\text{where: } s_{lin} = \text{linearised output weights} \\
f'(net) = \text{derivative of } f(\text{net}) \text{ at hidden layer neuron } j \\
s_j = \text{output layer weight from hidden layer neuron } j
\]

These linear weights simply represent the gradient of the hidden layer sigmoid activation functions, and are dependant on the training pattern that is being processed. These linear neurons can be implemented into the linearised network structure that is shown in Figure 3.5, which is used to determine the optimum hidden layer weights.

![Figure 3.5: ANN linearisation at output k for optimisation of hidden layer weights [104]](image)

Firstly, however, a new cost function must be derived for the hidden layer optimisation to account for linearisation error, and is shown in Eqn 3.14. The amount of change
required to achieve the optimum hidden layer weights ($\Delta R_{\text{opt}}$) can be derived by taking the partial derivatives of $E_{\text{linear}}$ and $E_{\text{pen}}$ with respect to $\Delta r_{ji}$ by using the chain rule, and setting them to equal zero. This is shown in Eqn 3.15.

$$E_{\text{hidden}} = E_{\text{linear}} + \mu E_{\text{pen}}$$  \hspace{1cm} \text{Eqn 3.14}$$

$$\frac{\partial E_{\text{hidden}}}{\partial \Delta r_{ji}} = \frac{\partial E_{\text{linear}}}{\partial \Delta r_{ji}} + \mu \frac{\partial E_{\text{pen}}}{\partial \Delta r_{ji}} = 0$$  \hspace{1cm} \text{Eqn 3.15}$$

where: $E_{\text{hidden}} = \text{overall error for hidden layer}$
$E_{\text{linear}} = \text{error for linearised activation functions}$
$E_{\text{pen}} = \text{penalty term to account for linearisation error}$
$\mu = \text{penalty constant}$
$r_{ji} = \text{individual hidden layer weight}$

The optimal change to the hidden layer weights can be expressed as a matrix by deriving and depicting a set of $(I+1).J$ linear equations into matrix form, as shown in Eqn 3.16 to Eqn 3.19. In this case the matrix $\tilde{A}$ has a size of $[(I+1).J][(I+1).J]$ dimensions, matrix $\tilde{b}$ of $[(I+1).J] \times 1$ and matrix $\Delta R_{\text{opt}}$ of $[(I+1).J] \times 1$.

$$\Delta R_{\text{opt}} = \tilde{A}^{-1} \tilde{b}$$  \hspace{1cm} \text{Eqn 3.16}$$

for $j \neq h$:

$$a_{ijm} = \sum_{p=1}^{P} \sum_{k=1}^{K} \left[ \left( \sin_{h}^{p} x_{i}^{p} \right) \left( \sin_{h}^{p} x_{m}^{p} \right) \right]$$  \hspace{1cm} \text{Eqn 3.17}$$

for $j = h$:

$$a_{ijm} = \sum_{p=1}^{P} \sum_{k=1}^{K} \left[ \left( \sin_{h}^{p} x_{i}^{p} \right) \left( \sin_{h}^{p} x_{m}^{p} \right) + \mu \left[ s_{ij} \right] f''(\text{net}^{p}_{j}) \right]$$  \hspace{1cm} \text{Eqn 3.18}$$

where: $s_{\text{lin} k}^{p}, s_{\text{lin} h}^{p} = \text{linearised weight from neuron } k \text{ at output layer to hidden neuron } j,h \text{ for training data } p$
$x_{i}^{p}, x_{m}^{p} = \text{input of neuron } i,m \text{ at the input layer}$
$s_{ij} = \text{output weight from hidden neuron } j \text{ to output neuron } k$
$f''(\text{net}^{p}_{j}) = \text{second derivative of } f(\text{net}) \text{ at hidden neuron } j$

This process requires each of the $\tilde{A}$ has $\tilde{b}$ matrices to be calculated for every training pattern, with the final summations used to calculate $\Delta R_{\text{opt}}$. The new hidden layer weights ($R_{\text{new}}$) can be calculated from Eqn 3.20.

$$R_{\text{new}} = R_{\text{old}} + \Delta R_{\text{opt}}$$  \hspace{1cm} \text{Eqn 3.20}$$
The linearisation error in calculating $\Delta R_{\text{opt}}$ means the process cannot be carried out in a single step. As such, an iterative procedure is needed to alternatively optimise the output and hidden layer weights to obtain the minimum error.

### 3.1.4.1 Algorithm

The OLL training procedure is then given in the following algorithm [104]:

1. Initialise the weights of the network with random values.
2. Use the ANN in a feedforward manner, by exposing it to certain process inputs with known process outputs with networks weights $R$ and $S$.
3. Compute the optimal output layer weights $S_{\text{opt}}$ (Eqn 3.10), and update the ANN with these values.
4. Use the ANN in a feedforward manner again with the new $S_{\text{opt}}$ weights, and calculate RMS error ($\text{RMS}_{\text{current}}$).
5. Compute the optimal hidden layer weight change $\Delta R_{\text{opt}}$ (Eqn 3.16), and update the ANN with $R_{\text{test}}$ based on Eqn 3.21.

$$R_{\text{test}} = R_{\text{old}} + \Delta R_{\text{opt}}$$  \hspace{1cm} \text{Eqn 3.21}

6. Use the ANN in a feedforward manner again with the new $S_{\text{opt}}$ and $R_{\text{test}}$ weights, and calculate RMS error ($\text{RMS}_{\text{test}}$).
7. If $\text{RMS}_{\text{test}} < \text{RMS}_{\text{current}}$ then:
   a. update the hidden layer weight matrix so $R = R_{\text{test}}$
   b. set $\text{RMS}_{\text{current}} = \text{RMS}_{\text{test}}$
   c. decrease the penalty constant by:

$$\mu_{\text{new}} = \mu_{\text{old}} \cdot \beta$$  \hspace{1cm} \text{Eqn 3.22}

where: $0 < \beta < 1$ (normally $\beta = 0.9$)
   d. continue to step 9
8. If $\text{RMS}_{\text{test}} \geq \text{RMS}_{\text{current}}$ then:
   a. decrease the penalty constant by:

$$\mu_{\text{new}} = \mu_{\text{old}} \cdot \gamma$$  \hspace{1cm} \text{Eqn 3.23}

where: $1 < \gamma$ (normally $\gamma = 1.2$)
   b. repeat from step 5
9. Cease training if the $\text{RMS}_{\text{current}}$ error is within a tolerable range, otherwise repeat step from step 2.
3.1.5 Training Considerations

The purpose of network training is to assign each of the ANN weights with a unique real number to enable the network to perform the transform that yields the required outputs with maximum accuracy. This means that every ANN must be programmed with different weights for different applications and, because it is impossible to compute the weights directly, the network must be trained. This involves presenting the network with a set of measured data from the system it is to model, which it then uses to assign its own values to the weighting factors and is referred to as “supervised training”. The way it does this depends on the network type, the acquired training data and the learning rules it uses, but is always a repetitive and iterative process that can be very time-consuming.

In order to produce both good performance and a reasonable learning rate from the ANN, a number of factors must be considered. In particular, the learning rules used can have a great effect on how quickly the ANN learns the process. The training data must offer an adequate representation all of the operating conditions of the system to be modelled, as well as being presented to the network in a particular way. Network architecture is also an important consideration. The number of hidden layers, the number of neurons within them and the way they are interconnected greatly effects network performance. [74, 96, 99]

Firstly, it is important that training data sufficiently represents all operational aspects of the system to be modelled. In particular, the dynamics of a system normally contains a number of input subgroups that have their own tendency towards a particular output prediction. As such, each subgroup must be adequately represented within the training set to allow ANN training of the complete system. Where noise is present within a system, each subgroup must contain enough data within it to include the effects of statistical variation of the process.

It is important to ensure that the order in which each subgroup is presented to the system is spread out. If the network is trained with just one example at a time (called a “pattern”) in the order that they were measured, the weights set meticulously for one fact could be drastically altered in the learning of another. In short, it may forget previous lessons when learning something new. This is undesirable, and the training set should ensure that the ANN learns everything together so it assigns weights that suit the entire
system. When learning off-line this is normally accomplished by randomising the order of training patterns, replacing the time series of data with a randomised series.

In addition to verifying that the training data is sufficiently represented, network performance can usually be improved by normalising the training data to ensure each input has similar magnitudes. This makes sure that the network is not biased towards inputs that are of a higher magnitude than others in the training set, which can create training problems. Normalising the output is also an important step towards improving network performance, since most training algorithms attempt to minimise the total error of the outputs. Using data that is not normalised will cause the network to train the output with the largest magnitude (and thus statistically the largest error) to be as accurate as possible, to the exclusion of the accuracy of other outputs.

The black box nature of the ANN models means that, once trained, their predictive performances must be observed to obtain estimates of their accuracy. In particular, the network could have made a number of generalisations within the test data that are not supported in reality. It is therefore important to gather a second set of data to be run through the ANN so that a comparison can be made between the desired output and the actual output. This is referred to as the “testing data”, and if the network cannot produce the desired accuracy using this data it may have to be redesigned or the training set may need to be broadened.

One important consideration is the number of internal neurons - too few will starve the network of the resources it needs, while too many will increase the training time and could cause overfitting. Overfitting can be a particular problem because it causes the network to memorise the training data, rather than generalise it, as shown below in Figure 3.6. In this case, the graph on the centre shows a good generalised fit to the somewhat noisy training data, while the graph on the right has created a curve that fits all of the training data very well but does not reflect the true data relationship, and the graph on the left has not learnt the process well. This also highlights one of the principle benefits of utilising a completely new set of data for ANN testing, because this should clearly show that the training generalisations are not sufficient, or if overfitting has occurred.
The number of input layer neurons can affect the accuracy of the network. In particular, the addition of input parameters that have little or no influence on the system outputs can significantly increase the network error because the ANN is forced to waste resources trying to identify relationships that are not there. In a similar manner, over-representing specific input parameters within the model can reduce accuracy. In this case using, say, engine crank speed and engine cam speed as ANN inputs will confuse the training process, because these parameters are related (crank speed = k * cam speed). It is therefore important to identify the minimum number of inputs required to successfully model the system for optimum performance [96, 99].

Therefore, the most appropriate ANN model of a process can be determined by adding training data as needed, iteratively altering the internal architectures of the neuron layers and iteratively removing or adding appropriate input parameters. This represents significant investigation, especially when considering that the training times for large ANNs can be in the order of days or weeks. As such, after an ANN has shown its capacity to model a system within reasonable error bounds, the process of finding the best ANN model can be very time consuming. However, once the best architecture is identified, ANN modelling becomes very simple and usually requires less processing power and time than traditional mathematical models. In the field of automotive technology the potential benefits that this can offer are large, and many investigations have been completed into different applications.

### 3.2 Automotive ANN Applications

Aside from the surface identification and stability control applications presented previously, ANN modelling and control has been pursued in a range of automotive technologies. In order to gain some additional background on where this investigation fits in, some of these are briefly discussed below.
3.2.1.1 Automobile Autopilot
In 1987 Shepanski et al [106] introduced a model called “an automobile autopilot” based on a highly simplified computer simulation of traffic conditions. The model assumed that when travelling on a wide shouldered two-lane freeway other vehicles would perform various pre-programmed manoeuvres, without consideration for the ANN controlled vehicle other than to avoid running into its rear end. The model sought to decide when the ANN driven vehicle should change lanes, and then to adjust vehicle speed and heading to perform the required manoeuvre. This decision was based on ANN inputs such as distances and relative speeds between objects and road curvature.

Two ANNs were then trained through back propagation to control the vehicle based on simulated data during lane changes. A steering angle network was designed purely to keep the car within the lane it was travelling in, while another network was used to decide when to change lanes and to control vehicle speed. The networks demonstrated that, while the control model must operate within specific limits, it could provide intelligent control of the vehicle. It also showed that the driving style for the driver during the training stage was reflected in the control algorithm of the ANN.

3.2.1.2 Driver Override in Crisis Situations
Research conducted in 1999 by Jayakumar et al [107] chose to further investigate the use of ANNs in vehicle control. The focus was on correcting driver mistakes when a crisis situation arose by overriding driver inputs and controlling throttle position, brake pedal position and steering angle in an effort to prevent road departures.

Again, the study used a simplified computer simulation of the vehicle dynamics to both train and test the control ANN. Training data sets from manually driven crisis scenarios within the computer simulation were used to program the neural network. Interventions were invoked within the developed ANN controller based on driver inadequacies, such as a delay in human response or inappropriate control input. The study showed reasonable results but was, however, limited to a travel speed of only 20km/hr and simple road curvature types to avoid tyre slip problems and other complex effects.

3.2.1.3 Static Suspension Tuning
Previous work by the author [90] presented an application of ANN to predict the optimum chassis arrangement for a steady state cornering condition for a racecar, which
3. Artificial Neural Networks

traditionally can require a large amount of practical testing to determine. Likewise, traditional computer modelling was identified as having a growing use in this field, but due to the extremely complex nature of the vehicle / driver / environment entity it was considered to have a number of practical limitations.

The study used a racecar equipped with a large sensor array, including engine speed, throttle angle, wheel speed, suspension position, steering angle, longitudinal and lateral acceleration and yaw rate. Chassis tuning was accomplished by varying caster, toe and front and rear tyre pressures. Data was collected for a total of six different chassis tuning combinations for the steady state cornering condition and a feed-forward back-propagation ANN model capable of predicting the lateral (centrifugal) acceleration of the vehicle for any given chassis tuning was produced. A numerical investigation was then completed with the ANN model to find the maximum lateral acceleration, and therefore speed, of the vehicle for each different possible chassis tuning combination. Each of the resulting 480 combinations were ranked and compared against the optimal combination found from extensive practical vehicle testing. A high degree of correlation was found between the ANN and traditional chassis tuning optimum arrangements, despite the ANN model requiring significantly less practical testing.

3.2.1.4 Vehicle Drivability Assessment

This work, presented by Schoeggl and Ramschak [72], attempts to provide vehicle developers with brand specific drivability characteristics using ANN. The ANN model can be used to estimate various parameters that represent drivability based on engine and vehicle design and tuning data. Manufacturers could use this tool to develop vehicles that meet the requirements of the market without extensive empirical assessment.

This method uses the ANN models to simulate complex and subjective observations of drivability relevant operating states to produce objective estimations of how they can be realised. These results can be used in vehicle research, development, calibration and quality tests to allow target drivability and marketing criteria to be optimised on the test bed. Furthermore, the study shows that this method results in improved quality of drivability, a reduction in the number of cost intensive prototypes and a reduction in calibration time of up to 40%.
3.2.1.5 Simulating Customer Response to Vehicle Noise
Jennings et al [108] also presents another ANN method for determining desirable vehicle characteristics, this time with the aim of evaluating sound quality within new vehicles. Normally this evaluation is completed by “jury” studies, whereby listening studies are completed in a sound laboratory at significant cost and time. The study presents an ANN approach that enables objective measurements of vehicle sounds to be converted into predictions of the subjective impression of potential customers, without the need to carry out the jury testing procedure. However, the study only focuses on the ANN design principles of the possible models, and does not provide any results.

3.2.1.6 Electric Car Fault Diagnosis
The research conducted by Kalogirou et al [109] shows another application for automotive ANN modelling. In this case an electric car is used that contains two electric motors, with the ANN model used to determine if faults are developed within the system. In particular, it is observed that a fault will cause the temperature of each motor to change in an unusual way, and the ANN model reflects this. Six inputs (covering energy used and motor temperature histories for each motor) are used within a single ANN model to predict the current motor temperatures, with these predictions then compared to the actual motor temperatures. Based on the observation that the ANN should only model the “no fault” case, any significant error between the predicted and measured values should then indicate a fault. No results for this research, however, are provided.

3.2.1.7 Prediction of Engine Torque and Emissions
ANN modeling has a use in engine applications. For instance, Arsie et al [103] investigated the capability of ANN modeling in predicting engine torque and emissions. In this work, air flow, fuel flow, ignition advance and engine speed are used as inputs to train an ANN model, which is then capable of predicting engine torque and hydrocarbon, carbon monoxide and NOx emissions with reasonable accuracy when the engine is used in the field. The work then argues that this ANN model can then be used to determine the rapid prototyping and optimal design of engine control strategies.

3.2.1.8 Engine Misfire Detection
ANN modeling has shown potential in detection of engine misfire events, as Tawel et al [110] shows. Cylinder misfire is normally determined by observing the acceleration variations of the engine crankshaft, where a misfire should result in a brief deceleration.
However, this observation is made significantly difficult due to the torsional dynamics of the crankshaft, and this method of identification is limited in practice. Instead, an ANN model is present that predicts if a misfire has occurred (no misfire when output = 1, misfire when output < 1) based in inputs of crank acceleration, crank speed, air flow and a signal identifying which cylinder combustion should take place. The model then was shown to exhibit a misclassification error of only 1%, which was considered to be well within the acceptable range.

3.2.1.9 Prediction of Engine Air-Fuel Ratio
The final example attempts to utilise ANN modelling as an aid in providing fine control over air-fuel ratios. In this work, Alippi et al [111] observes that even a 1% variation in air-fuel ratio can lead to a 50% reduction in the efficiency of catalytic converts in reducing pollutants, and that better control is required to meet more stringent regulations. In this way, an ANN model is proposed that predicts the current air-fuel ratio based on inputs of engine speed, throttle position, fuel injection pulse width and manifold pressure for times t, t-1 and t-2, and previous air-fuel ratios for times t-1, t-2 and t-3. The model is then shown to predict air-fuel ratios with 1.3% mean error.

The research goes on to develop an ANN controller of the fuel injection pulse width to implement fine control of the air-fuel ratio. In this case, the ANN presented previously is rearranged to operate as an inverse ANN model, where engine speed, throttle position, manifold pressure and air-fuel ratio for times t, t-1 and t-2, and previous fuel injection pulse width for times t-1, t-2 and t-3 are used to determine the required fuel injection pulse width. The ANN model is then capable of determining the appropriate injection pulse width to provide the desired air-fuel ratio, which is then used as a control signal. The resultant controller was tested to produce “encouraging” results, with its performance highly comparable to traditional control methods.

Real-time control, however, often requires specialised hardware to take advantage of the high speed parallel computing potential of ANNs.

3.2.2 ANN Hardware
At present hardware that has strong parallel processing features are rare, with most devices utilising the serial computational techniques used in computers. However, if the very high modeling and control speeds that ANN are capable of are to be realised,
hardware that reflects their structures are necessary. In fact, the limitations of implementing ANN models with serial processors has been identified as a significant contributing factor to the limited development of real-time ANN embedded control systems [86]. Sitte [112] and Tawel et al [110], however, present two different types of ANN hardware for automotive applications.

Sitte, for instance, observes that the “high levels of parallelism means high computing speed (can be achieved), and what is more significant, the computing speed is practically independent of the size of the network”. As such, a “Local Clustern Neural Network” is developed within an analogue integrated circuit, and achieves a high degree of parallelism. Furthermore, the ANN application of the circuit means it can be designed so that the required computing structures use currents, and not the voltages normally required for digital logics circuits. Therefore the circuit can be built with only a few transistors in a small and low cost silicon chip, with minimal power consumption.

In practice, the LCNN chip accepts six analogue inputs, produced one output and contains eight clusters of LCNN neurons, although additional neurons can be implemented by connecting additional chips in parallel. The ANN weights are also stored on the chip within digital storage cells, and are downloaded to the chip using serial communications with a computer. As such, the chip contains no training capability, and ANN training must be completed on an external computer.

The work by Tawel et al [110] presented above on engine misfire detection requires very fast prediction of misfire events. As such, it places very high demands on the ANN hardware, which traditional serial computation would have trouble meeting. To this end, the work presents a “Neuroprocessor” chip with the aims of providing a mass marketable, flexible and accurate product. The operation of the chip is described as:

The architecture consists of: (1) a global controller; (2) a pool of 16 bit-serial neurons; (3) a ROM based bipolar sigmoid activation lookup table; (4) neuron state registers; and (5) a synaptic weight RAM.

In this design, both inputs to the network as well as neuron outputs are stored in the neuron state RAM. When triggered by the global controller, each of the 16 neurons performs the multiply and accumulate operation. They receive in a bit serial fashion as input the synaptic weights and activations from either (a) input nodes or (b) outputs from other neurons and output the accumulated sum of partial products onto a tri-stated bus which is commonly shared by all 16 neurons.

Tawel et al [110]
Each of the 16 neurons embedded within the neuroprocessor operate in parallel, and a significant increase in ANN computational speed is realised (with a 4 input, 15 1st layer, 7 2nd layer, 1 output neural architecture observed to take less than 80μs to compute). Furthermore, when numerous hidden layers are required the 16 neurons can be reused for each layer, which reduces the cost of the chip.

3.3 Remarks
This section illustrates that ANNs have a number of performance benefits that show a potential in vehicle control methods. The abilities of neural networks to be trained, rather than programmed, were shown to provide an opportunity to model complex systems that would otherwise be extremely hard to model using conventional techniques. They have also been shown to provide model solutions with minimal computation, and reasonable result accuracy as sensory inputs fail. Previous research has been shown to attempt to apply these benefits to various automotive applications to varying levels of ability. However, it seems that there is still a wide gap between ANN models used in simplified examples and models that can be used in real life vehicle scenarios. Finally, with the impending widespread development of neuroprocessors the advantages of high speed ANN modelling and control of complex systems will be realised in practice. As such, the potential of artificial neural networks in wide-ranging applications looks hopeful.
This section will present the test vehicle and the process of installing and configuring a data logging system. This data logging system is exported from a previous test vehicle, modified and installed onto the new one, and includes an array of chassis sensors, the data logging device and a number of ancillary devices.

This installation forms the first of three stages that will be discussed later. Specifically, it provides the data logging capability required to complete the ANN road surface identification research, but is insufficient for the ANN stability controller development. It is completed first because the simple access to the hardware and previous experience with the system enabled rapid progress to be made in the early stages of investigation at minimal cost and with a high degree of flexibility, while providing a means to test sensors for installation reliability.
4.1 Test Vehicle

Previous work within the “Intelligent Car” project saw the design and construction of a racecar incorporating a comprehensive data logging arrangement for off-line ANN investigation into a variety of automotive applications. The principle arrangement for this data logging was through MoTeC’s Advanced Dash Logger (ADL), which powered and measured an array of chassis sensors, as well as providing a programmable display of engine and chassis parameters to the driver. The installation measured parameters such as wheel speed, suspension position, brake pressure, steering wheel angle, attitude and acceleration – as well as incorporating a radio modem for remote data collection and some addition devices to enable RS232 communication [40]. The previous test vehicle and ADL installation is shown in Figure 4.1.

![Figure 4.1: Advanced dash logger (ADL) in previous Intelligent Car](image)

These parameters were required for this investigation and, because the original vehicle no longer had use for them, the entire system (excluding brake pressure sensors and
wiring loom) was transferred to the test vehicle used here. This required installing all non factory fitted sensors, as well as designing and constructing a new wiring loom for the ADL and fitting a number of switches, indicator lights and buzzers. The new test vehicle, a 2002 Toyota Corolla Ascent hatchback with 1ZZ-FE VVTi engine, is shown in Figure 4.2, and details of the components given in the Appendix.

### 4.2 Chassis Sensors

Chassis sensors are defined here as any sensor that provides information on the performance of the vehicle as a whole, including sensors that detect driver demands. This basically includes any sensor that is not directly related to engine control, although it is observed that throttle opposition falls into both categories.

In this investigation, the chassis sensors were used to measure:

- Wheel speed at each wheel
- Suspension position at each wheel
- Steering wheel angle
- Longitudinal, lateral and normal acceleration
- Roll, pitch and yaw angles
- Roll, pitch and yaw angular rates

This represents a relatively comprehensive array of sensors, especially when all of the data that can be derived from them is considered. For instance, suspension position can be calibrated to give suspension spring force, and can be differentiated to gain suspension velocity, which is useful in evaluating damper characteristics.

The placement and installation quality of the sensors is arguably more important than their quantity, as increased noise and vibration may require excessive data filtering and make these secondary measurements worthless. To this end, it is important to understand how the sensors operate, and to install them in a way that is most suited to their design.

#### 4.2.1 Wheel Speed Sensors

The wheel speeds of the test vehicle are measured using a number of Hall effect sensors placed in appropriate positions within the wheel rims. These sensors directly measure
the angular velocity of each rim in a digital manner, which can be translated into linear wheel speeds (at the ground) using the tyre geometry.

The sensors operate on the principle that when a current carrying conductor is placed into a magnetic field, a voltage will be generated perpendicular to both the current and the field. This is termed the Hall effect, after Dr E. Hall who discovered it in 1897. When a current is flowing perpendicular to a magnetic field, the Hall effect is observed as a potential difference developed across the material, with a value proportional to the current and the magnetic field intensity. Further, if there is no magnetic field present, the system will not induce any Hall effect voltage. The consequence of this is that a digital sensor, based on the Hall effect, can be used to measure whether or not a ferrous material is in close proximity to it. Such a sensor benefits from no moving parts, or any need to come in contact with the object being measured, meaning it can expect a long life and can operate at a very high switching rates [113].

Moreover, by incorporating the sensor into a ferrous gear tooth arrangement, as shown in Figure 4.3, the sensor can be used to measure rotational speeds for many different applications. The benefits of Hall effect sensors are clear, and they are becoming widely used within the automotive industry.

![Figure 4.3: Gear tooth Hall effect sensor [113]](image)

The Hall effect wheel speed sensors used to measure wheel speeds on the test vehicle were manufactured by Honeywell (GT1 series) and supplied through MoTeC. The sensors operate with a maximum switching time of 15μsec on a rising edge and 1.0μsec on a falling edge, a minimum tooth width of 2.5mm and a minimum tooth spacing of 10mm. The sensor is composed of an integrated circuit that is made up of discrete capacitors and a bias magnet, as can be seen in Figure 4.4. It is sealed in a probe type, non-magnetic plastic casing.
The Hall effect sensor incorporates three wires. One wire is the supply voltage (5V), the second is the ground (0V) and the third is the signal wire that provides the sensor measurement value. As each gear tooth passes, the sensor detects the change in magnetic flux level away from its in-built bias magnet, as depicted above, and digitises the result. This creates a stepped output within the signal wire, with the digital output switching between the supply voltage when it passes a gap in the gear and the saturation voltage (0.4 V) when it passes a tooth.

### 4.2.1.1 Sensor Considerations

The performance and accuracy of the Hall effect sensors are dependent mainly on the way they are positioned in relation to the target material, and on its magnetic characteristics. The shape and number of teeth on the gear-tooth wheel can significantly affect the accuracy of the results, as can sensor clearance with the teeth and sensor vibration. Obviously, incorrect sensor installation can give rise to erroneous results, and care must be taken to ensure this is avoided. However, the operation of the sensor is such that improper installation will produce clearly erroneous results, and as such its accuracy is almost wholly dependent on the geometry of the gear-toothed disc. Different gear-toothed discs will produce different measurement properties, as can be seen from the calculations in Figure 4.5.

![Figure 4.4: Honeywell Hall effect sensor configuration [113]](image)

![Figure 4.5: Effect of number of gear teeth on sensor properties (50km/h)](image)
As the numbers of teeth on the disc are increased, there is a corresponding increase in the sampling resolution of the sensor. This is desirable because it means that small velocity variations as the wheel rotates will be able to be measured. However, with increasing sampling resolution of the sensor, the time it has to “switch” from one tooth to the next decreases. Since the switching time of the Hall effect sensor is constant this means that the error will increase linearly with increasing the number of teeth. By measuring on the “Falling Edge” this error can be substantially reduced. Finally, the graph shows the minimum possible diameter of the gear-toothed disc for the number of teeth from the specifications. This is also an important consideration, as the disc must fit within the vehicle rims.

4.2.1.2 Sensor Installation
The wheel speed sensors are mounted differently front and rear. The rear Hall effect sensors operate off custom machined mild steel gear-toothed discs that fit over the brake rotor “hat” and provide 40 teeth/revolution, while the front sensors utilise the rim itself with 16 teeth/revolution. This was because space confinements limited mounting options considerably, with the front wheel assemblies containing much larger brake rotors and calipers than the rear.

4.2.1.2.1 Driven Wheel Speed Sensors
The final mounting position of the front (driven) wheel speed sensors is shown in Figure 4.6. The Hall effect sensor is positioned between the rim and the brake rotor, sits next to the brake caliper and points outward. This is because the internal contour of the rim is such that the entire outer profile of the brake caliper is contained within it with small clearances of approximately 10mm, leaving very little space to mount a gear-toothed disc. As a result of this, it was decided to use the ventilation holes in the rim as the gear-toothed disc, and the Hall effect sensors was positioned to this effect. A 20mm wide and 5mm thick aluminium bar was chosen to hold the sensor in rigidly place and was mounted off one of the caliper bolts, which was lengthened to accommodate the addition. The aluminium bar then bends up and around the brake rotor with sufficient clearance to avoid excessive heat, and the Hall effect sensor is bolted in place using two mounting holes. This provides a sufficiently rigid bracket to hold the sensors in place, but also allows easy sensor placement adjustment by altering the bracket geometry.
Further, the rim (which now also acts as the gear-toothed disc), contains 16 evenly spaced ventilation holes in a circular pattern, at a 140mm radius. Each of the holes are of a circular shape, and have an internal diameter of 28mm. The spacing between the hole edges are thus 25mm, and comprises of 3mm thick mild steel sheet.

It should be noted, however, that this is a less than optimal solution and contains three additional sources of error. The first is that it is not recommended that the gear teeth be of circular shape, as is the case here. This is because as a “tooth” passes the sensor, the exact position of the rising or falling edge is not clearly defined. The second problem is that the quality of the machining of the rim is unknown, and there may be some variation between the width of each “tooth”. The third problem is that the 3mm thick steel rim is too thin based on MoTeC’s specifications. None of these factors are considered significant sources of error as the first is repeatable, the second is assumed to be negligible, and the third is offset by the excessive tooth width and target thickness of the “teeth” in the rim. The accuracy of the sensor alone in measuring the angular velocity is, therefore, given in Eqn 4.1 and expressed in Figure 4.7.

\[
Error (rad/ sec) = \dot{\theta} - \left( \frac{2\pi}{\text{No.Teeth}} \right) - \left( \frac{1}{\theta} \times \frac{2\pi}{\text{No.Teeth}} \right) + \text{SwitchingTime}
\]

Eqn 4.1
4.2.1.2.2 **Rolling Wheel Speed Sensors**

The rear (rolling) wheel speed sensors are installed in a different manner than the front, and are shown in Figure 4.8. This is because, although the rims are the same size, the rear brake rotors and calipers are smaller than at the front, giving more clearance and more mounting options.

The Hall effect sensor was mounted in a similar arrangement as on the front wheel, but that the custom built gear-toothed disc has been placed over the brake rotor “hat”, and is held in place when the rim is installed. The final manufactured mild steel disc is given in Figure 4.9, which highlights its “C” shaped cross section. This cross section was chosen to provide significant target thickness for the gear teeth and to allow the disc to be held in place by the rim, aiding simple removal if the need arose. By utilising tight clearances within the design, it was possible to ensure that the disc was placed in alignment with the wheel and “on-centre”. It should be noted, however, that this design requires the rim to be offset marginally due to the thickness of the disc over the “hat”. This 3mm disc
thickness was the minimum allowable to ensure accurate machining of the disc, and the resulting rim offset is considered to be negligible.

**Figure 4.9:** Production of rear (rolling) gear-toothed discs

**Figure 4.10:** Geometry of rear gear toothed discs

Inspection of the geometry of the disc, given in Figure 4.10, shows that the gear-toothed disc contains 40 teeth, with a target thickness of 20mm, tooth height of 5.9mm, tooth width of 4.4mm, and tooth spacing of 11mm. All of these values exceed the specification given by MoTeC, and it is acknowledged that up to 50 teeth could have been machined into the disc at minimum specification. The choice of 40 teeth, however, was based on machining difficulties. Firstly, it was observed that the manual controlled milling machine to be used was designed to easily machine 40 evenly spaced segments into circular objects. While it is indeed possible to machine any number of teeth into the disc, doing so requires more effort on the machinist’s part and the likelihood of small
deviations in tooth width and tooth spacing greatly increases. The second consideration was based on how many passes were required by the cutting tool to create the “teeth”. It was decided the easiest method was to use a single pass technique, whereby the cutting tool cuts the tooth spacing in one cut. This is simpler than the two pass option, where each side of each tooth is cut individually, but restricts the tooth spacing cuts to the width of the cutting tool. In turn this restricts the design options, and produces “dovetail” shaped teeth, neither of which were considered problematic. The estimated accuracy of the sensor arrangement in measuring angular velocity is therefore given in Figure 4.11, which is based on Eqn 4.1 above.

![Figure 4.11: Maximum Hall effect sensor error on rear (rolling) wheel – 40 teeth per revolution](image)

### 4.2.1.3 Sensor Calibration

In automotive applications, the raw angular velocity measurements of each wheel can be difficult to directly interpret into real-world meaning. It is therefore normal practice to convert wheel angular velocity to the corresponding linear velocity of the vehicle in km/hr. This initially looks like a straightforward task based on the circumference of the wheel, but as MoTeC explains, it is not that simple.

The wheel circumference for correct calibration is usually somewhere between the measured wheel circumference using a tape measure and the distance measured by rolling the vehicle for one turn of the wheel. This must be determined experimentally as it can depend on the vehicle and tyre.

The circumference of a tyre is difficult to determine because the tyre deflects at the contact patch. The tyre is pneumatic, so air pressure and temperature can affect its diameter. It can also have a wide variety of loads placed on it, which deform its shape, and therefore its circumference. It can also wear out. Any simple calibration technique will be in error occasionally, and the calibration methodology must seek to find the middle ground.
The most widely advertised method for determining wheel speed calibration data is to measure how many times the wheel rotates over a given distance. This must be done on the vehicle with a condition of negligible slip (i.e. at low accelerations and speeds), and must be done with the vehicle loaded to its average weight and weight distribution. The tyre pressures must also be at their rated values.

It was observed in practice, however, that it is very difficult to accurately measure a specific distance using the available equipment. Instead, it was noted that a highway near Hobart contains an “odometer check” facility, which comprises of a near level stretch of road with measured markings every 1km for 4km. Using this location over the full 4km it was found that the front and rear wheel circumference was 1.830m.

It is noted, however, that these values will change when tyre loads and pressures are altered, and significant error may propagate through the results. The level of this error under normal driving conditions is difficult to establish, but can be estimated using the following argument. Under normal driving it is possible (although improbable) to completely unload one or more wheels, whereby the tyre will adopt a circular shape. In such a case it is unlikely that the wheel will be spinning with such velocity as to significantly alter its geometry. Tyre pressure variation is unlikely to significantly change the shape of the tyre. It can therefore be assumed that the maximum error in tyre circumference likely to be observed will be the difference between the unloaded tyre circumference and the loaded (calibration) tyre circumference. Moreover, when the tyre is placed under excessive load, its circumference will reduce in the same manner as above. However, due to the compression of the air inside it and the construction of the tyre, it is expected that this error will be of similar or smaller magnitude to the no load case. The expected maximum error can therefore be specified, and is show in Eqn 4.2 to Eqn 4.4 below.

\[
C_{\text{MaxError}} (m) = C_{\text{NoLoad}} - C_{\text{Calibration}} \quad E\text{qn 4.2}
\]

\[
V_{\text{MaxError}} (km/hr) = \frac{3.6 \dot{\theta} C_{\text{MaxError}}}{2\pi} \quad E\text{qn 4.3}
\]

\[
E_{\text{Velocity}}(\%) = \frac{C_{\text{MaxError}}}{C_{\text{Calibration}}} \quad E\text{qn 4.4}
\]
By substituting the appropriate values and noting the no load circumference = \( C_{\text{No Load}} = 1.843 \text{m} \) we get the maximum circumference error = \( C_{\text{Max.Error}} = 0.013 \text{m} \), linear velocity maximum error at 50km/hr = \( V_{\text{Max.Error}} (\text{km/hr}) = 0.06 \text{km/hr} \) and percentage maximum velocity error = \( E_{\text{Velocity}} (%) = 0.7\% \) approximately front and rear. It is observed, however, that this is a maximum value, and for most stable driving conditions this level of error is unlikely to be reached. Nonetheless, this error clearly overshadows the switching error of the Hall effect sensor.

### 4.2.2 Suspension Position Sensors

The suspension position of each wheel was measured using linear potentiometers installed on individual suspension linkages. Potentiometers are analogue sensors, and operate on the basis that their electrical resistance is proportional to length of the resistor, and that the output voltage is proportional to resistance. They generally consist of a moveable component that makes contact with an internal resistor and forms a circuit, with the resistance amount defined by its position. The movement of this contact is normally either in translation (linear) or rotation (angular). When a resistance element has a voltage applied to it the motion of the moveable contact results in a change in output voltage across the sensor that is linearly proportional to the contact position, as is depicted in Figure 4.12.

![Potentiometer operation](Figure 4.12: Potentiometer operation [115])

Four 100mm stroke linear potentiometers were selected for this application, and were supplied by Gefran (model PZ12A). These sensors are considered a standard size for automotive suspension measurement. They consist of anodised aluminium cylinder cases with internal moveable control rods made from stainless steel, and have a useful electrical stroke of 100 mm and a mechanical stroke of 105mm. At the end of the
moveable rod and at the bottom of the sensor there are two M5 self-aligning rod ends used for mounting, with a minimum “eye to eye” length of 228mm due to the size of the sensor.

The sensors can survive speeds of up to 10m/s and forces of less than 0.5N. The 40Ω resistor can also withstand up to 10mA (although 0.1µA is recommended) and 60V, and has an independent linearity error ± 0.1% with infinite resolution. As with the Hall sensors, these sensors also entail the use of a supply voltage wire (brown), a ground wire (blue) and the signal wire (yellow).

4.2.2.1 Sensor Considerations
The placement of the linear potentiometers to measure suspension movement is critical in obtaining useful, reliable and accurate data. Nonetheless, mounting these sensors to the suspension system is a difficult proposition, and many design compromises need to be made.

Since wheel position relative to the vehicle in the vertical plane is being measured, it is obvious that mounting the sensor directly from the body of the vehicle to the wheel in a vertical direction would provide direct results. This is, however, fraught with difficulties. Firstly, the wheel moves on its suspension much more that 100mm which would require a much larger sensor. Secondly, having the sensor mounted at the wheel would expose it to adverse operating conditions, with a strong likelihood of failure. Thirdly, the steering action of the front wheels would alter the geometry of the sensor, affecting results. Finally, physically mounting the sensor in the centreline of the wheel is impossible.

As a result of these problems potentiometers are often mounted to the suspension springs or dampers. This has the benefit of providing spring and damper position and velocity (which plays a very important role in evaluating suspension performance [116]), but with knowledge of the suspension geometry or experimental testing it can be used to estimate the wheel position – with the assumption of no free play in joints. This often reduces the displacement for measurement, places the sensors in a safer position, negates steering effects and is much easier to mount.
4.2.2.2 Sensor Installation
While the above is the case on most racecars, the suspension on the test vehicle makes mounting suspension sensors directly to the springs or dampers difficult. As a result, it was decided to mount the sensors from the vehicle body to an arbitrary point on the suspension system that moved with the wheel. Using this method it was possible to easily mount the sensors in a way that maximizes their stroke length, and places them in a reasonably safe position. Using an experimental technique it would then be possible to calibrate the sensed displacement to movement at the wheel, or movement at the spring/damper with a reasonable degree of accuracy. This is a common mounting method.

4.2.2.2.1 Front Suspension Position Sensors
The front wheel suspension is independent and of a “Macpherson Strut” type [20], and as such, comprises of a lower pivot and a spring/damper from which the wheel is suspended, as depicted in Figure 4.13. The installation of the front linear potentiometers for suspension movement measurement is shown in Figure 4.14. The lower end of the linear potentiometer is attached to the lower pivot about half way along its length, and the upper end is connected to the chassis through the wheel well.

![Figure 4.13: Front suspension](image)

This choice of placement was made based on observations of the suspension movement, and on possible mounting points. The wheel well contains a small number of nuts welded into the body of the car that provide mounting possibilities, and one of these was observed to be in a useful location. The upper rod end of the potentiometer was thus installed in this location using a custom built “double threaded” bolt, as shown in Figure
4.15. This bolt was designed to convert the \( \frac{1}{4} \)” thread of the vehicle to a M5 thread to suit the sensor, as well as to provide sufficient clearance for the sensor to operate correctly.

![Figure 4.14: Front suspension position linear potentiometer mounting](image1)

![Figure 4.15: Front suspension “double threaded” bolt for mounting](image2)

![Figure 4.16: Front suspension lower strut mounting method](image3)

The installation of the lower potentiometer mounting required some small modification to the strut, and is shown in Figure 4.16. The mounting was produced by drilling and tapping a M5 thread into the hollow strut. It is noted that this will reduce the strength of the strut, but a location was chosen whereby the effect would be negligible. A M5 bolt was installed into the hole so the thread was pointing outwards and “Loctite®” used to provide a stud to mount the sensor on. This was difficult in practice because there was no way to access the bolt head, so the threaded end of the bolt was modified to allow
Finally, it was observed that this mounting arrangement took full advantage of the length of the sensor for the full range of suspension movement. The sensor was installed with approximately 5mm extra travel with the suspension in full droop and has some additional travel to compensate for unknowns when in full rebound. It is noted, however, that even though the sensor has been installed to accommodate extremes of movement to avoid failure, its normal operating range will be limited to approximately 10% of its travel because full droop and full rebound are rarely encountered in practice.

4.2.2.2 Rear Suspension Position Sensors
The rear suspension is not independent, and is of a “Trailing Twist Axle” type [20]. This type of suspension is increasingly used in small front wheel drive vehicles, and consists of two trailing arms connected by a single transverse member, as shown in Figure 4.17.

Unfortunately, this means that a movement from one wheel can directly affect movement of the other, so it is difficult to consider each as separate entities. They are not rigidly connected, however, with the bar between them designed to be rigid in bending (locating the wheels in plan view) and torsionally compliant during roll (providing anti roll and camber gain characteristics). It could be argued that it is fair to approximate the workings of the rear suspension to an independent trailing arm arrangement, with an anti-roll bar linking each wheel. This assumption reduces the degrees of freedom of the
system, and as a result it can be assumed that the wheels move in the arc of the trailing arm. This makes it possible to relate the movement at the trailing arm as proportional to movement at the wheel. By making this assumption a very small amount of error would be introduced into the wheel position calibration, as the arc of movement of the trailing arm would vary slightly in reality.

Nonetheless, the placement of the suspension sensors must be chosen to measure only the movement of the wheel, and not incorporate any of the suspension flex into the readings. This means that the sensors must be mounted to a part of the trailing arm not likely to flex with movement of the transverse member. This is compounded by the desire to use existing features around the suspension to provide rigid mounts, and to avoid mounting options that would reduce suspension strength. Nonetheless, the final mounting points are shown in Figure 4.18, and fulfil these requirements.

The lower mount bolts directly into the trailing arm. This location was chosen to provide a vertical mount for the sensor as well as to avoid reducing the strength of the arm (being mounted at the edge of the formed metal sheet, well away from the welds). The upper mount is installed in a similar way, with an M5 thread tapped into an overhanging edge of metal sheet. Both mounts also utilise “Loctite®”, and have spacing washers installed to maximise the rod end movement potential.

Figure 4.18: Rear suspension position linear potentiometer mounting

4.2.2.3 Sensor Calibration
Suspension position sensors can be calibrated to determine three important, yet proportional, variables. When evaluating damper performance it is important to be able
to measure the position of the damper, which can be differentiated into speed. In this case it is useful to reference the calibration to the damper. When evaluating kinematic suspension performance, however, it is important to determine the load on the tyre, which for the most part is determined by the spring position. In this case it is useful to reference the calibration to the spring, which is often in the same position as the damper. When evaluating the dynamics of the vehicle, as is the aim of this investigation, it is useful to determine the position of the wheel relative to the body. These three variables are related to each other by the “Motion Ratio” of the suspension (MR = Wheel Movement/Spring Movement), which is approximately constant, and, for example, allows the force at the spring to be related to the force at the wheel.

The preferred method of racecar suspension position calibration is to remove the suspension spring and anti-roll bar and measure the position and sensor voltage relationship as the suspension is moved through its entire range. Removing the spring makes it much easier to move the suspension and collect data. In the case of the test vehicle, however, the type of suspension makes this difficult, as the spring/damper unit is integral to the suspension at the front (in the Macpherson Strut arrangement) and at the rear, the transverse member makes independent movement impossible. As such, combined with the amount of effort required, a different method was chosen for suspension calibration.

It was decided that the simplest method of calibration was to model the suspension in CAD. This would not only provide the geometry necessary to complete the calibration, but would compile all required data to undertake a very thorough suspension analysis if required. The resulting CAD models are shown above in Figure 4.13 and Figure 4.17.

By including the suspension linear potentiometer into the models it was possible to relate a movement at the wheel to movement at the sensor and at the spring/damper. By assuming that the linear potentiometers are in fact linear, data was acquired to calibrate the suspension parameters at each wheel, with the results shown in Table 4.1. It can be seen from the data that the suspension position has been zeroed so that the normal ride height when unloaded reads as zero. This was done to provide results that were easy to understand and interpret, although it is noted that the zero is arbitrary. Further, the data for one side of the car is different from the other. This arose from the fact that it was difficult to position the sensors in an identical arrangement on both sides of the car.
Finally, the motion ratios where calculated from Table 4.1 as approximately 0.99 for the front wheels and 1.25 for the rear, the ratios decrease with increasing deflection.

<table>
<thead>
<tr>
<th>FL Voltage (V)</th>
<th>0.00</th>
<th>0.01</th>
<th>0.50</th>
<th>0.86</th>
<th>1.00</th>
<th>1.15</th>
<th>1.50</th>
<th>1.73</th>
<th>2.00</th>
<th>2.13</th>
<th>2.50</th>
<th>2.89</th>
<th>3.00</th>
<th>3.53</th>
<th>4.04</th>
<th>5.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspension Pos. (mm)</td>
<td>-83.0</td>
<td>-52.4</td>
<td>-48.1</td>
<td>-20.5</td>
<td>-11.5</td>
<td>0.0</td>
<td>26.2</td>
<td>43.4</td>
<td>64.5</td>
<td>74.7</td>
<td>104.0</td>
<td>115.8</td>
<td>145.5</td>
<td>193.8</td>
<td>245.4</td>
<td>371.2</td>
</tr>
<tr>
<td>Spring/Damper Pos. (mm)</td>
<td>0.0</td>
<td>60.8</td>
<td>124.3</td>
<td>156.8</td>
<td>221.4</td>
<td>284.3</td>
<td>342.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR Voltage (V)</td>
<td>0.00</td>
<td>0.26</td>
<td>0.50</td>
<td>1.00</td>
<td>1.13</td>
<td>1.37</td>
<td>1.50</td>
<td>1.97</td>
<td>2.00</td>
<td>2.38</td>
<td>2.50</td>
<td>3.00</td>
<td>3.13</td>
<td>3.78</td>
<td>4.28</td>
<td>5.00</td>
</tr>
<tr>
<td>Suspension Pos. (mm)</td>
<td>-96.6</td>
<td>-98.7</td>
<td>-98.7</td>
<td>-27.2</td>
<td>-17.9</td>
<td>0.0</td>
<td>10.0</td>
<td>46.0</td>
<td>48.0</td>
<td>77.3</td>
<td>86.6</td>
<td>127.2</td>
<td>138.4</td>
<td>196.4</td>
<td>247.9</td>
<td>337.1</td>
</tr>
<tr>
<td>Spring/Damper Pos. (mm)</td>
<td>0.0</td>
<td>60.8</td>
<td>124.3</td>
<td>156.8</td>
<td>221.4</td>
<td>284.3</td>
<td>342.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RL Voltage (V)</td>
<td>0.00</td>
<td>0.06</td>
<td>0.50</td>
<td>1.00</td>
<td>1.37</td>
<td>1.50</td>
<td>1.54</td>
<td>2.00</td>
<td>2.00</td>
<td>2.50</td>
<td>2.61</td>
<td>3.00</td>
<td>3.19</td>
<td>4.18</td>
<td>4.86</td>
<td>5.00</td>
</tr>
<tr>
<td>Suspension Pos. (mm)</td>
<td>-83.4</td>
<td>-80.6</td>
<td>-57.8</td>
<td>-30.2</td>
<td>-9.5</td>
<td>-2.0</td>
<td>0.0</td>
<td>26.1</td>
<td>26.3</td>
<td>54.3</td>
<td>60.5</td>
<td>82.7</td>
<td>93.5</td>
<td>157.3</td>
<td>213.4</td>
<td>225.1</td>
</tr>
<tr>
<td>Spring/Damper Pos. (mm)</td>
<td>0.0</td>
<td>54.7</td>
<td>82.9</td>
<td>111.3</td>
<td>139.9</td>
<td>196.7</td>
<td>250.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR Voltage (V)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.50</td>
<td>1.00</td>
<td>1.32</td>
<td>1.50</td>
<td>1.57</td>
<td>1.95</td>
<td>2.00</td>
<td>2.50</td>
<td>2.56</td>
<td>3.00</td>
<td>3.14</td>
<td>4.13</td>
<td>4.83</td>
<td>5.00</td>
</tr>
<tr>
<td>Suspension Pos. (mm)</td>
<td>-85.6</td>
<td>-85.2</td>
<td>-59.6</td>
<td>-32.0</td>
<td>-14.0</td>
<td>-3.8</td>
<td>0.0</td>
<td>21.7</td>
<td>24.3</td>
<td>52.5</td>
<td>56.0</td>
<td>81.0</td>
<td>89.0</td>
<td>152.8</td>
<td>208.9</td>
<td>225.3</td>
</tr>
<tr>
<td>Spring/Damper Pos. (mm)</td>
<td>0.0</td>
<td>54.7</td>
<td>82.9</td>
<td>111.3</td>
<td>139.9</td>
<td>196.7</td>
<td>250.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Suspension and spring/damper position calibration data

4.2.3 Steering Angle Sensor

The steering angle sensor of the test vehicle was measured using a rotary potentiometer acting on a pulley system, as depicted in Figure 4.19. The sensor acts on the same principle as the linear potentiometers above, but is rotary in nature and has a 10 turn limit. The pulley system is used to provide increased mounting options and to alter the sensor motion ratio for improved resolution.

4.2.3.1 Sensor Considerations

The steering angle sensor is designed to measure the steering wheel angle directly, is independent of steering free play and provides a direct record of driver input. By placing the larger pulley onto the shaft the steering wheel is mounted on (i.e. on the steering column) and connecting it to the smaller pulley of the potentiometer via a toothed belt, accurate measurements of steering wheel angle can be made. The smaller pulley must
have the facility to adjust the tension of the belt to ensure it retains its position, and must also move with the steering column.

4.2.3.2 Sensor Installation
A photograph of the steering angle potentiometer is provided below in Figure 4.20. A large diameter hole was machined into the large pulley and is held in place in the steering column using a number of grub screws. The installation of this pulley required the removal of the steering linkage. Further, the small pulley is mounted to a steering column flange so it moves with the steering column. It is held in place using a custom built bracket, which incorporates a sliding mechanism to allow pulley belt tensioning.

4.2.3.3 Sensor Calibration
The steering wheel angle sensor was calibrated into degrees from straight. That is, when the vehicle is travelling in a straight line the steering angle should read 0°. When the wheel is turned clockwise a positive angle is recorded, with a full turn equal to 360°, and when turned anti-clockwise a negative angle is measured.

Determining the 0° position of the steering wheel was the most difficult part of the calibration procedure, as it required data on the alignment of the wheels. To this end calibration was performed on the “Tasman Bridge”, which is located in Hobart, Tasmania, and includes a 600m straight and inclined section of road. By logging the steering wheel angle over this section of road it was possible to take the average value as
the 0° calibration position. With this data obtained, it was a simple task of marking the steering position at 0° and turning the wheel 360° in either direction to find these calibration points. By assuming a linear relationship, it was possible to complete the calibration for the entire steering range, the results of which are given in Table 4.2.

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>0.000</th>
<th>1.230</th>
<th>2.461</th>
<th>3.708</th>
<th>4.955</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steering Wheel Angle (°)</td>
<td>-720</td>
<td>-360</td>
<td>0</td>
<td>360</td>
<td>720</td>
</tr>
</tbody>
</table>

*Table 4.2: Steering wheel angle calibration data*

By assuming no free play in the steering system, it is possible to calibrate the “steering wheel angle” data into “steered angle at each wheel”. This proves more information from a vehicle dynamics point of view, and as such is an important relationship to determine. The relationship is not linear however, so it becomes necessary to either accurately model the system, which is difficult, or to obtain experimental data. The latter method was chosen, as obtaining the experimental data is a reasonably straightforward task requiring information to be gathered on how much the tyres are turned (relative to the ground) for any input measurement of the steering sensor voltage.

This calibration procedure required the angle that the tyre was turned to be accurately measured from a reference point. To do this a flat board was fixed to the face of the wheel rim, and a reference edge was fixed to the ground. With the car sitting normally on a flat surface it was possible to measure the angle between the board and the reference (by measuring two horizontal distances between them). This produced a calibration data for both of the front wheels, which can be seen in Table 4.3.

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>0.000</th>
<th>1.230</th>
<th>2.461</th>
<th>3.708</th>
<th>4.955</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steering Wheel Angle (°)</td>
<td>-720</td>
<td>-360</td>
<td>0</td>
<td>360</td>
<td>720</td>
</tr>
</tbody>
</table>

Further still, it is useful to sometimes consider a vehicle using the “bicycle” model. This approximates the four-wheel system into a two-wheel system, and is a standard tool in suspension investigation as discussed by Dixon [20]. To produce this bicycle model it is necessary to use the steered angles of the front wheels to determine the imaginary front wheel angle of the bicycle, which is located through the centreline on the vehicle. This is done by considering the “Instantaneous Turn Centre (ITC)” [16] of each of the steered wheels with no slip. This is defined by looking at the vehicle in plan view and projecting perpendicular lines from each of the tyres. The point at which the line from the front tyre intersects the line from the tyre behind it is the ITC of that side of the car. By calculating the ITC for both sides of the car, the steered angle for the bicycle model can be determined from the average ITC. The resulting calibration table is shown in Table 4.4.
4. Chassis Sensors and Data Logger Installation

### Table 4.3: Individual steered wheel angle calibration data

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>0.41</th>
<th>0.72</th>
<th>1.22</th>
<th>1.68</th>
<th>1.96</th>
<th>2.26</th>
<th>2.45</th>
<th>2.45</th>
<th>2.46</th>
<th>2.63</th>
<th>2.96</th>
<th>3.42</th>
<th>3.69</th>
<th>3.99</th>
<th>4.42</th>
<th>4.52</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left Steered Angle (°)</td>
<td>-37</td>
<td>-31</td>
<td>-21</td>
<td>-13</td>
<td>-8.1</td>
<td>-3.3</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0</td>
<td>2.7</td>
<td>7.8</td>
<td>19</td>
<td>31</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Right Steered Angle (°)</td>
<td>-33</td>
<td>-28</td>
<td>-20</td>
<td>-12</td>
<td>-8</td>
<td>-3.2</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0</td>
<td>2.8</td>
<td>8.1</td>
<td>16</td>
<td>27</td>
<td>36</td>
<td>38</td>
</tr>
</tbody>
</table>

### Table 4.4: Bicycle model steered wheel angle calibration data

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>0.41</th>
<th>0.72</th>
<th>1.22</th>
<th>1.68</th>
<th>1.96</th>
<th>2.26</th>
<th>2.45</th>
<th>2.45</th>
<th>2.46</th>
<th>2.63</th>
<th>2.96</th>
<th>3.42</th>
<th>3.69</th>
<th>3.99</th>
<th>4.42</th>
<th>4.52</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle Steered Angle (°)</td>
<td>-35</td>
<td>-29</td>
<td>-20</td>
<td>-13</td>
<td>-8.1</td>
<td>-3.3</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0</td>
<td>2.8</td>
<td>8</td>
<td>16</td>
<td>25</td>
<td>33</td>
<td>35</td>
</tr>
</tbody>
</table>

#### 4.2.4 Attitude and Heading Reference Sensor

Vehicle acceleration and yaw rate have been identified in previous work as important parameters to measure to gain information on chassis movement. In order to measure these an “Attitude and Heading Reference Sensor” (AHRS) supplied by “Crossbow Technology Inc.” was installed on the test vehicle, and is designed principally for the avionics industry.

Figure 4.21: Attitude and heading reference sensor supplied by Crossbow (DMU model 400CA-200)

The DMU AHRS400CA-200 sensor, which is pictured in Figure 4.21, measures acceleration in three directions (x, y and z), yaw, pitch and roll rates and yaw, pitch and roll angles. If the sensor is placed at the centre of gravity of the vehicle, these parameters are sufficient to easily establish the absolute linear and polar acceleration of the vehicle chassis, and also to determine its orientation in three-dimensional space. Due to the compact nature of the sensor, it can be assumed that each of these readings can be referenced from the same position within the test vehicle.
4.2.4.1 Sensor Considerations
The AHRS is a clustered measurement system, designed to measure nine parameters, including:

- dynamic acceleration on three axes (± 0.012G bias, 1% full scale non-linearity);
- stabilised roll, pitch and yaw angular rates (± 1.0°/sec bias, 0.3% non-linearity);
- roll, pitch angles (± 1° static and ± 2.5° dynamic accuracy); and
- yaw angle (± 2° static and ± 4° dynamic accuracy).

The AHRS measures these variables by using a combination of micro-machined silicon three-axis accelerometers, three-axis rotational rate sensors and three-axis magnetometers and can output the results in an analogue form through nine channels or through a single RS232 digital output. Figure 4.22 depicts the block diagram of the DMU AHRS cluster measuring system and reference axes.

Figure 4.22: AHRS block diagram [117]

4.2.4.1.1 Tri-Axial Accelerometer Operation
Accelerometers can be thought of as measuring the relative displacement of a spring mass system to measure acceleration. The three micro-machined silicon micro electrical mechanical system (MEMS) accelerometers used in this sensor are no different, using differential capacitance to sense acceleration. These accelerometers operate by sensing a change in electrical capacitance with respect to applied acceleration through the use of a distorting diaphragm sandwiched between two plates. The two plates form the capacitor unit, and detect changes as they are separated by the movement of the diaphragm due to acceleration in a single plane. Since capacitance is inversely proportional to the distance between the two plates, and the diaphragm displacement is a non-linear function, and resulting capacitance differentials are not a linear function with applied acceleration. Therefore, the AHRS performs signal conditioning to the results to achieve linearisation.
4.2.4.1.2 **Tri-Axial Accelerometer Gyroscope Sensor Operation**

The three angular rate gyroscopic sensors used in the AHRS are made up of a number of vibrating ceramic plates that use a silicon MEMS structure to measure the Coriolis force \[ F = 2\omega \cdot V \cdot \sin(\phi) \], where \( \omega \) = angular velocity of the axis, \( V \) = relative velocity and \( \phi \) = angle between vectors \( \omega \) and \( \phi \) induced by the dynamic movement of the test apparatus. This data can be used to gain the rotation rate around the given axis. The advantage of this approach is that the output angular rate is independent of the acceleration output. One problem that arises is that a change in the direction around one axis of a driving transducer induces a vibration in the detection transducer on another axis. To overcome this problem an oscillator circuit is used to control the vibration. However, during testing in previous research, it was observed that when supply power became sufficiently low (at about 11V) this oscillator circuit began to fail and as such the angular rates were the first parameters to lose accuracy. To overcome this problem the AHRS should be installed so input voltages stay continually above 11V.

4.2.4.1.3 **Tri-Axial Accelerometer Magnetometer Operation**

The three magnetometers within the AHRS are constructed as miniature fluxgate sensors and are used to provide the AHRS heading angles with respect to the Earth’s magnetic field. These headings are only used for reference, however, with the angular rate integrals and gravity angle used to stabilise the results, as depicted in Figure 4.23. Unfortunately, this signal conditioning introduces result errors into the angle measurement due to the nature of measurement, and can be further exacerbated by the erroneous angular rate measurements at low supply voltages.

![Figure 4.23: Angle value signal conditioning block diagram [117]](image)

4.2.4.2 **Sensor Installation**

The ideal place to measure vehicle acceleration and dynamic angles of the chassis is at the centre of gravity of the vehicle, and away from magnetic/ferrous objects. This placement means that the sensor can directly associate the measurements to the entire body and requires no interpretation of the results. It also means that the sensor should be mounted away from the chassis to avoid adverse affects. This is rarely possible in...
automotive applications due to vehicular design, and other positions must be found. Unless major structural modifications are to be carried out, this obviously requires the sensor to be placed away from the centre of gravity of the vehicle. The result is that the induced misalignment of the measured coordinate system with respect to the established coordinate system of the test vehicle can lead to significant error. If the vehicle chassis is considered rigid, the measured angles and angular rates from the sensor can be assumed identical regardless of placement, although the linear accelerations cannot. This effect is highlighted in Figure 4.24 below.

![Figure 4.24: Some effects of sensor placement on acceleration values](image)

Nonetheless, using the AHRS linear and polar acceleration outputs and referencing them mathematically to the centre of gravity of the vehicle can negate this effect. By using this data in conjunction with the vehicle geometry and the assumption of a rigid body, it is possible to place the sensor anywhere on the vehicle and then calculate the linear acceleration at any other point because the acceleration of the body is fully defined.

Although this appears to be an excellent solution, it has one severe drawback that stems from financial concerns. The AHRS is an expensive sensor, and is used here simply because it was freely available within the University. Such a sensor is unlikely to be used in automotive applications for some time, and thus using the advanced features it contains is a luxury that cannot be warranted in this research. This means that, although it is possible to reference the data to the centre of gravity, the sensor should be utilised in the same manner as it would in commercial implementation.
This means that the AHRS should be mounted as close as possible to the centre of gravity of the vehicle, and any misalignment problems that arise should be factored into the sensor error. To this end, the AHRS was installed into the centre console of the car between the two front seats, as pictured in Figure 4.25. The centre of gravity of the normally loaded vehicle was determined along its longitudinal axis (front to back) by measuring the tyre loads, and its position along the lateral axis (left to right) was assumed symmetric. The position of the centre of gravity along the vertical axis was not calculated, as this requires significantly more equipment.

The sensor unit was installed into the centre console using four brass screws through the mounting holes in its base. The screws were installed rigidly into the centre console and in effect supply four studs to mount the sensor on. By using a number of spacer washers it was possible to finely adjust the orientation of the sensor for calibration.

It is observed that the voltage outputs of the AHRS are in the 0-5V range for acceleration (assuming 5G max) and the ±5V range for angles and angle rates. Since the data acquisition devices to be used in the test vehicle only measure positive voltages it is necessary to condition the AHRS signal. Fortunately, previous research using the AHRS also encountered this problem and developed a signal-conditioning device specifically for this purpose. This simple device ensures positive voltages by adding 5V to the AHRS angle and angle rate signals to overcome this problem, and also amplifies acceleration by a factor of two to provide similar range. It does, however, produce two new problems. The first is the noise the device adds to the signal, the second is the fact that the new signal is now in the 0-10V range, which some data acquisition devices cannot handle. This, combined with the fact that the AHRS voltage signal inherently
contains error (as a result of the digital to analogue conversion within the sensor), means that the voltage outputs of the sensor contain significant error in anything other than static conditions. Nonetheless, all of this error can be overcome by using the AHRS serial communications that attract none of these problems. It is noted, however, that many data acquisition devices cannot communicate through serial communications, and that even though the ADL contains this functionality, it is not compatible with the AHRS due to firmware limitations.

### 4.2.4.3 Sensor Calibration

The orientation of the sensor axes installed in the vehicle are the same as the SAE vehicle defined axes used in automotive applications, as discussed previously. The AHRS calculates its values based on these directions, and is pre-calibrated to an extent. This means it was only necessary to calibrate the sensor for a number of ‘zeros’ and ensure level placement.

A flat and level surface to place the sensor on was produced for the calibration procedure. The sensor was calibrated for linear acceleration by progressively orientating each of the axes into a vertical direction, so they would each see 1G (9.81m/s²), while the others 0G (0m/s²). By assuming a linear relationship, as given in the specifications, these values could be used to determine the calibration for the full range.

<table>
<thead>
<tr>
<th>Linear Acceleration (G)</th>
<th>-10</th>
<th>-5</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Axis Voltage (V)</td>
<td>8.644</td>
<td>6.959</td>
<td>5.611</td>
<td>5.274</td>
<td>4.940</td>
<td>3.604</td>
<td>1.934</td>
</tr>
<tr>
<td>Y Axis Voltage (V)</td>
<td>8.654</td>
<td>6.969</td>
<td>5.621</td>
<td>5.284</td>
<td>4.953</td>
<td>3.629</td>
<td>1.974</td>
</tr>
<tr>
<td>Z Axis Voltage (V)</td>
<td>8.562</td>
<td>6.912</td>
<td>5.592</td>
<td>5.262</td>
<td>4.912</td>
<td>3.557</td>
<td>1.852</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Angle (°)</th>
<th>-180</th>
<th>-90</th>
<th>0</th>
<th>90</th>
<th>180</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roll Voltage (V)</td>
<td>1.065</td>
<td>3.262</td>
<td>5.165</td>
<td>7.252</td>
<td>9.278</td>
</tr>
<tr>
<td>Pitch Voltage (V)</td>
<td>1.130</td>
<td>5.077</td>
<td>8.687</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yaw Voltage (V)</td>
<td>1.037</td>
<td>3.235</td>
<td>5.432</td>
<td>7.500</td>
<td>9.223</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Angular Rate (°/s)</th>
<th>200.2</th>
<th>0</th>
<th>200.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roll, Pitch &amp; Yaw (V)</td>
<td>1.054</td>
<td>5.150</td>
<td>9.246</td>
</tr>
</tbody>
</table>

*Table 4.5: Attitude and heading reference sensor (AHRS) calibration data after signal conditioning*

Calibrating the angles from the AHRS required a similar procedure, whereby values of 0 and 90° are easy to determine. On the other hand, it is difficult to calibrate anything other than the zero values for the AHRS polar accelerations because it would require spinning the sensor at known accelerations (with the cable attached). Instead it was observed in the specifications that 1V = 1°/s. Using this information it was possible to
calibrate all of the polar accelerations. The final calibration table for the voltage outputs is given in Table 4.5.

4.3 Advanced Dash Logger

The Advanced Dash Logger is supplied by MoTeC, which is described as a “fully featured, self contained, programmable data logger, device controller and display unit”. It uses a 32-bit microprocessor and can log analog, digital, RS232 and CAN bus channels, as well as provide a display of any measured or calculated parameters to the driver and control a number of external devices. Of particular interest here, it measures analog channels at 12-bit resolution and digital channels at a maximum rate of 3200Hz (with 1MHz counting), and can filter and log these parameters at up to 1000Hz into flash memory.

It can be upgraded in functionality with the purchase of “upgrade” codes. In this case the ADL was equipped with the “30 I/O”, “1 Mbyte Memory” and “Telemetry” upgrades, which increased the functionality of the base ADL to incorporate:

- 10x Analog voltage inputs (AV1 to AV10);
- 4x Analog temperature Inputs (AT1 to AT4);
- 4x Switch inputs (Sw1 to Sw4);
- 2x Digital inputs (Dig1 & Dig2);
- 4x Wheel speed inputs (Spd1 to Spd4);
- 2x 0 to 1 volt inputs using the LA1 & LA2 pins (without Lambda functionality);
- 4x Auxiliary digital outputs (Aux1 to Aux4);
- 6x CAN bus addresses (can convert to RS232 using Real Time Clock modules);
- 1x RS232 communication channel;
- Radio and GSM telemetry transmission (via RS232); and
- 1 Megabyte flash logging memory.

Finally, all the wiring to the ADL is via a 79-pin autosport connector, which is located at the back of the 385gram unit.
4.4 Advanced Dash Logger Installation

Even though the ADL installation was essentially to be a copy of the racecar setup, this proved not as simple as expected. This was due to three reasons; the first being that the ADL did not work in the racecar as a result of some unidentified wiring problems, making initial investigation difficult. The second reason, which proved a much larger hurdle, was that there was no documentation of the installation – and all of the wiring was concealed in large taped electrical looms. The third was that the racecar ADL installation with regard to the AHRS sensor was partially flawed, which in previous research had resulted in the need to be selective of acquired data and mindful of the extent of introduced noise. As such, the ADL installation had to be started essentially from scratch, relying only on the MoTeC documentation [114] and significantly increasing the project complexity.

Excluding the sensor installation, described above, the ADL installation could be split into a number of tasks. These included installing the user interface of the ADL (which included a computer interface, as well as control buttons, control switches, indicator lights and an indicator buzzers), the radio modem, the radio antenna, the CAN bus and the CAN devices, as well as correcting the AHRS problems and designing the general...
4. Chassis Sensors and Data Logger Installation

wiring arrangement and routing. The final installation of which can be seen in Figure 4.26.

4.4.1 Interface, Switches, Buttons, Indicator Lights & Buzzer

The ADL, like any device, requires an interface of some description. To this end it has a computer interface (for configuration and programming), inputs for a number of buttons and switches (to control operation modes) and outputs for additional driver information. These interfaces were installed into the test vehicle in a neat and convenient manner in two ways. The first consisted of mounting an aluminum plate to the dash that was fitted with the computer interface socket, four switches and three push buttons. The second involved fitting a number of coloured LEDs and a buzzer into the air vent the ADL was mounted to. These installations are shown in Figure 4.28 and Figure 4.27 respectively, and are relatively self-explanatory.

![Figure 4.27: ADL output LEDs](image)

There are a few points that are worthy of note, however, and are listed below.
The computer interface is via the CAN bus and requires a special CAN to parallel cable to connect to a PC;

- The LEDs above the switches simply indicate if the switch is on or off, and are powered directly from the battery;
- The push buttons are capable of more functionality than specified by using them in combination and/or holding them down for a period of time;
- The LEDs below the ADL turn on in complete rows, and represent three separate and programmable outputs from the ADL (e.g. shift lights);
- The buzzer (which is not pictured) is fitted behind the ADL and is controlled in an identical manner to the three LED outputs; and
- The LED and buzzer output intensities and flashing frequencies can be programmed to vary to indicate separate functions.

### 4.4.2 Radio Modem

The radio modem installed on the racecar was transferred to the test vehicle, and allows data to be transmitted from the ADL to a PC within a maximum 30km radius. While this has very little use in this investigation, it was installed to further increase the functionality of the vehicle. In this way, third parties can monitor any measured parameters on an external PC as the vehicle drives on the road, which may be a useful tool for traffic investigation. Further, it would not be a difficult task to alter the direction of the signal, whereby data can be sent from the roadside to the test vehicle for projects such as speed zone detection and direct traffic updates to GPS devices.

![Radio Modem Installation Image](image)

*Figure 4.29: Radio modem installation*
The device consists of a transmitter and aerial fitted to the test vehicle, as shown in Figure 4.29 and Figure 4.30, which the ADL streams serial data to via RS232 (19200 baud, 8 data bits, no parity, 1 stop bit). The signal is then sent via 900MHz radio transmission to the receiver, which converts it back to serial and to the roadside PC. MoTeC provide their “Telemetry Monitor” software to display the result on the PC, although the data throughput rate can be considered too slow for effective chassis data logging.

Finally, the radio modem functionally requires the ADL “Telemetry” upgrade, which is already installed. This upgrade enables the RS232 signal from the ADL, which can be programmed to transmit any measured or calculated variables at up to 115200 baud. It is possible, therefore, to incorporate ADL data transmission to any RS232 device. Further, by using the ADL’s serial inputs, it is possible to configure two-way RS232 communications to provide more inputs and allow additional control of the ADL display and outputs.

### 4.4.3 CAN Bus and Additional Devices

The Controller Area Network (CAN) Bus of the ADL provides a high-speed network communications system that allows multiple devices to be connected and communicate with each other at up to 1Mbit. The bus consists of two wires (Hi and Lo), connected at each end by two 100Ω resistors, that form the CAN “backbone” as depicted in Figure 4.31. A number of devices can then be “hung” off the backbone by simply connecting to the Hi and Lo wires and ensuring that the maximum wire length and wire twist requirements are meet. Each device can then be given an address, and whenever two devices wish to communicate with each other they simply specify this number. In this way, each of the devices observe all of the data that is passed around the network, but only act if the data is specifically addressed to them.
The ADL incorporates CAN bus functionally, as well as the means through which devices can be addressed on the network. It also contains a number of protocols to allow the specific functions of devices to be integrated into the ADL. In this way, the ADL controls the CAN bus and allows a number of devices to work together. In the case of this installation, the CAN devices include the ADL itself, the CAN user interface plug, the MoTeC ECU (which will be discussed later, and is used for engine control) and two real time clocks – all of which are supplied by MoTeC and depicted in Figure 4.32.

The CAN interface plug can be used to configure the ADL (as described above), and also to configure the MoTeC ECU. This allows the ADL and ECU to communicate directly, enabling the ADL to incorporate engine data for logging and display (although this is limited to an extent by MoTeC software limitations). Lastly, in addition to precise data and time, the real time clocks provide “Async Expander” functionality. This basically provides one additional serial input to the ADL per clock, and allows a range of devices such as GPS receivers to be added very simply. It would also allow, for instance, additional computer serial outputs to be connected indirectly to the ADL.

### 4.4.4 Attitude and Heading Reference Sensor Rectification

The AHRS was installed into the test vehicle in an identical manner to the racecar installation, as discussed previously. Unfortunately, the serial output of the sensor was not compatible with the serial inputs of the ADL or CAN bus due to MoTeC software
restrictions, which necessitated the use of the sensor’s nine analogue voltage outputs. These voltages were conditioned using the amplifier unit designed for the racecar.

Once the unit was installed, however, its limitations became evident. Instantly noticeable was that a periodic noise (of frequency 22kHz and peak to peak amplitude of 0.5V) was produced on the AHRS outputs, and further increased in the amplifier unit. Installing 5.6nF polyester capacitors into the amplifier unit, which effectively introduced a 1.2kHz low pass filter to remove the noise, rectified this issue.

More significantly, however, it was found that the amplifier unit had been manufactured on an incorrect premise. This premise was based on MoTeC’s assertion that all of its analogue inputs measured values in the 0-15V range. It was found, however, that when the sensor applied more than 8.5V to any of the “analogue temperature” inputs (which contain a 1000Ω resistor between the input and the ADL 5V supply to allow the use of two wire variable resistance sensors) the ADL would malfunction and register incorrect readings for almost all analogue inputs. This is explained as follows:

The problem was due to the system having a common sensor voltage reference signal, which was shared among a number of sensors, combined with the fact that some of the sensor signals were actually (erroneously) influencing the sensor reference voltage. Unfortunately there are no circuit diagrams for the MoTeC ADL unit, however a number of test-measurements have revealed that the MoTeC 5V reference has a high “Thevenin resistance”, and is severely affected by reverse current. The interference path is from the angle/angle rate custom amplifier outputs, through the MoTeC “Temperature Inputs”, to the MoTeC 5V reference via the temperature input’s 1000Ω pull up resistor. As the voltage on the temperature input passes 8.5V there is a severe shift in the MoTeC 5V reference, which immediately causes radical change in any sensor signal which is derived from this 5V reference.

The solution to this was straightforward and involved moving all of the AHRS inputs to the ADL “analogue voltage” input pins, which do not contain the pull up resistors. Unfortunately, this produced another problem in that the suspension position voltages had to be moved to the analogue temperature inputs when the AHRS is in use, which are not capable of 1000Hz logging rates.

4.4.5 Wiring Loom

In contrast to the wiring loom on the racecar, the loom for the test vehicle was fully designed prior to construction and installation. The design schematic is shown in Figure 4.33. This ensured that the wiring was sufficiently simple and easy to understand by
minimising the occurrence of “add ons”, which are commonly installed in an *ad hoc* manner. This was taken a step further by trying to predict future uses of the ADL and including this into the loom construction, resulting in the inclusion of all necessary wires for the “50 I/O” upgrade (not shown in diagram).

Figure 4.33: ADL wiring diagram

(ADL pin function lists and explanations provided in [114])
The resulting loom is depicted in Figure 4.34 and, apart from some initial issues with regard to determining Rx and Tx serial wiring, operated as expected upon installation.

4.5 Advanced Dash Logger Configuration

The ADL configuration was done through the MoTeC “Dash Manager” software (shown in Figure 4.35), and was a simple process involving assigning input/output pins to the correct functions, calibrating sensors, configuring the display and data logging and writing simple programs for the output channels. This process is clearly documented in the ADL help menu and associated documentation and, since the entire installation was not unusual in any way, will not be discussed here.
4.6 Interpreter Software

The ADL contains no functionality to display logged data. Instead the logged data must be downloaded onto a PC via the CAN interface cable. Further, the data was not in a common file format and was viewed using MoTeC’s freely available “Interpreter” software, of which a screen shot is provided in Figure 4.36. The software contains a large number of data analysis tools for motor racing, most of which are unfortunately only available with the purchase of the “Pro Logging” upgrade.

In the course of this investigation, however, this software was only used to perform one function – that being to export the logged data to comma separated variable (*.csv) text file. The logged data could be exported to Microsoft “Excel” for further formatting and analysis, which will be discussed later.

![Figure 4.36: MoTeC “Interpreter” software (logged data chart shown)](image)

4.7 Remarks

This chapter broadly covers three important areas with regards to the installation of data logging hardware. These are: the installation and configuration of chassis sensors; the installation and wiring of the data logging unit; and the installation of ancillary systems.

In particular, the installed sensors include: wheel speed hall effect sensors and suspension positions potentiometer at each wheel; a steering wheel angle potentiometer; and the “attitude and heading reference sensor” to measure acceleration and angular rates near the center of mass of the vehicle. The operation, estimates of possible modes of
error and calibration of these sensors are discussed. The “MoTeC Advanced Dash Logger” data logger is then presented with reference made to the installation documentation, followed by details of the data logging ancillary systems. These systems include: user inputs; ADL outputs; CAN bus; ADL CAN interface; real-time clocks for addition RS232 capability; radio modem; and AHRS signal rectifier. The wiring loom that was designed and built for all of these sensors and systems is shown, followed by brief presentations of the ADL configuration method and software.

Of final note, however, is that this installation was only intended as the first stage in a more comprehensive data acquisition and control system to be installed later within this investigation. In particular, the readily available equipment meant that ANN investigations could begin at an early stage, and that the sensors could be installed and tested for accuracy before being utilised in later systems.
Modern vehicle stability controllers are limited in application because there is no simple and inexpensive way to identify the type of road surface the vehicle is travelling on all of the time. A common approach is to determine the road coefficient of friction while the vehicle performs severe manoeuvres by comparing wheel slip to the vehicle attitude and acceleration. This approach, however, delays efficient control and does not allow the controller to anticipate the road ahead, thus limiting its performance on a variety of road surfaces and in a range of driving conditions.

The artificial neural network models presented here attempt to solve this problem by measuring wheel speed and suspension travel on a single wheel, and monitoring the effects of vibration on these parameters and their derivatives. Vibration characteristics are used as inputs to the ANN model, which then characterises the road surface.
5. Pavement Feature Recognition during Stable Driving Conditions

5.1 Surface Prediction using ANN

When considering modern automotive stability controllers, such as antilock brakes, traction control and vehicle dynamics control, pavement characteristics have a large impact on operation and performance [8, 29]. The effect of pavement coefficient of friction on tyre grip has been discussed previously, and it is this grip that provides a vehicle with any level of stability. When surface characteristics change so does the grip, and the controller must firstly be able to account for these changes and secondly be able to predict the surface features that lie immediately ahead for efficient operation.

5.1.1 Current System Improvements

Gaining this information using current techniques is very difficult because surface friction is only gauged by comparing tyre slip to vehicle acceleration and attitude during severe manoeuvres [14, 37, 62]. Surface coefficient of friction (adhesion) must be able to be identified during stable conditions to improve systems in this respect, but this can be difficult, as stated below.

With the sensor set typically used by vehicle stability enhancement systems, it is not possible to determine the surface coefficient of adhesion as long as the vehicle remains within the linear range of operation. In this range the tyre lateral forces depend mainly on tyre elastic properties (cornering stiffness) and not on the properties of the surface, thus the vehicle response to a given steering input remains nearly independent of the surface coefficient of adhesion. This makes it impossible to determine surface properties from measured vehicle response in the linear range.

A. Hac and M. Simpson [22]

Using conventional methods of comparing acceleration to slip this statement is true. Nonetheless, stability controllers are generally only required during severe manoeuvres (i.e. outside the linear range of the tyre), meaning this technique of estimating grip from vehicle response covers most of the controller operation and generally offers adequate performance.

Even so, it should be noted that controllers utilizing this type of surface evaluation inherit a number of intrinsic limitations. Potentially of greatest significance is the initial delay of efficient control, which is a consequence of the control strategy employed. For example, the road condition identifier presented by G. Mauer et al for ABS [69] initially assumes a dry road condition. If a wheel suddenly becomes locked, it will run through a series of tests based on brake pressure and tyre slip to categorise the road condition as icy, then wet and finally as dry. Surface identification can only be made in unstable
conditions, and the controller will only come into operation during these critical or unstable situations. This means that the required road surface condition data is not available to the controller initially; instead it must guess the initial condition and monitor the responses before efficient control can be accomplished. This produces a delay in the response of the controller in the first stages of operation. A delay of this sort is a significant problem because small controller errors during the early stages of a critical manoeuvre can lead to large detours from the vehicle’s optimum trajectory. A method of identifying the road surface before the vehicle’s stability system comes into operation (i.e. during stable driving conditions) would solve this problem.

Another difficulty that arises through the traditional method of surface estimation is that it is very problematic to predict any sort of future surface variation during the time of controller operation. This can be valuable data and would allow the controller to reduce unstable operation during transient situations, as the following example illustrates.

Consider a traction controller in operation on a sealed road that is intermittently wet and dry. Imagine that a traction controller with no capacity to predict future conditions is implemented on a dry section of the road and the surface suddenly becomes wet. For a short period of time the tyre will slip excessively under acceleration on the wet road until the controller reduces torque enough to bring the wheel under control again. This is obviously undesirable because, during this short period of instability, the vehicle is much more susceptible to external influences which may force it to go out of control, such as a violent steering input from the driver. This situation could be avoided if the controller had a capacity to identify the pavement condition immediately ahead of the tyre because transient conditions could be identified before they are encountered. In this example, this extra road condition data could be used to marginally reduce the acceleration of the vehicle in the dry (i.e. with less tyre slip) in anticipation of the impending wet pavement. When the wet pavement is reached, the period of instability will be shorter than in the above case because the tyre will have had less slip and less torque applied to it in the dry, meaning the controller can adapt more rapidly to the wet condition.

If a sensor were installed on the vehicle ahead of the tyre to produce this data it would have obvious application. Z. Fan et al [23] discusses such a sensor, which measures the coefficient of friction of the road using electric resistance strain gauges with reference to work performed by T. Muro [118]. However, it is considered that this type of sensor
would be expensive and difficult to implement and, therefore, its use in commercially available vehicles is prohibitive. Another option would be to produce historical models from the surface data gathered during manoeuvres involving unstable driving conditions. However, these periods are normally very short and not sufficient to produce reliable statistical models.

Another option would be to evaluate the road surface condition while the vehicle is driving normally (i.e. in stable conditions) and use this to construct a surface history, from which the controller will be able to make statistical judgments on future conditions. While it is observed that this method will not produce the “instantaneous” performance gained from the sensor option discussed above, it will, however, allow the controller to predict the likelihood of an impending road surface variation.

### 5.1.2 The Tyre as a Sensor

A large amount of work has been done in the past to estimate the friction between the tyre and the road, either based on the effects of the friction process itself, or on the parameters affecting the frictional processes [59]. This includes using the tyre as a road surface sensor. The tyre is the only part of vehicle in contact with the road, and it is a natural conclusion that by monitoring various aspects of the tyre’s behaviour an indication of pavement characteristics can be gained. Previous work, which has been discussed above, has shown that this can be done in a number of ways, and for a number of different applications.

T. Shiotsuka et al [75] has shown that the shape of the Power Spectral Density (PSD) graphs of suspension acceleration can be used to identify and classify different road surfaces based on their roughness. T. Umeno et al [62] conducted research based on the frequency characteristics of the wheel speed vibration, incorporating PSD, to identify the slope of the linear region of the tyre longitudinal force vs. slip graph. W. Pasterkamp et al [59] used measured tyre force and torque to obtain coefficient of friction estimates, based on the “Magic Formula” tyre model of L. Palkovic et al [119].

The tyre interacts with the road on many levels of complexity, and by monitoring these it is possible to gain information on a number of pavement features. This, however, is made quite difficult by the structure of the tyre itself. While they may not look it, tyres are extremely complicated components of a vehicle, and are extremely difficult to model
accurately [26, 120, 121]. Using these models to gain data on the type of road surface is simply another complexity to be added to the overall process. Combined, this represents a formidable problem to be tackled using conventional techniques.

Conventional models rely on the programmer being able to understand the processes within the system and how they relate to the parameters of importance. For the problem of using the tyre to predict the pavement features, this means that the programmer must, firstly, concentrate on one feature at a time and, secondly, understand how the tyre reacts to this feature generally. This is a complicated process by itself, but when the number of parameters that may have a bearing on characterising a pavement surface are included this turns into a significant investigation. This is compounded by the complicated interplay of parameters in dynamic situations (which are considered normal conditions for a tyre), and as such little progress has been made in this field to date. Further, if a model was created using this approach, it would undoubtedly be very large and require a considerable amount of processing. This is undesirable because it will increase the development costs and debugging time necessary to ensure reliable operation, as R. Bannatyne [28] suggests.

5.1.3 Potential of ANN
Artificial Neural Networks, as discussed previously, can offer a number of features that traditional models lack. Most important here is their ability to be trained through experimentation, their ability to incorporate a large number of inputs and their use for input parameter importance analysis.

By using ANNs to determine surface features it is possible to use their benefits to great advantage. Because ANN models can be trained through experimentation, the programmer doesn’t need to understand the processes involved in the tyre’s operation to construct a model. Because ANN models can take an excessive number of parameters as inputs, the programmer isn’t limited to simplified tests where only a few parameters are taken into account. Further, because ANN models can be readily used for input parameter importance analysis, the programmer can make a model with many inputs and deduce which parameters have a large effect for future focus.

Using this method it is theoretically possible to construct an ANN model using as input every conceivable parameter to do with the tyre, and as output every pavement feature of
potential benefit, and train it to model the tyre over every imaginable condition. This model, if properly trained, could be used to predict the pavement features accurately and in all conditions. Further, to save computational time, and possibly improve model accuracy, an input importance analysis could be carried out to remove parameters that were seen to have little or no influence. This would have the benefit of highlighting which are the most important parameters for surface characterisation if the programmer had a desire to construct a conventional model of the process.

Of course, such an ANN model is highly impractical. At present ANN models are severely limited by computational power during the training process and architectures are still only a pale imitation of the neural processes of the brain from which they are based, meaning they have a number of large restrictions in practice. The ANN mentioned above would be too large to train using current equipment and by sheer complexity would probably be highly erroneous. By simplifying it to a manageable level, however, the same methodology could be used to take as input, a large number of relevant parameters to gain an estimation of major surface features in all required conditions, within acceptable error bounds. Parameter importance analysis could then be used in the same way to reduce model size and as an aid in conventional model programming.

5.2 Choice of ANN Model

The design of ANN models are based on the type and how many input parameters are used, how the outputs signals should be utilised, and the desired internal structure of the ANN. Previous work has shown that resonance strength in wheel speed vibration can provide information on tyre-road friction [62] and that resonance of suspension acceleration can be used to estimate road surface roughness [75]. These two independent results alone provide compelling evidence towards the development of a robust and accurate surface predictor. The use of ANN models for road surface identification is a partially proven method, and was used by W. Pasterkamp et al [59] and T. Shiotsuma et al [75] for pattern generalisation and complex curve fitting for highly simplified cases.

It is therefore a reasonable assumption that, by using vibration as a key source of data and by including additional input parameters to previous research, a more general and more accurate ANN model could be constructed to estimate similar output parameters.
This could produce a model that can quickly and accurately predict a number of surface features during stable conditions, and this is the goal of this investigation.

5.2.1 Model Inputs
Wheel speed and suspension position are two self-evident measurable parameters of tyre condition, as are tyre temperature and pressure, and wheel force and torque. These parameters can all provide valuable data on tyre condition, but combined represent an excessive sensor array, which dramatically increases cost, calibration time and data analysis complexity.

Commonly available tyre temperature and pressure sensors are one option, but are designed to measure overall changes in the tyre condition [122] and are therefore not able to accurately measure high-speed variation, as is necessary for vibration measurement. Further, the data available from these two sensor types react very slowly to surface conditions and, if such data were to be used in the proposed model, it would produce a slow model, which is not desired. While there is no doubt that tyre pressure and temperature can provide information on the surface condition (e.g. a low tyre temperature could indicate a wet road), the sensors currently available cannot produce the vibration data needed here. The data available from these sensors would undoubtedly be of benefit to the model, but would provide little helpful data when compared to the outlay required to purchase and install them on the test vehicle and in the marketplace. As a result, tyre temperature and tyre pressure are ignored here.

Forces and torques applied to the wheel could also be measured, but these parameters are difficult to measure and require expensive instrumentation. W. Pasterkamp and H. Pacejka [59] state that the instrumentation required to directly measure the force in three dimensions and moment in one, which is necessary to gain the overall picture of a single tyre’s load, is extremely expensive and requires the use of rotating wheel dynamometers. Instead, they suggest indirectly measuring these parameters using a combination of suspension position potentiometers and wheel upright mounted strain gauges, as well as a steering link load cell and steering wheel angle potentiometer. This is also an expensive and complicated proposition and, while the data available from these sensors would no doubt be highly useful, the cost and complexity of such a system places it outside the bounds of near future passenger vehicle incorporation. As a result, tyre forces and torque are ignored in this model also.
These exclusions leave just wheel speed and suspension position as useful measurable quantities. Wheel speed measurement on each wheel is now common practice, as is it necessary for most ABS, TCS and VDC. Suspension position sensors, however, are not currently used on passenger vehicles, but since they simply consist of a potentiometer mounted to the suspension, they would not constitute a significant expense for future addition. As stated above, they have the added benefit of being able to estimate the normal load placed on the tyre, which can be of direct benefit to stability controllers. These sensors have been installed on the test vehicle, and have a sampling capacity of 100Hz and 1000Hz for wheel speed and suspension position respectively, which means they can accurately measure high frequency vibrations. Further, by using the data from these sensors, it is possible to gain their differentials such as wheel acceleration and rate of acceleration and suspension velocity and acceleration for analysis.

It should be noted that by excluding parameters that define the vehicle’s overall dynamics (i.e. steering wheel angle, acceleration, slip angles, etc.) it is being assumed that these factors have little or no influence on the tyre’s operation as a road surface sensor. This is clearly not the case in many conditions, most notably in critical and near critical situations, but their influences decrease considerably in stable conditions. When the vehicle is operating well into the linear region of its tyre properties (i.e. at low speed, with small steering inputs and low acceleration) its slip angles will become negligible, and so will the effects of the overall vehicle condition. This would eliminate the need for these additional sensors, but also limit the conditions available for testing, which must then be limited to highly stable conditions. Once results are obtained in these specialised cases, it may be possible to include data from other sensors (such as steering angle and acceleration sensors which are integral to stability controllers) to produce models that are applicable over a wider range of conditions.

By limiting testing to highly stable conditions (i.e. at very low slip and slip angle) it is now possible to use wheel speed and suspension position sensors only to construct a significant history of the tyre condition, which can be used to estimate the road surface features. For example, frequent and long sustained suspension accelerations may suggest a very bumpy road, while consistent very low acceleration could indicate a smooth road. This can also be said of the interaction of the tyre with the road, where vibrations induced in the wheel angular velocity are indicative of the current road surface
conditions. Vibration is the key source of information here, and a method of effectively using it as input to a model must be chosen.

There are a number of methods for characterising vibration, and all have potential benefit to the proposed model. These fall into the three broad categories of statistical tools, harmonic analysis and power spectrum density features. These are listed below:

<table>
<thead>
<tr>
<th>Statistical Tools</th>
<th>Harmonic Analysis</th>
<th>Power Spectral Density (PSD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean / RMS</td>
<td>Fundamental frequency</td>
<td>Graph feature</td>
</tr>
<tr>
<td>Moment about mean</td>
<td>Fundamental amplitude</td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>Harmonic frequencies</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>Harmonic amplitudes</td>
<td></td>
</tr>
<tr>
<td>Max. and Min.</td>
<td>Total harmonic distortion (THD)</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>Total harmonic distortion plus noise (THD + noise)</td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This provides a broad list of parameters worth considering, however the list can be shortened significantly. Firstly, fundamental and harmonic frequencies can be considered integral components to the PSD graph. Secondly, THD is a function of fundamental and harmonic amplitudes. Further, many of the statistical tools are functions of each other (e.g. variance = standard deviation^2) and some are irrelevant here, such as the median and mode. T. Umeno et al [62] focused on the observation that at 60km/hr there is a resonant frequency of wheel speed present at around 40Hz on asphalt roads, a feature that is not present on icy roads. As discussed previously, this suggests that significant data can be derived from observing fundamental and harmonic frequencies alone. Furthermore, T. Shiotuka et al [75] observed that the entire suspension acceleration PSD graph altered depending on the roughness of the road, and as a result used 20 and 40 evenly spaced points to characterise the PSD curve as input to the surface prediction model.

These findings strongly suggest that the PSD of a number of parameter histories can be highly useful in estimating road surface features, as it highlights resonant frequencies and can produce a ‘fingerprint’ of the surface characteristics. The problem then arises of how to include PSD as an input to the model. The best method would be to have as input the PSD value of every frequency possible, however considering the complexity of each
PSD graph, this would produce a very large number of inputs. A method that can provide a similar level of information with fewer inputs is necessary.

One option is to calculate each of the resonate frequencies and input them into the model. These should be unique for different characteristics and provide a method for significantly reducing the number of inputs. This would allow for the inclusion of harmonic amplitudes as input, a feature that is yet to be explored. In addition, as there are only a few extra inputs, values such as mean, standard deviation, maximum and minimum and THD and THD + noise of the data sets could also be used.

Another option would be to separate the PSD into a number of specific segments and provide as input statistical information from within each segment. For a sampling frequency of 1000Hz, as used for suspension position, the PSD graph will extend from 0Hz to 500Hz. By breaking this graph into a number of segments and taking mean values for each, it is possible to approximate the shape of the graph in a similar manner (although not the same) as T. Shiotsuka et al [75]. This approximation could be improved by providing further statistical information for each segment, such as the standard deviation, using the process shown in Figure 5.1. Further statistical information of the data history could be included as input.

![Figure 5.1: Method of PSD approximation for use as model input](image)

**5.2.2 Model Outputs**

The goal of the surface predictor is to identify relevant pavement features that may affect the control strategy of a stability controller, and there are a number of options to be considered.
Prediction of the maximum coefficient of friction between the tyre and the road is the most evident choice. This parameter has a direct relationship to the force a tyre can transmit to a vehicle, and its addition to a stability controller would be of clear benefit. Theoretically, it should be possible to construct a model to act as a ‘virtual sensor’, where the model, using a number of other inputs, replaces an actual coefficient of friction sensor. This method, however, has two clear drawbacks. Firstly, maximum coefficient of friction is not an ‘all encompassing’ parameter for stability controllers. Depending on the method used to evaluate the amount of friction it can, for example, have serious limitations on unsealed roads. The surface friction that the detector would measure is much less than if the wheel was allowed to slip more and ‘dig’ into the road. Such an error would result in the controller allowing much less vehicle acceleration than is possible, which is undesirable.

The second, and more substantial, problem is that a coefficient of friction sensor is not available. Without such a sensor it is impossible to obtain the training data necessary for most ANN models, as well as being very difficult to establish model error. Another option to construct a model with coefficient of friction as output is to predict the static coefficient of friction, a parameter similar in application to the dynamic coefficient of friction. This has the same drawback discussed above on unsealed roads, but it is easier to measure, as it only requires a stand alone device (i.e. not mounted to the vehicle) that places a specific load on an object and measures the force necessary to move it based in the equation; \( Force = \text{Coeff. of Friction} \times \text{Normal Load} \). This produces a new problem, however, as it would require baseline studies of each test road to establish average coefficient values. This is not particularly appealing because it would require a large amount of effort to gain the coefficient values, which themselves would be highly approximate along the length of road, and would be variable over time. While using coefficient of friction values would undoubtedly be a very useful model output, it was decided not to use this approach based on the difficulties in deriving values for model training.

As demonstrated, there are very few options for directly measuring features of the road surface. It would be difficult, therefore, to produce training data for the ANN models based on measurable outputs. Another approach considered was to simply characterise different features heuristically. For example, it can be observed that a road is rough or
smooth, wet or dry, or sealed or unsealed. If these features can be defined well, and testing conditions found where it is possible to gain sufficient training data, they could be used to produce a heuristic surface approximator. This has a significant application to current stability control systems, which for example, often operate highly erroneously on unsealed roads [14]. If the model could tell the controller that the vehicle was travelling on an unsealed road then it could change its control strategy to significant benefit. Other heuristic parameters could also be used in similar ways to provide the controller with additional pavement feature information for improved performance. Some heuristic parameters that could be considered are:

- Sealed / unsealed
- Wet / dry
- Smooth / rough
- Surface type (e.g. asphalt / cement)
- Icy / not
- Oily / not
- Potholes / not

These parameters are binary in nature but could be expanded to include sliding scales, such as replacing “smooth / rough” with “degree of roughness” or a similar parameter. It is observed, however, that characterising a surface with a sliding scale would be much harder to implement than using binary values. This is because it is much easier to observe if a road is ‘smooth’ or ‘rough’ than it is to say how rough it is. Implementing a sliding scale would make compiling the model training data difficult because it would require gathering data on many different surfaces to obtain enough data variation. Small surface changes along each test track would also adversely affect the results, making the selection of the track difficult.

This is not the case for binary parameters. For each parameter it is only necessary to gain training data for two cases (true or false), meaning that compiling enough data for ANN training will be a simpler process. The use of binary parameters means that the selection of test tracks would be much simpler. This is because, instead of having to find a number of surfaces with constant values of roughness, say, it is only necessary to find a road that is only either constantly categorised as ‘smooth’ or constantly ‘rough’. When considering a model with multiple outputs, the process of producing a surface predictor incorporating a number of sliding scales would also become exponentially more difficult, whereas the binary multi-output models would simply require a more thoughtful choice of test track. For these reasons it was decided that the use of heuristic binary outputs was
the most appropriate choice. This left only the choice of which parameters to use as outputs of the model, and two methods were adopted.

The first method relied on predicting each of the surfaces travelled on individually. These surfaces are listed below, and are based on easily accessible roads in proximity to the University of Tasmania Hobart campus.

- Dry smooth asphalt
- Dry rough asphalt
- Dry rough cement
- Dry rough unsealed gravel
- Wet smooth asphalt
- Wet rough asphalt
- Wet rough cement
- Wet rough unsealed gravel

The second method, similar to the first, focused on predicting the surface features individually, as listed below.

- Dry or wet
- Sealed or unsealed
- Smooth or rough
- Surface type (asphalt or not asphalt)

These four tests could be used to construct the same outputs as the first method, but can also be used to provide more information on how well the model is performing in separate areas. In addition the training data for the second method can be derived from the first, making comparisons simple.

Of the four surface features studied using the second method, three have obvious affects on tyre traction. A wet road has less grip than a dry one. The way a tyre grips a sealed road is very different to an unsealed one. A rough surface may provide greater friction at the contact patch, increasing performance. The surface type, however, is difficult to identify using the binary method selected. Using this method it will only provide information on whether or not the surface is asphalt or cement (asphalt / gravel is already determined by the sealed / unsealed test). This output was included because it was expected that the tyre would grip an asphalt road in a slightly different way to a cement one, and that environmental conditions will have different effects on each. In addition it
is necessary to generate the outputs of the first method with the outputs of the second, enabling direct comparisons to be made.

5.2.3 ANN Structure
The ANN model proposed here is a multi-input, multi-output type. There are a number of choices for the model architecture, including the type of ANN to use, the way the ANN is internally structured and the method by which to gain the desired outputs.

Back-Propagation Feed-Forward (BPFF) artificial neural networks have been proven over a wide variety of fields and provide simple and robust models. Optimised Layer by Layer (OLL) ANN models, on the other hand, belong to the newer family of ANN models and can dramatically reduce training time and improve accuracy. Both models will be investigated here, first the BPFF to evaluate the optimum network architecture, followed by the OLL to attempt to improve model accuracy and training time, and provide an indication of how useful it is in this application. The most appropriate internal structure for both ANN types can be derived through experimentation, and will be discussed later within this chapter.

This leaves the way the model produces the desired outputs as the final question with regard to ANN structure. In the work performed by T. Shiotsuka et al [75], which attempts to identify different test tracks based on their roughness, a method was used whereby each surface had a corresponding ANN model output. For the seven surfaces tested, this gave an ANN with seven outputs. When the ANN identified a particular surface it was trained to output a ‘1’ to the corresponding output and a ‘0’ to all the others. The same binary method could be applied here, both with regard to the prediction of the different surfaces, and to the prediction of the different surface features. It should be noted, however, that by having a large number of outputs the situation might arise whereby they adversely affect the accuracy of the model. This could become a problem if more outputs were added to the model in the future.

The other option would be to use just a single output neuron, from which all of the required data could be derived. For example, a system where 1=dry smooth asphalt, 2=dry rough asphalt, 3=dry rough cement, etc could be used. Such a system would simplify the ANN structure and make data transfer easier. However, the internal ANN processing required for this one neuron would be greater and the values near the
minimum and maximum outputs would seldom be reached, as the ANN would naturally
tend to the median output values. This is a consequence of the ANN architecture and
would adversely affect model accuracy. Finally, such a result could be gained using the
multi-output approach discussed above with some post processing. As a result, the
binary multi-output approach is considered better than the single-output method.

5.3 Data Aquisition
The surface predictor model proposed here requires no control feedback signals, meaning
that the data can be logged during normal driving conditions for later analysis. As a
result, the level of instrumentation for this part of the research is minimal. Also, all track
testing was performed on open public streets, all in proximity to the University of
Tasmania Hobart campus. This meant that the testing conditions had to be selected to
abide by roads and traffic legislation to ensure safety, as well being appropriate to the
research performed. For testing to be completed in highly stable conditions, it was
decided to perform all surface tests at 50±5km/hr to meet these goals, where the speed
variation is a consequence of the driver’s ability to maintain a constant speed. This
choice of speed was made to ensure speed limits were not violated and to avoid traffic
problems created by travelling too slowly. This speed is well in the linear region of the
tyres as long as the roads do not contain sharp corners or require heavy braking.

5.3.1 Sensor Selection and Data Logging
As stated previously, only wheel speed and suspension position sensors would be logged.
Further, it was decided to only log data for a single wheel to simplify the research and, as
a result, the rear right tyre was selected for data acquisition. This choice was made based
on a number of influences. Most importantly, because the test vehicle is front wheel
drive, the rear wheels have no torque applied to them unless braking. This, at least
partially, removes a variable. Further, because the rear wheels are not steered, the effects
of camber change from steering can be ignored. Lastly, most surface variation of roads
occur towards the verges, so by selecting the driver’s side wheel (right hand drive) this
level of variation was reduced. It should be noted at this stage that the rear suspension of
the test vehicle is not independent. As explained previously, the two rear wheels utilise a
“Trailing Twist Axle” [20] suspension system, which connect both wheels with a single
transverse member and, as such, influence each other’s movement and introduce a
potential source of error that is unavoidable.
Data logging was completed through the MoTeC Advanced Dash Logger, as described previously. This enabled simple data logging and meant that research could be undertaken at an early stage, and with minimal installation cost and effort. Wheel speed was calibrated into units of km/hr, based on wheel angular velocity and circumference, and has an estimated accuracy of ±0.15km/hr at 50km/hr (based on wheel angular position accuracy of 9° for each tooth ≈ 40mm of tyre circumference). However, the way the MoTeC system calculates wheel speed is unknown and is not clear in the documentation, and this error has been observed to be higher in some circumstances. Suspension position was calibrated to measure intervals of 3.74mV, which corresponds to the resolution of the ADL using its analogue voltage input pins 1 to 4 (the only input pins capable of measuring at 1000Hz). A supply voltage of 8V was applied to the sensor, giving it approximately 2100 measurement intervals over its 100mm travel range. This measurement is directly proportional to the position of the potentiometer, meaning that the 3.74mV accuracy translates to 0.0468mm, plus the linearity error of the sensor itself of 0.1mm, making the total sensor error approximately 0.15mm. At this stage of the investigation no effort was made to calibrate the suspension potentiometer to the wheel position, as this information is irrelevant in ANN models and is difficult to obtain. Wheel speed was measured at its maximum value of 100Hz and suspension position at its maximum value of 1000Hz. This produces a large amount of data, restricting the 1Mb ADL to around 7 minutes of logging before the memory had to be cleared.

Worthy of note is the method employed to identify road surface changes. One option would be to download the measured data from the ADL for each surface tested with a note on the surface characteristics. This would be very tedious and would not allow data to be logged over the transition from one surface to another (as would be necessary to test the reaction speed of the ANN model to surface changes) because there would be no method of noting the change. Data was logged in this way towards the beginning of the investigation, but the shortfalls soon became apparent. To solve this problem the ADL program was modified to log trip distance, with one of the toggle switches programmed to reset the counter. Each time the surface changed, the trip distance was manually reset and this could be used to identify which surface was which (with appropriate notes). This was a very helpful and simple addition, although it had a low accuracy of around ±0.4 seconds (≈ 6m at 50km/hr) because the reset button was operated manually. However, this reaction time error was not considered a problem for the bulk of this
research because it is irrelevant to the data selection method employed for model training, to be discussed below. It should be noted that if this arrangement was found to be too slow, further accuracy could be accomplished through the use of trackside beacons.

5.3.2 Data Conditioning

Data from the test tracks was uploaded to a notebook PC from the ADL as its flash memory became full. Notes were made to each uploaded file to state which surfaces the data was logged on and the conditions of the day. Once all of the required data was acquired the notebook PC was removed from the vehicle and data conditioning performed elsewhere.

All of the logged data could be viewed through “MoTeC Interpreter” and specific regions selected and saved to “Comma Separated Variable” (*.csv) text files. Using this method each surface type for each test could be isolated into its own text file for later use. Regions were only saved to file where they meet the required speed restrictions of $50 \pm 5 \text{km/hr}$ and did not include large accelerations, decelerations or steering inputs.

Once all of the files were saved to *.csv format they were then reformatted into “Tab Delimited” (*.txt) files. This was done, firstly, to remove all of the superfluous data Interpreter places into the files and, secondly, to place the data in a more useful format for later use. The final format of each *.txt file was two columns, the first containing the measured suspension position, the second containing the measured wheel speed. Both were written to file at 1000Hz (i.e. 1000 rows per second of logging), which meant that the wheel speed measurements were repeated over each 1/100 second segment. The file manipulation was performed in “Microsoft Excel” for simplicity, but because Excel can only cope with around 60,000 rows this meant that the file sizes had to be restricted to this length. When this was the case, multiple files were used to cover the entire data range for each test track.
5.3.3 **Road Selection**
The selection of different test roads was based on a number of criteria, listed below:

- Must represent the desired surface very well;
- Must have a constant surface that extends for at least 1km;
- Must have light traffic to enable constant travel at 50km/hr;
- Must not contain any sharp corners or obstacles that require braking;
- Must not have any external influences on road condition (e.g. road works); and
- Should be in proximity to the University of Tasmania Hobart campus and each other.

Bearing these points in mind a number of test tracks were chosen in three locations, as shown in Figure 5.2, Figure 5.3 and Figure 5.4. The Sandy Bay testing area also extends further northwest along the same road, and is separated into two regions as the surface changes from cement (pictured) to smooth asphalt.

The Lower Sandy Bay road offers smooth asphalt and rough cement testing conditions, although the smooth asphalt sections have been repaired multiple times and as a result are a little ‘patchy’ in places. The road follows the shoreline and as a result is mostly level, and also contains very few corners. The Dynnyrne track, conversely, is constantly sloped and has a number of sweeping corners. This section consists of a newly resurfaced road and is characterised as smooth asphalt. This section was included in addition to the Lower Sandy Bay section because neither tracks were of sufficient length for ANN training. Lastly, the Fern Tree test track consists of a rough asphalt section, followed by unsealed and rough asphalt sections. The first asphalt section, on average, slopes slightly downhill. The adjacent unsealed section slopes steeply downhill and about halfway along becomes level for the remainder. These two sections are reasonably winding, but were not considered to be of an extent to significantly affect the model results. Along the unsealed road there are some very sharp corners, which were excluded from the logged data. The final rough asphalt section is level and has only a few sweeping corners.
5. Pavement Feature Recognition during Stable Driving Conditions

Figure 5.2: Lower Sandy Bay testing area [123]

Figure 5.3: Dynnyrne testing area [123]
These test tracks were chosen to provide a large variation in the data presented to the model. This ensures that the ANN models will generalise the relationships that are desired, and not overfit for certain specific features that may be present on individual roads. Additionally, data was logged with the vehicle travelling in both directions along
the roads, effectively doubling the test track lengths. The same tracks were used for testing in wet and dry conditions.

### 5.3.4 Track Testing
Wet testing was performed only when it had been raining for an extended period of time, to ensure a totally wet track. Conversely, dry testing was only to be accomplished after an extended period of no rain to ensure the roads (especially unsealed) were dry. This meant that it was not possible to gain the testing data for both conditions within a short period of time. Additionally, testing was performed in two separate sets with the vehicle set up slightly differently, both situations separated by a large period of time.

Testing data was first gathered for initial surface testing on 31st July 2003 for wet and 11th August 2003 for dry. This data was gathered over a short period of time and used for initial ANN model investigation to gain an idea of the processes involved, which will be presented later. This data was not used in any other research and was not integrated in any way into the data gathered later, as it was observed that the modifications to the test vehicle performed afterwards would void future comparison.

The testing data used for the majority of this research was compiled on 12th March 2004 for dry and 23rd April 2004 for wet. In both circumstances testing was completed within a two-hour period, so condition variation was expected to be low during each day. It should be noted that the weather on both of these days was considerably different from each other, and as a result there were large variations in some aspects. During dry testing there was no cloud cover and the temperature was approximately 24°C, whereas during wet testing it was overcast and lightly raining, with a temperature of approximately 18°C. This provides some variation, such as in pavement temperature, which may have a bearing on the communality between wet and dry results, but could not be avoided.

### 5.4 Software
All of the software used in this investigation was developed in LabVIEW 6i [96], and was created over a long period of time. Investigation into this area was initiated at an early stage to gauge likely outcomes and identify avenues worth following. Initial programs were created as a reflection of this and were designed not to produce reliable, repeatable data, but to give indications of what was likely to work, and what wasn’t. It
should be stressed that these initial programs were developed as an aid to understanding the problem, and not as a consequence of it. They are not comprehensive and any results derived from them were taken as an indication of performance only. These programs follow, all of which provided valuable insight.

- “Frequency Power Spectrum.vi”;
- “Create Surface Test Training File_Original.vi”; and
- “ANN Training from File_Graphical.vi”.

The lessons learnt from these programs were used at a later date to complete more thorough research. They provided insight into the problem at hand, and enabled accelerated research in a particular direction, rather than a broad reaching approach that would have been necessary otherwise. It was decided to write additional LabVIEW programs to investigate theses avenues, and the following files were produced:

- “Create Surface Tests Training File.vi”;
- “ANN Training from File.vi”; and
- “Use ANN from File.vi”.

The purposes of each of these programs will be individually discussed here, and a diagram of the simplified operation of the program will accompany each description.

### 5.4.1 Frequency Power Spectrum.vi
The frequency power spectrum of the measured data is an important source of road surface characteristics, as discussed above. This program was written to observe the frequency power spectrum under a number of different conditions to examine their effects and gain an understanding of the required signal conditioning. Figure 5.5 shows the summary of the program, and demonstrates that there are a number of data processing options. Firstly, once the required data is read from file, the user can filter it using a number of different filters and filter options. These filters include “Chebyshev”, “Inverse Chebyshev”, “Elliptic” and “Bessel”, and can be bypassed. The filtered array is written to screen and compared to the unfiltered data before being passed to a decimating function, where it is possible to effectively reduce the sampling rate of the data to the required amount. This function has the capacity to filter the data using an averaging technique.
LabVIEW VI: Observe the Frequency Power Spectrum of a Logged Signal and its Derivatives - 19/8/03

Note: This program shows the power spectrum density of a measured parameter over time. Results are ONLY printed to screen.

Input Data from File & User Options

Filter Measured Signals

Select Required Data (specify time period)

Decimate Data

Compute Power Spectrum

Power Spectrum Graph

Figure 5.5: “Frequency Power Spectrum.vi” basic functions and summary
The filtered data is sent to the LabVIEW “FTT Power Spectrum.vi” and the power spectrum written to screen with a number of user options such as data “windowing”. The same data is differentiated and passed through another decimating process to produce a filtered data graph and a power spectrum graph for the 1st derivative of the measured data. This is repeated for the 2nd derivative as well. All results are written to screen only, and allow the user to modify parameters to produce the desired power spectrum properties and observe their effects quickly and simply.

5.4.2 Create Surface Test Training File_Original.vi
The “Create Surface Test Training File_Original.vi” was produced as a preliminary program for analysing the ability of ANN models to predict road surface characteristics. It started out as a simple program but grew as more features were added and further avenues explored, and is not an ‘efficient’ program. Figure 5.6 shows the program summary, and has been written to produce a file for later ANN training, and can also be used for online ANN testing.

The program first reads data from up to 20 files and assigns a user selectable output to each, which corresponds to the desired surface output. This means each file must contain data for one output value only. Arrays of wheel speed, suspension position and desired output are produced and in a While Loop, which simulates the process over time. The required segment of data history to be analysed is selected for each loop, which has a user selectable length in seconds. This data passes through a filtering process as above, is differentiated to produce 1st and 2nd derivatives and is then passed to the “Signal Analysis_SubVI_1-9-03_.vi” for custom signal analysis. This produces a number of signal parameters including harmonic amplitudes and frequencies, as well as signal THD, THD + noise, mean and standard deviation of the data segments for wheel speed and suspension position, and their derivatives. This produces an array of 144 parameters that, in conjunction with the desired outputs (of which there is a single-output or a multi-output option), are saved to file for later ANN training.
LabVIEW VI: Produce Surface ID Training File and Test ANN from Logged Data Files (preliminary) - 15/09/03

Summary

Input Data from File & User Options

Filter Measured Signals

Signal Derivatives

Signal Analysis

Write Training Data to File

Use ANN to Test ANN Prediction

Figure 5.6: “Create Surface Test Training File_Original.vi” basic functions and summary
If the data is to be used for ANN testing, the ANN training architecture is read from file and used to condition the “Use ANN_Single Pattern_Graphical_SubVI.vi” subVI (which is described below). The required 144 inputs are produced the same way as above, normalised, and then feed into the subVI. The output of the ANN is displayed on screen in real time and also saved to file. If a binary multi-output model is used a “Confidence Level” is also provided, which was added to give an indication of the expected accuracy of the ANN output by observing the model deviation from a “0” or a “1” result.

5.4.2.1 Use ANN_Single Pattern_Graphical_SubVI.vi

The “Use ANN_Single Pattern_Graphical_SubVI.vi” computes the forward pass of the BPFF ANN using the LabVIEW graphical code, and is summarised in Figure 5.7. The input data, ANN architecture and 1st, 2nd and output layer weights are taken as input and used to determine the ANN predicted output. The program is designed to accept one or two hidden layer architectures, where defining the number of 2nd layer neurons as 0 identifies a single hidden layer ANN.

![LabVIEW SubVI: Use ANN Single Pattern Graphical SubVI.vi](image)

**Figure 5.7: “Use ANN_Single Pattern_Graphical_SubVI.vi” basic functions**

The subVI is structured in such a way as to calculate the ANN output using a serial computation method. The output of each 1st layer neuron is calculated within a For Loop before being passed to the 2nd layer (if applicable) and to the output layer. The outputs of the 2nd and output layer neurons are calculated in the same manner, which includes a summation of all the weighted neuron inputs followed by a sigmoid function of the
result. It should be highlighted that this serial computation method would be much slower than using the parallel computing method that ANN models allow. Nonetheless, using conventional PCs and programming languages this is unavoidable.

5.4.3 ANN Training from File_Graphical.vi

The “ANN Training from File_Graphical.vi” program, as summarised in Figure 5.8, was created as a generic tool for BPFF ANN training from file. Process input and output data is supplied from a correctly formatted text (*.txt) file and the user is allowed to enter a number of parameters to affect the training strategy employed. The program then proceeds to train the ANN to minimize error, saves the architecture to file and completes an input importance analysis. The data written to file can be incorporated into other programs to run the ANN models. This program is useful for training BPFF ANN for any process because it has a generic structure and can handle multi-input, multi-output models with one or two hidden layers. It has the capacity to train the ANN multiple times using different architectures as an aid in identifying the model that produces minimum error.

The user firstly enters the 1st and 2nd layer neuron range desired and the program produces a 2D array of all of the possible architecture combinations. The user enters the file training location, the number of input and the number of outputs, and a subVI is used to extract the input parameter names, the input data array, the input min. and max., the output parameter names, the output data array, and the output min. and max. The program then separates this data into “Training” and “Testing” sets, where the training set is used only for ANN training and the testing set only for training evaluation, such as the determination of model error and input parameter importance analysis.

The training and testing data is normalised for use in the ANN models and the initial weights randomly generated into 2D arrays of the required size. ANN training is completed iteratively within a While Loop, where the training array is randomly indexed and used as input to the “Train ANN_Graphical_SubVI_vi”, which is described below. This subVI then updates the arrays of the ANN weights and the ANN weight adjustments (for training momentum) to be used in the next iteration. The rate of convergence is controlled by the “Learning Constant” and the “Momentum Constant”, which can be modified at each iteration if desired. A facility is provided to progressively
reduce these values along an exponential curve during training to provide large weight updates initially and small adjustments towards the end of training.

A capability is included within the training While Loop to display the ANN training progress. After a user selectable number of loop iterations a function is initiated to run the ANN (“Use ANN_Single Pattern_Graphical_SubVI_vi” detailed above) with the current weights through all of the testing data to display the results on a graph and calculate RMS error. If the error is the lowest yet seen the program will save these weights separately for later use.

Once ANN training has been completed (the number of iterations has reached the user’s limit) the program outputs the final weights to “1st Layer Weights (Last Iteration)”, “2nd Layer Weights (Last Iteration)” and “Output Layer Weights (Last Iteration)” and the minimum error weights to “1st Layer Weights (Min. Error for Display Output)”, “2nd Layer Weights (Min. Error for Display Output)” and “Output Layer Weights (Min. Error for Display Output)”. These weight arrays are saved to file in a specific format for later use. The weights are used again through the “Use ANN_Single_Pattern_Graphical_SubVI_vi” for all of the data to produce the ANN predictions, and the results written to file. The same process is repeated for the testing data only for error analysis.

The last stage of the program is to conduct the input importance analysis. Here the normalised testing input array is taken and one of the inputs replaced with its half magnitude value (0.5) for the entire set. This is analogous to that particular input sensor failing. The “Use ANN_Single_Pattern_Graphical_SubVI_vi” is again used to run the data through the ANN model, and the resultant RMS error compared to the RMS error of the ANN with all working inputs. This process is repeated for all of the ANN inputs, and the results written to file and to the display.
LabVIEW VI: Train ANN and importance analysis (original program) - 17/11/03

Note: This program allows the user to select a number of different architectures for ANN. It separates the input file into a training and a testing set and then performs the ANN training. A parameter importance analysis is also performed.

Summary:
- Input Data from File and User Options
- Separate Training and Testing Data Randomly
- Normalise Data
- Train ANN at Selected Learning Rate
- Display Training Progress
- Finish ANN Training
- Save Weights to File
- Use ANN
- Save ANN Prediction to File
- Perform Importance Analysis
- Write Importance Analysis to File

Figure 5.8: “ANN Training from File_Graphical ANN.vi” basic functions and summary
5.4.3.1 Train ANN_Graphical_SubVI_.vi
This subVI is designed to compute the weight updates required for BPFF ANN training, and is summarised in Figure 5.9. This is achieved in three stages, the first stage is identical to the “Use ANN_Single Pattern_Graphical_SubVI_.vi” detailed above. It also has more program inputs, which include learning rate, momentum constant and 1st, 2nd and output layer weight adjustments for the previous iteration.

The second stage of the program is used to back propagate the output error through the ANN layers. The scaled error for each of the output layer neurons is calculated by comparing the ANN output to the desired output using a For Loop, before passing the resultant information to the 2nd layer neurons to calculate the 2nd layer scaled error (if applicable). This information is then passed to the 1st layer neurons for scaled error evaluation using the same method.

The third stage updates the ANN weights based on their scaled error values. This function includes the learning rate to augment training convergence speed and momentum (incorporating the 1st, 2nd and output layer weight adjustments from the previous iteration) to help avoid local minima. These final functions output the new, updated, ANN weights and the ANN weight adjustments to be used for further iterations. The ANN predicted output is provided if needed.

**Figure 5.9: “Train ANN_Graphical_SubVI_.vi” basic functions**

The graphical nature of this subVI made it very complicated to follow and program, as seen from the code. Indeed, the debugging phase of the development of this subVI was extremely long and there is still a possibility of small errors, which if they exist are...
considered to have a small to negligible effect on training speed and accuracy. This fact highlights the complexity of graphical programming in some situations, and should be avoided in future work.

5.4.4 Create Surface Tests Training File.vi

It was decided that the “Create Surface Test Training File_Original.vi” discussed above was too complicated and not general enough for further research, so a new program was written to replace it. The way files were input into the original program was tedious and, given the way the model outputs were input by the user, there was a strong possibility of error. Functions such as decimating the arrays were found to be of little benefit, and were omitted. Data filtering was found to have some advantage occasionally and so was implemented into the new program in a simplified form. It was also found that the binary multi-output ANN models performed better than the single-output models, so the single-output option was removed from the program. The ability for online ANN testing was found to have little use at this stage of research and was omitted in favour of a separate program for this purpose.

LabVIEW VI: Create Surface Training File from Multiple User-Selectable Measurement Files

The new “Create Surface Tests Training File.vi” is summarised in Figure 5.10. This program was designed to allow the user to select data files (each containing data for single output values only) in an easier manner and to obtain model training outputs from the file names, which must be formatted correctly. This reduces the likelihood of user error significantly and simplifies the use of the program.
The input arrays of wheel speed and suspension position are filtered (if desired) using a simplified method. The wheel speed measurements are extrapolated to 1000Hz equivalent by means of an averaging method between data points to avoid high frequency noise in the PSD. This is different to the previous model, which simply read the 100Hz data at 1000Hz producing a severely stepped waveform. This could be improved in future to reduce the level of calculation required, but was not considered an important condition at this stage.

The next stage in the program differentiates the wheel speed and suspension position data to gain their derivatives. An extra feature was included here to allow the user to select any number of derivatives to remove the confines of only analysing up to the 2nd derivative in the previous program. The data was then used to analyse the wheel speed and suspension position signals and their derivatives. The signal analysis was completed by presenting measured data history segments of a user selectable length (in seconds) to a separate subVI called “Measure Pattern Properties_ Power Spectrum Segments_SubVI_.vi”, which is detailed in Figure 5.11. The results from this subVI were written to file, along with the desired model output, for use in a separate program to complete ANN training.

![Figure 5.11: “Measure Pattern Properties_ Power Spectrum Segments_SubVI_.vi” basic functions](image)

5.4.4.1 Measure Pattern Properties_ Power Spectrum Segments_SubVI_.vi

This subVI was written to provide a flexible tool for selecting which signal analysis parameters were to be written to the ANN training file. As it stands, the program takes the desired 1D array segment as input and produces statistical, harmonic and PSD data.
that can be incorporated into the subVI output array. The possible outputs of this program include, but are not limited to:

- Array mean;
- Array standard deviation;
- Array minimum and maximum;
- Harmonic amplitudes and frequencies;
- Total harmonic distortion and THD + noise; and
- PSD curve.

A great deal of data can be derived from the PSD curve, and within this subVI the mean and standard deviation of the curve for 0-1, 1-5, 5-20, 20-50, 50-100, 100-200, 200-300, 300-400, 400-500Hz frequency segments are calculated and used as subVI output.

5.4.5 ANN Training from File.vi

The original “ANN Training from File_Graphical.vi” program was found to have a number of limitations in practice, and was replaced. Most notably, the program was written with little programming experience with LabVIEW, and as a result, it is far more complicated than necessary. This complexity made the program difficult to follow, producing debugging problems that may not have been resolved. The program required the user to enter a number of redundant parameters and included a number of superfluous calculations that slowed it down.

The new “ANN Training from File.vi” solves these problems. The program still performs the same function as its predecessor, but is faster, easier to use and more reliable. Firstly, the number of user inputs has been reduced so the user only needs to enter the training file path and the number of outputs within the file for the program to gather much of the required data. Other required inputs include the number of hidden layer neurons, what percentage of data is to be set aside for testing, the number of training iterations to use and parameters defining learning constant and momentum. There is a control over the magnitude of the initial random weights to improve convergence speed.

The first function of the program is to open the training file and separate it into input and output, and training and testing arrays. At this stage a new directory is created and the
training file copied to this location for future reference. There is also the option of using separate, pre-formatted, training and testing files to reduce the necessity to repeat this long process when re-running the program and to allow new testing data to be used. The user 1D array inputs of “Number of 1st Layer Neurons” and “Number of 2nd Layer Neurons” are read and used to create a 2D array of every combination of desired hidden layer neurons (if “Number of 2nd Layer Neurons” = 0 then a single hidden layer architecture is used).

Once data has been established, the program enters a For Loop for each of the hidden layer neuron combinations. Here the learning rate and momentum parameters are entered and the ANN weight arrays initialised using random values of a user selectable magnitude.

ANN training is completed within a While Loop, which finishes when a selected number of iterations have been completed. Here the learning rate and momentum values are progressively reduced and ANN training is completed by iteratively presenting the required data to the ANN training subVI “Train ANN with Output_0 to 1_SubVI_.vi” or “Train ANN with Output_-1 to 1_SubVI_.vi”. These two subVIs perform slightly different operations, as discussed below, and allow the user to select between two ANN models depending on the characteristics of the input data. Both of these subVIs take as input the ANN weights, previous ANN weight adjustments, input parameters, desired output and learning rate and momentum. They provide the ANN output prediction, updated ANN weights and ANN weight adjustments as output, as shown in Figure 5.12. The training While Loop updates these values iteratively, and so progressively trains the ANN model using randomly indexed data from the training array.

Another function is imbedded within the training While Loop to display the training progress on screen. After a specified number of iterations the “Use ANN_0 to 1_with Normal_SubVI_.vi” or “Use ANN_-1 to 1_with Normal_SubVI_.vi” subVI (which ever is selected) is applied to the testing data using the current ANN weights. Using this data the ANN RMS error is calculated and a graph of ANN predicted output is compared to the desired output on screen.
LabVIEW VI: Train a Backpropagation Feedforward ANN from File and Perform Importance Analysis.

**Note:** This program is developed for generic BP training from a correctly formatted file. Training options include the type of BP, the range of 1st and 2nd hidden layers (zero 2nd layer neurons = single hidden layer ANN), number of training iterations and and learning constant variation.

**SubVI:** Train ANN with Output_0 to _1_SubVI_vi, Train ANN with Output_0 to _1_SubVI_vi, Use ANN_0 to _1/SubVI_vi & Use ANN_0 to _1/SubVI_vi.

Figure 5.12: “ANN Training from File.vi” basic functions and summary.
Once training was completed the program then exits the training While Loop and uses the final weights for further analysis and writes them to file. The input parameter importance is calculated first by replacing one input parameter with its median value through the whole testing array, and running it through “Use ANN_* to *_with Normal_SubVI_.vi” to monitor the effects of this change on the RMS error. This was repeated for all of the input parameters, and the difference between each RMS error and the RMS error of the unmodified testing data is saved as the parameter importance. This produces a 1D array of the values that represent how much the ANN predictive error will increase if that input fails. A small increase corresponds to a parameter that has little importance, while a large increase suggests a high importance.

The final function of the program is to produce the ANN predictions for the testing set (using “Use ANN_* to *_with Normal_SubVI_.vi”) and to write the results to file. The process input values, the desired outputs, the output error, the input parameter names, the input parameter importance and the input and output minimum and maximum values are written to the same file.

Further functionality was installed at a later date, to enable OLL ANN modelling. To do this, the sub-programs “Train ANN with OLL SubVI group” and “Use ANN with OLL SubVI group” were added to implement the OLL algorithm that has been described in the theory previously. Although the OLL training method is very different to backpropagation, these programs were designed to fit into the existing program with very minimal modifications.

5.4.5.1 Train ANN with Output_* to *_SubVI_.vi
The “Train ANN with Output_0 to 1_SubVI_.vi” and “Train ANN with Output_-1 to 1_SubVI_.vi” subVIs are both summarised in Figure 5.13. They accept all the data required for training, normalise the input and output parameters and present the result to a LabVIEW Formula Node. The node is embedded inside a Case Structure and contains C++ code for one and two hidden layer ANN training. The C++ code computes the entire training process of one iteration, replacing the highly complex graphical programming of “Train ANN_Graphical_SubVI_.vi” with text, greatly reducing the possibility of errors and making the program more versatile and easy to follow. The Formula Node results are passed to the subVI outputs and the ANN prediction is “un-normalised”.

The normalisation process was included within the subVI in this case to simplify the function of the higher-level programs. It is this normalisation process that is the only difference between “Train ANN with Output_0 to 1_SubVI_.vi” and “Train ANN with Output_-1 to 1_SubVI_.vi”. The former program normalises the data into the 0 to 1 range, while the later normalises the data into the -1 to 1 range. This is done to allow the user to offset the position of the “0” values depending on the type of inputs required. This is useful because if the ANN inputs contain “0” values, the weights have diminished ability to offset the neuron result in that region. By specifying a normalisation process that replicates the input function (i.e. if the inputs are in the 0-x range then normalise to 0-1, if the inputs are in the –x-x range then normalise to –1-1) it is possible to reduce this effect and improve ANN accuracy.

![Diagram](image)

**Figure 5.13: “Train ANN with Output* to *_SubVI_.vi” basic functions**

### 5.4.5.2 Use ANN* to * with Normal_SubVI_.vi
The “Use ANN_0 to 1 with Normal_SubVI_.vi” and “Use ANN_-1 to 1 with Normal_SubVI_.vi” are very similar to “Train ANN with Output* to *_SubVI_.vi” except they perform the ANN prediction part of the C++ code only. As in Figure 5.14, the subVIs take as input the input array and the ANN weights and the max and min values. They normalise the input array into the 0-1 or –1-1 regions and finally pass the data to a Formula Node containing the C++ code. Separate Formula Nodes are selected depending on if the ANN of a one or two hidden layer architecture is desired, and the resultant ANN prediction is “un-normalised” and passed to the subVI output.
5. Pavement Feature Recognition during Stable Driving Conditions

5.4.5.3 Train ANN with OLL SubVI group addition

The original version of “ANN Training from file.vi” consisted of ANN training through backpropagation only, with a Case Structure utilised to determine the type of backpropagation model used as discussed above. As such, the addition of OLL training within this framework required an extra case to be constructed within the Case Structure, and is shown in Figure 5.15.

Unlike backpropagation training OLL models require data to be presented to them in a different way and can only work with one hidden layer. The existing program was programmed to either randomly index input/output patterns or index them repeatedly in sequential order, both of which are valid backpropagation ANN training methods. In contrast, OLL models require the input/output data to be sequentially run through the ANN model three times for a single weight update to be made, with different functions performed each time.

In this manner, the OLL training case comprises of three “rules”, which determine which of the three OLL training phases are taking place. Within each of these rules, the input/output data is run through the OLL training functions as discussed in the theory in previous sections. The three stages (loops) include:

1. Compiling the $A_{ij}$ and $B_{jk}$ matrices based on each input/output set to determine new output layer weights $V_{opt}$.
2. Compiling the $A_{ij}$ and $B_{ji}$ matrices based on each input/output set, using the new output layer weights, to determine new 1st layer weights $W_{test}$.
3. ANN testing with the new $V_{opt}$ and $W_{test}$ for each input/output set to evaluate if new weights increase or decrease error. If error reduces then $W_{test}$ replaces the old 1st layer weights, weight factor ($\mu$) is reduced and process starts again from...
step 1. If error increases the old 1st layer weights are kept, μ is increased and the process start again from step 2.

**LabVIEW VI: Train ANN with OLL SubVI Group**

The internal structure of the OLL ANN is slightly different to backpropagation models, although they both consist of the same input and output variables. As such, a new case was inserted into each of the existing “Use ANN_* to * with Normal_SubVI_.vi” Case Structures to contain the new OLL model. This was a simple process, and essentially takes a number of SubVIs from the “Use ANN with OLL SubVI group” to complete ANN predictions based on current weights, as shown in Figure 5.16. Furthermore, because the case input and output variables have identical functionality as the backpropagation cases, no changes needed to be made in this regard.

**5.4.5.4 Use ANN with OLL SubVI group addition**

The internal structure of the OLL ANN is slightly different to backpropagation models, although they both consist of the same input and output variables. As such, a new case was inserted into each of the existing “Use ANN_* to * with Normal_SubVI_.vi” Case Structures to contain the new OLL model. This was a simple process, and essentially takes a number of SubVIs from the “Use ANN with OLL SubVI group” to complete ANN predictions based on current weights, as shown in Figure 5.16. Furthermore, because the case input and output variables have identical functionality as the backpropagation cases, no changes needed to be made in this regard.
5.4.6 Use ANN from File.vi

“Use ANN from File.vi”, as summarised in Figure 5.17 is a simple program and performs the function that this name suggests. The user enters the location of the “Testing Data” file, the “ANN Use Weights” file, the “ANN Use Max. and Min.” file, the “Number of 1st Layer Neurons” and the “Number of 2nd Layer Neurons” and the program extracts the ANN architecture and model input and desired output data arrays. The testing input data is run through “Use ANN_* to *_with Normal_SubVI_.vi” or the “Use ANN with OLL SubVI Group” using a For Loop and the ANN predictions computed. The desired output data is used to calculate the ANN error and the results written to file and to screen.

This is a generic program that takes any file of ANN inputs and desired outputs and computes the ANN output predictions using a specified architecture. This means that the input data must be formatted correctly, which in this case means it must be run through the “Create Surface Tests Training File.vi” before being used.

5.5 ANN Model Analysis

5.5.1 Initial Investigation

As stated above, the initial part of the investigation of this surface prediction problem was conducted as an aid to understand the problem, rather than to solve it. The goals were to identify which vibration parameters had a bearing on the predictive ability of the ANN models, to gain knowledge the accuracy to expect and to obtain an indication of the most promising avenues to explore. This argument lead to the selection of almost all
parameters LabVIEW could provide through its various signal analysis subVIs for input into the ANN model initially. These subVI functions included; “Auto Power Spectrum.vi”; “Harmonic Analyzer.vi”; “Std Deviation and Variance.vi”; “Extract Single Tone Information.vi”; “Harmonic Distortion Analyzer.vi”; “SINAD Analyzer.vi”; “Power and Frequency Estimate.vi”; “Mode.vi” and; “Moment About Mean.vi”.

Applying the results from these subVIs for suspension position, suspension velocity, suspension acceleration, wheel speed, wheel acceleration and wheel rate of acceleration an ANN model was produced with 210 inputs, as shown in Figure 5.18. The results from this model were promising, but it became apparent that the computation required to produce all of these inputs was intensive and would significantly slow the model. The size of the ANN model was considered too large for effective training.

<table>
<thead>
<tr>
<th>MODEL INPUTS:</th>
<th>ANN MODEL:</th>
<th>MODEL OUTPUTS:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs for suspension position, suspension velocity, suspension acceleration, wheel speed, wheel acceleration and wheel rate of acceleration including:</td>
<td>40 – 200 1st hidden layer neurons</td>
<td>Outputs for surface type formatted depending on ANN model, including:</td>
</tr>
<tr>
<td>• Measured value</td>
<td>0 – 40 2nd hidden layer neurons (&quot;0&quot; neurons means the 2nd hidden layer is removed from the model).</td>
<td>1. V. rough gravel (wet)</td>
</tr>
<tr>
<td>• Sine phase</td>
<td>1 or 9 output layer neurons, depending on the method employed.</td>
<td>2. Rough gravel (wet)</td>
</tr>
<tr>
<td>• Fundamental frequency</td>
<td>If using the one model output option a single neuron will produce an output in the range of 1 to 9 to correspond to each output.</td>
<td>3. Rough gravel (dry)</td>
</tr>
<tr>
<td>• THD</td>
<td>If using the nine model output option each neuron corresponds to each output. A binary response is used to specify which surface the model predicts to be correct.</td>
<td>4. Rough asphalt (wet)</td>
</tr>
<tr>
<td>• SINAD</td>
<td></td>
<td>5. Rough cement (wet)</td>
</tr>
<tr>
<td>• %THD</td>
<td></td>
<td>6. Smooth asphalt (wet)</td>
</tr>
<tr>
<td>• %THD + noise</td>
<td></td>
<td>7. Rough asphalt (dry)</td>
</tr>
<tr>
<td>• Line frequency interval</td>
<td></td>
<td>8. Rough cement (dry)</td>
</tr>
<tr>
<td>• Estimated power spectrum frequency peak</td>
<td></td>
<td>9. Smooth asphalt (dry)</td>
</tr>
<tr>
<td>• Estimated power spectrum power peak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Standard deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Variance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Mode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Moment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 1st to 5th harmonic amplitudes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 1st to 5th harmonic frequencies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 1st to 10th component levels</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With this in mind, a number of input parameters were removed that were either seen to have little impact on model accuracy, or to be repetitious in nature. This process left 102 model inputs, which included the measured value, mean, standard deviation, sine phase, SINAD, 1st to 5th harmonic amplitudes, 1st to 5th harmonic frequencies, %THD and
%THD + noise. Using these inputs alone, it was observed that the model error did not alter considerably.

The next stage of investigation attempted to identify whether the wheel speed parameters had a larger or smaller impact than suspension position parameters on the model accuracy. This was considered an important consideration because, if it were possible, one of these two sets of parameters could be ignored and model complexity reduced significantly. To this end, the model inputs were split between two separate ANN models of 51 inputs each, one using just wheel speed data, and the other just suspension position. The error produced by these two models substantially increased, indicating that data from both sensors was vital in providing the level of accuracy desired.

The effects of data filtering was considered next, as problems with the data acquisition were discovered. It was observed that the way the sensors collected the data produced a “stepped” waveform. This was because, firstly, when “MoTeC Interpreter” writes the 100Hz wheel speed data to file at the 1000Hz of the suspension data it simply repeated the same measurement 10 times. The stepped waveform that this produced gave erroneous vibration characteristics. The second problem that arose was from a hardware error. The “MoTeC ADL” is designed to measure four inputs at 1000Hz and the rest at 500Hz. This was not evident at the time of installation, and as a result the suspension position sensor was wired to produce a maximum sampling rate of 500Hz only. This was not highlighted in the logging software, which continued to log at 1000Hz, producing another stepped waveform. Rectifying this error within the program was considered a difficult task, so an ability to filter the data was included instead. By observing the effects of various low pass filters it was possible to gain smooth curves without diminishing the accuracy of the data significantly.

Using this filtered data the ANN models were tested again for dry test data only, wet test data only and combined wet and dry data. This was done to observe the effect of the data filtering, as well as the model accuracy between wet and dry conditions. Remarkably, the data filtering was found to produce only a small increase of performance, and the wet/dry tests showed favourable results. It was observed that accuracy improved reasonably in dry only tests, as it did in wet only tests - except to a smaller extent. The combined wet and dry tests produced more error than the separate
ones, but this was taken as a positive result as it showed the model was capable of distinguishing between a wet road and a dry one, on a variety of pavements.

The ANN inputs were modified after it was observed that the 1st to 5th harmonic frequencies were confined to a small frequency range. It was decided to further remove some of the seemingly superfluous inputs and replace them with higher levels of harmonics. This resulted in the selection of 146 input parameters, including the measured value for wheel speed and suspension position only, mean, standard deviation, 1st to 10th harmonic amplitudes, 1st to 10th harmonic frequencies, and %THD and %THD + noise as the ANN inputs. These inputs were left unmodified for the remainder of this initial investigation.

Next, the type of model output was investigated. Up until this stage a single output neuron was used to produce an output in the range 1 to 9, where each integer represented a different surface, in a wet or dry condition. The method was observed to work adequately, but had a number of perceived drawbacks. Most notably, by using this method the ANN output has a tendency to predict values in the median of its range (i.e. values around 5) because this type of ANN model has trouble driving outputs to its extremities. This was a significant problem and other output methods were investigated.

The output method that was selected in the end was a “binary” nine-output model. Using this method, a single output neuron was used to correspond to each surface. A “0” indicated a false test, while a “1” indicated a true test. If a neuron showed a “True” result it was used to identify the road surface. This allowed the user to gain more information from the output data, such as the “confidence” of the model. This “confidence” level can be derived by how much the analogue model output deviates from the binary condition desired. A model output near a “0” or a “1” indicates a high level of confidence, while a value far from this binary condition of, say, 0.7 shows a model that is not sure of its output and less accuracy can be assumed. Using this method it is possible to gauge the level of likely error from the ANN model alone, which could be a valuable tool in later investigation.

Further research was carried out into the internal structure of the ANN. The first test was designed to gauge the impact of the inclusion of a small number of linear neurons into the ANN models, as suggested by O. Sørensen et al [124]. While this was intended to
increase the curve fitting abilities of the ANN model by increasing its linear response, the opposite proved the case and the inclusion of linear neurons was abandoned. Performance was also evaluated when a different normalisation process was used, whereby the 0 to 1 normalisation used previously was replaced with a –1 to 1 normalisation. This process shifts the “0” value, which can have a large effect on the operation of the ANN model depending on the type of inputs and outputs. In this case, however, the modification proved to have a negative effect and was also abandoned.

The last avenue to be explored during the initial investigation was the length of the data array to be used for signal analysis. The signal analysis needs a history of data to be able to operate, and the length of that data has a large impact on the accuracy and speed of the model. At this initial stage, two history lengths were investigated for this purpose of 3 seconds and 5 seconds (or 3000 samples and 5000 samples). It was found that the 5 second option produced a slightly more accurate model of 21.8% error compared to 22.4% of the 3 second model, but it was observed that a significant amount of extra processing was required to produce this small gain. As a result the 3 second option was selected for subsequent work.

In all, 26 different tests were performed at this initial stage to provide a basis for future study. While it is clear it does not constitute a systematic approach of problem identification and solution, they provided a large amount of useful insight into the problem at hand, and allowed a number of apparent “dead ends” to be explored without a large degree of effort. Further, in-depth, investigation could then be conducted using the lessons learnt here to converge towards the desired model at a greater pace.

### 5.5.2 Input Parameter Importance

Four months passed between the initial investigation above, and the subsequent work below. This large time period enabled a fresh approach and the inclusion of new ideas, which proved to be extremely valuable. Firstly, during this period, a new ANN training program was produced which was more efficient and greatly reduced the likelihood of internal errors. The nature of vibration was also studied and a number of additional insights were gained. Comprehensive coverage of the modeling results are given in the Appendix.
Of greatest significance, it was observed that the harmonic amplitudes and frequencies were not likely to contain much of the data necessary for surface prediction. It was the shape of the PSD curve that was of most interest, and not necessarily the frequency peaks. For example, a region in the PSD curve of a comparably low value can provide as much data on road surface type as a region of large magnitude. To this end it was decided to replace the existing model harmonic inputs with PSD mean and standard deviation values for different segments of the curve, as shown previously in Figure 5.1. This increased model accuracy, and was adopted throughout future work.

In combination with the previous investigation results, an ANN model was constructed as shown in Figure 5.19. To produce this model a number of new programs were written to increase their efficiencies and enable simple modification in future. The suspension position sensors were re-installed to provide a sampling rate of 1000Hz and improved accuracy. This higher sampling rate eliminated the stepped suspension waveform, making low pass filtering unnecessary. The wheel speed sensor still required signal conditioning to convert the 100Hz signal to 1000Hz without the stepped waveform, and
this was done using a simple averaging technique over 10 samples. This data was then used for input to the signal analysis subVIs using 3 second segments, producing the required data for the 138 input ANN.

5.5.2.1 Investigation Goals
138 inputs were considered too many inputs for the ANN model, therefore the number of inputs was reduced while maintaining low error. To this end, further work was carried out to approximately halve the number of model inputs.

A simple method of accomplishing this task was to perform an input importance analysis of all of the inputs for the model and remove the inputs of least importance. It was anticipated, however, that this might not produce the most reliable results. This was based on the observation that a single input may be highly important in predicting a single model output in special conditions, but could have a low overall importance and may be excluded. This was undesirable so a different approach was planned.

The approach that was selected was based on training a number of single output models instead of one multi-output model. That is, a separate model was used to predict whether a road was smooth asphalt or not. Another model could then be used to predict whether the road was wet or dry. It was anticipated that by doing this more, ANN accuracy could be achieved for each output, making any calculation of input importance more useful. It would also enable the user to observe which parameters had the most relevance in determining particular surface features. The results of each of these tests could then be compiled and the inputs selected for future work.

5.5.2.2 Investigation Method
ANN models were to be trained and tested for a large variety of input data and output results. These models included single outputs of only:

- Smooth asphalt in dry
- Rough asphalt in dry
- Rough cement in dry
- Rough gravel in dry
- Asphalt surface in dry
- Smooth asphalt in wet
- Rough asphalt in wet
- Rough cement in wet
- Rough gravel in wet
- Asphalt surface in wet
- Smooth asphalt in wet & dry
- Rough asphalt in wet & dry
- Rough cement in wet & dry
- Rough gravel in wet & dry
- Asphalt surface in wet & dry
- Wet road or dry road
The accuracy of these outputs could then be separately evaluated, and the strengths and weaknesses of the surface prediction model identified. Once each of the models was trained it would be a simple problem of determining individual input importance. Once this data was obtained it could be used to compile the list of which inputs to include and exclude for future work.

Further testing was conducted at this stage to gauge the accuracy of these single-output ANN models to the original multi-output models. By training the multi-output models to produce smooth asphalt, rough asphalt, rough cement, rough gravel and (if needed) wet or dry road outputs within a single model its accuracy could be compared against the single-output models.

A specific ANN internal architecture was selected at this stage to accelerate investigation and enable meaningful comparison. This architecture consisted of a single hidden layer ANN with 100 neurons within the hidden layer, all containing sigmoid activation functions. The outputs were of a binary nature, as discussed above, and consisted of any combination of smooth asphalt, rough asphalt, rough cement, rough gravel or wet/dry test as necessary for the selected model. The normalization process was always of the 0 to 1 type.

5.5.2.3 Input Importance and Selection of Parameters
The resultant importance analyses for the single-output and multi-output models are given in the Appendix. It was found that the single-output and multi-output models produced similar importance on the same inputs, and that these did so consistently when re-trained. It is therefore evident that the models are consistently converging to similar solutions, which indicates that the ANN models were identifying clear relationships between the input parameters and the desired outputs. The result shows that the single-output models are providing the same information as the multi-output models, and can consequently be avoided in future. This would dramatically reduce investigation time and effort, because it means only one multi-output model is needed instead of four or five single-output models to produce the same data.

This discovery weakens the argument for the need to identify the importance of each input on each output, as the multi-output model has shown itself as being capable of such
an application. However, the argument still exists that some important parameters for specific features may not be important enough in an overall sense to avoid being excluded. It was therefore decided that two methods of exclusion be adopted and compared to produce the final list of input parameters for further model investigation.

The first method is very simple, and required only the overall importance analysis for the wet and dry multi-output model. The parameters were ranked in importance and, once a desired level of importance was reached, the less important parameters were excluded.

The second method, as originally proposed, is more complicated. It involved ranking the top ten inputs for each of the single-output models, as shown in Table 5.1, and including them into the list of input parameters to be used for further research. The top twenty parameters for the wet/dry test were included. This was done because it is not as highly represented as the other model outputs within the list.

Further investigation showed that a number of these high-ranking parameters are repeatedly high-ranking throughout each single-output model. This is an indication of their relative importance to the predict ability of the overall model, and using this observation it is possible to rank all of the model input parameters based on how high they rank, and how often. The result of this method is given in Table 5.2.

The importance rank given in Table 5.2 was directly compared to the importance rank derived from the multi-output model. Apart from some small variation in the ranking of parameters, both tables showed similar importance placed on most inputs. This is a good sign, and means that the selection of parameters for exclusion was based on two sets of independent results.
### Table 5.1: Top ten most important input parameters for individual outputs (in descending order)

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Dry Smooth Asphalt</th>
<th>Wet Smooth Asphalt</th>
<th>Wet &amp; Dry Smooth Asphalt</th>
</tr>
</thead>
<tbody>
<tr>
<td>104 WS d/dt 20-50Hz Mean</td>
<td>58 SP d2/d2 20-50Hz Mean</td>
<td>36 SP d/dt 20-50Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>127 WS d/d2 20-50Hz Mean</td>
<td>37 SP d/dt 50-100Hz Mean</td>
<td>37 SP d/dt 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>95 WS d/dt</td>
<td>38 SP d/dt 50-100Hz Std. Dev.</td>
<td>58 SP d2/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>118 WS d/d2</td>
<td>36 SP d/dt 20-50Hz Std. Dev.</td>
<td>81 WS 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>65 SP d/d2 50-100Hz Mean</td>
<td>127 WS d/d2 20-50Hz Mean</td>
<td>41 SP d/dt 200-300Hz Mean</td>
<td></td>
</tr>
<tr>
<td>38 SP d/dt 50-100Hz Std. Dev.</td>
<td>98 WS d/dt 0-1Hz Mean</td>
<td>79 WS 5-20Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>37 SP d/dt 50-100Hz Mean</td>
<td>11 SP 5-20Hz Std. Dev.</td>
<td>84 WS 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>55 SP d/d2 1-5Hz Std. Dev.</td>
<td>56 SP d/d2 5-20Hz Mean</td>
<td>56 SP d/d2 5-20Hz Mean</td>
<td></td>
</tr>
<tr>
<td>29 SP d/dt 0-1Hz Mean</td>
<td>60 SP d/d2 50-100Hz Mean</td>
<td>36 SP d/d2 50-100Hz Mean</td>
<td></td>
</tr>
<tr>
<td>133 WS d/d2 200-300Hz Mean</td>
<td>50 SP d/dt 20-50Hz Mean</td>
<td>38 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Dry Rough Asphalt</th>
<th>Wet Rough Asphalt</th>
<th>Wet &amp; Dry Rough Asphalt</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 WS d/dt 1-5Hz Mean</td>
<td>39 SP d/dt 200-300Hz Mean</td>
<td>36 SP d/dt 1-5Hz Mean</td>
<td></td>
</tr>
<tr>
<td>3 SP Std. Dev.</td>
<td>105 WS d/dt 20-50Hz Std. Dev.</td>
<td>61 SP d/d2 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>35 SP d/dt 20-50Hz Mean</td>
<td>58 SP d2/d2 20-50Hz Mean</td>
<td>81 WS 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>31 SP d/dt 1-5Hz Mean</td>
<td>36 SP d/dt 50-100Hz Std. Dev.</td>
<td>37 SP d/dt 50-100Hz Mean</td>
<td></td>
</tr>
<tr>
<td>59 SP d/d2 20-50Hz Std. Dev.</td>
<td>56 SP d2/d2 5-20Hz Mean</td>
<td>13 SP 20-50Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>69 SP d/d2 400-500Hz Std. Dev.</td>
<td>98 WS d/dt 0-1Hz Mean</td>
<td>97 WS d/dt THD+Noise</td>
<td></td>
</tr>
<tr>
<td>104 WS d/dt 20-50Hz Mean</td>
<td>126 WS d/d2 5-20Hz Std. Dev.</td>
<td>86 WS 200-300Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>9 SP 1-5Hz Std. Dev.</td>
<td>55 SP d/d2 1-5Hz Std. Dev.</td>
<td>58 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>55 SP d/d2 1-5Hz Std. Dev.</td>
<td>82 WS 20-50Hz Std. Dev.</td>
<td>59 SP d/d2 20-50Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>45 SP d/dt 400-500Hz Mean</td>
<td>72 WS Std. Dev.</td>
<td>10 SP 5-20Hz Mean</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Dry Cement</th>
<th>Wet Cement</th>
<th>Wet &amp; Dry Cement</th>
</tr>
</thead>
<tbody>
<tr>
<td>38 SP d/dt 50-100Hz Std. Dev.</td>
<td>58 SP d2/d2 20-50Hz Mean</td>
<td>37 SP d/dt 50-100Hz Mean</td>
<td></td>
</tr>
<tr>
<td>37 SP d/dt 50-100Hz Mean</td>
<td>38 SP d/dt 5-100Hz Std. Dev.</td>
<td>38 SP d/dt 5-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>50 SP d/d2 THD</td>
<td>98 WS d/dt 0-1Hz Mean</td>
<td>43 SP d/dt 300-400Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>60 SP d/d2 50-100Hz Mean</td>
<td>31 SP d/dt 1-5Hz Mean</td>
<td>38 SP d/dt 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>3 SP Std. Dev.</td>
<td>72 WS Std. Dev.</td>
<td>31 SP d/dt 1-5Hz Mean</td>
<td></td>
</tr>
<tr>
<td>98 WS d/dt 0-1Hz Mean</td>
<td>8 SP 1-5Hz Mean</td>
<td>43 SP d/dt 200-300Hz Mean</td>
<td></td>
</tr>
<tr>
<td>101 WS d/dt 1-5Hz Std. Dev.</td>
<td>99 WS d/dt 0-1Hz Mean</td>
<td>49 SP d/d2 Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>58 SP d/d2 20-50Hz Mean</td>
<td>11 SP 5-20Hz Std. Dev.</td>
<td>61 SP d/d2 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>31 SP d/dt 1-5Hz Mean</td>
<td>57 SP d/d2 5-20Hz Std. Dev.</td>
<td>11 SP 5-20Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>95 WS d/dt Std. Dev.</td>
<td>126 WS d/d2 5-20Hz Std. Dev.</td>
<td>10 SP 5-20Hz Mean</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Dry Gravel</th>
<th>Wet Gravel</th>
<th>Wet &amp; Dry Gravel</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 SP d/d2 20-50Hz Mean</td>
<td>58 SP d2/d2 20-50Hz Mean</td>
<td>56 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>31 SP d/dt 1-5Hz Mean</td>
<td>56 SP d2/d2 5-20Hz Std. Dev.</td>
<td>59 SP d2/d2 20-50Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>35 SP d/dt 20-50Hz Mean</td>
<td>31 SP d/dt 1-5Hz Mean</td>
<td>35 SP d/dt 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>59 SP d/d2 20-50Hz Std. Dev.</td>
<td>72 WS Std. Dev.</td>
<td>34 SP d/dt 5-20Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>26 SP d/dt</td>
<td>8 SP 1-5Hz Mean</td>
<td>31 SP d/dt 1-5Hz Mean</td>
<td></td>
</tr>
<tr>
<td>55 SP d/d2 1-5Hz Std. Dev.</td>
<td>3 SP Std. Dev.</td>
<td>41 SP d/d2 200-300Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>32 SP d/dt 1-5Hz Std. Dev.</td>
<td>32 SP d/dt 5-20Hz Std. Dev.</td>
<td>36 SP d/d2 20-50Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>8 SP 1-5Hz Mean</td>
<td>98 WS d/dt 0-1Hz Mean</td>
<td>8 SP 1-5Hz Mean</td>
<td></td>
</tr>
<tr>
<td>54 SP d/d2 1-5Hz Mean</td>
<td>9 SP 1-5Hz Std. Dev.</td>
<td>43 SP d/d2 300-400Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>11 SP 5-20Hz Std. Dev.</td>
<td>11 SP 5-20Hz Std. Dev.</td>
<td>3 SP Std. Dev.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Dry Asphalt Surface</th>
<th>Wet Asphalt Surface</th>
<th>Wet &amp; Dry Asphalt Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>127 WS d/d2 20-50Hz Mean</td>
<td>58 SP d/d2 20-50Hz Mean</td>
<td>58 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>3 SP Std. Dev.</td>
<td>36 SP d/dt 20-50Hz Std. Dev.</td>
<td>36 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>58 SP d/d2 20-50Hz Mean</td>
<td>38 SP d/dt 50-100Hz Std. Dev.</td>
<td>38 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>60 SP d/d2 50-100Hz Mean</td>
<td>72 WS Std. Dev.</td>
<td>37 SP d/dt 50-100Hz Mean</td>
<td></td>
</tr>
<tr>
<td>104 WS d/dt 20-50Hz Mean</td>
<td>33 SP d/dt 5-20Hz Mean</td>
<td>36 SP d/d2 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>101 WS d/dt 1-5Hz Mean</td>
<td>36 SP d/d2 100-200Hz Std. Dev.</td>
<td>61 SP d/d2 50-100Hz Std. Dev.</td>
<td></td>
</tr>
<tr>
<td>95 WS d/dt Std. Dev.</td>
<td>100 WS d/dt 1-5Hz Mean</td>
<td>35 SP d/d2 20-50Hz Mean</td>
<td></td>
</tr>
<tr>
<td>37 SP d/dt 50-100Hz Mean</td>
<td>59 SP d/d2 20-50Hz Std. Dev.</td>
<td>82 WS 20-50Hz Std. Dev.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Wet &amp; Dry Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>56 SP d2/d2 5-20Hz Mean</td>
<td>36 SP d/dt 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>57 SP d/d2 5-20Hz Mean</td>
<td>43 SP d/dt 300-400Hz Std. Dev.</td>
</tr>
<tr>
<td>9 SP 1-5Hz Mean</td>
<td>13 SP 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>90 WS 300-400Hz Std. Dev.</td>
<td>66 SP d/d2 300-400Hz Mean</td>
</tr>
<tr>
<td>37 SP d/dt 50-100Hz Mean</td>
<td>59 SP d/d2 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>92 WS 400-500Hz Std. Dev.</td>
<td>84 WS 50-100Hz Std. Dev.</td>
</tr>
<tr>
<td>35 SP d/dt 20-50Hz Mean</td>
<td>9 SP 1-5Hz Mean</td>
</tr>
<tr>
<td>81 WS 20-50Hz Mean</td>
<td>79 WS 5-20Hz Mean</td>
</tr>
<tr>
<td>82 WS 20-50Hz Std. Dev.</td>
<td>10 SP 5-20Hz Mean</td>
</tr>
<tr>
<td>49 SP d/d2 Std. Dev.</td>
<td>43 SP d/d2 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>35 SP d/dt 20-50Hz Mean</td>
<td>82 WS 20-50Hz Std. Dev.</td>
</tr>
</tbody>
</table>
The final list of input parameters used in following investigation is given in Table 5.3. Most of this list was complied by integrating the two independent results together. Some parameters such as THD (which had a comparably low importance) were excluded based on a desire to reduce computational complexity. Wheel speed parameters over 100Hz were also excluded based on the observation that since the sensor only logged at 100Hz no notable PSD data should exist above this value.

Table 5.2: Importance rank for all inputs, using single-output models

<table>
<thead>
<tr>
<th>Rank</th>
<th>Input No.</th>
<th>Input Name</th>
<th>Input Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58</td>
<td>SP d2/dL2 20-50Hz Mean</td>
<td>SP d2/dL2 0-1Hz Mean</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>SP d/dt 50-100Hz Std. Dev.</td>
<td>WS 400-500Hz Std. Dev.</td>
</tr>
<tr>
<td>3</td>
<td>59</td>
<td>SP d2/dL2 20-50Hz Std. Dev.</td>
<td>SP d2/dL2 Std. Dev.</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>SP d/dt 50-100Hz Std. Dev.</td>
<td>SP d2/dL2 Std. Dev.</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>SP d/dt 20-50Hz Std. Dev.</td>
<td>SP d2/dL2 Std. Dev.</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>SP d2/dL2 5-20Hz Mean</td>
<td>SP d/dt 1-5Hz Std. Dev.</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>SP d2/dL2 5-20Hz Std. Dev.</td>
<td>WS d/dt 1-5Hz Std. Dev.</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>SP d/dt 20-50Hz Mean</td>
<td>SP d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>9</td>
<td>31</td>
<td>SP d/dt 1-5Hz Mean</td>
<td>SP d/dt Mean</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>SP 5-20Hz Std. Dev.</td>
<td>SP 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>11</td>
<td>55</td>
<td>SP d2/dL2 1-5Hz Std. Dev.</td>
<td>SP THD</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>SP Std. Dev.</td>
<td>SP THD</td>
</tr>
<tr>
<td>13</td>
<td>10</td>
<td>SP 5-20Hz Mean</td>
<td>SP THD</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>SP d/dt 5-20Hz Std. Dev.</td>
<td>SP THD</td>
</tr>
<tr>
<td>15</td>
<td>32</td>
<td>SP d/dt 1-5Hz Std. Dev.</td>
<td>SP THD</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>SP 1-5Hz Mean</td>
<td>SP 0-1Hz Mean</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>SP 1-5Hz Std. Dev.</td>
<td>SP Mean</td>
</tr>
<tr>
<td>18</td>
<td>60</td>
<td>SP d2/dL2 50-100Hz Mean</td>
<td>SP 0-1Hz Std. Dev.</td>
</tr>
<tr>
<td>19</td>
<td>35</td>
<td>SP d/dt 20-50Hz Mean</td>
<td>SP 1-5Hz Std. Dev.</td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>SP 1-5Hz Std. Dev.</td>
<td>SP Mean</td>
</tr>
<tr>
<td>21</td>
<td>46</td>
<td>SP d/dt 1-5Hz Std. Dev.</td>
<td>SP Value</td>
</tr>
<tr>
<td>22</td>
<td>61</td>
<td>SP d2/dL2 50-100Hz Std. Dev.</td>
<td>SP Value</td>
</tr>
<tr>
<td>23</td>
<td>33</td>
<td>SP d/dt 20-50Hz Mean</td>
<td>SP Value</td>
</tr>
<tr>
<td>24</td>
<td>30</td>
<td>SP d/dt 1-5Hz Std. Dev.</td>
<td>SP Value</td>
</tr>
<tr>
<td>25</td>
<td>14</td>
<td>SP d/dt 5-20Hz Std. Dev.</td>
<td>SP Value</td>
</tr>
<tr>
<td>26</td>
<td>72</td>
<td>WS 20-50Hz Std. Dev.</td>
<td>WS 20-50Hz Std. Dev.</td>
</tr>
<tr>
<td>27</td>
<td>26</td>
<td>WS 5-20Hz Mean</td>
<td>WS 5-20Hz Mean</td>
</tr>
<tr>
<td>28</td>
<td>80</td>
<td>WS 5-20Hz Std. Dev.</td>
<td>WS 5-20Hz Std. Dev.</td>
</tr>
<tr>
<td>29</td>
<td>81</td>
<td>WS 20-50Hz Mean</td>
<td>WS 5-20Hz Mean</td>
</tr>
<tr>
<td>30</td>
<td>45</td>
<td>WS 400-500Hz Std. Dev.</td>
<td>WS 0-1Hz Value</td>
</tr>
<tr>
<td>31</td>
<td>105</td>
<td>WS 5-20Hz Std. Dev.</td>
<td>WS 5-20Hz Std. Dev.</td>
</tr>
<tr>
<td>32</td>
<td>69</td>
<td>SP d2/dL2 400-500Hz Mean</td>
<td>WS 400-500Hz Mean</td>
</tr>
<tr>
<td>33</td>
<td>77</td>
<td>WS 1-5Hz Mean</td>
<td>WS 1-5Hz Mean</td>
</tr>
<tr>
<td>34</td>
<td>42</td>
<td>SP d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>35</td>
<td>29</td>
<td>SP d/dt 0-1Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>36</td>
<td>64</td>
<td>SP d2/dL2 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>37</td>
<td>43</td>
<td>SP d/dt 300-400Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>38</td>
<td>78</td>
<td>WS 1-5Hz Std. Dev.</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>39</td>
<td>40</td>
<td>SP d/dt 100-200Hz Std. Dev.</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>40</td>
<td>84</td>
<td>WS 50-100Hz Std. Dev.</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>41</td>
<td>85</td>
<td>SP d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>42</td>
<td>63</td>
<td>SP d2/dL2 100-200Hz Std. Dev.</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>43</td>
<td>66</td>
<td>SP d2/dL2 300-400Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>44</td>
<td>51</td>
<td>SP d2/dL2 THD+Noise</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>45</td>
<td>52</td>
<td>SP d2/dL2 THD+Noise</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>46</td>
<td>50</td>
<td>SP d2/dL2 THD+Noise</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>47</td>
<td>91</td>
<td>WS d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>48</td>
<td>92</td>
<td>WS d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>49</td>
<td>93</td>
<td>WS d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
<tr>
<td>50</td>
<td>94</td>
<td>WS d/dt 200-300Hz Mean</td>
<td>WS d/dt 200-300Hz Mean</td>
</tr>
</tbody>
</table>
The new 67 input, 5 output model was tested for accuracy, with limited improvements found. It was therefore considered necessary to focus the investigation on improving the accuracy of this type of model, rather than in continuing conceptual advances. As such, a thorough analysis of the properties of the model was conducted to see how accuracy was affected by ANN architecture and PSD time interval length, as well as further testing the input importances. It was considered important to test these effects on the data used for ANN training, as well as for completely new testing data. These results are given in Figure 5.20 to Figure 5.25.

Many observations can be taken from these results, but there are four notable ones. Firstly, the level of error between training and testing data is very different, and indicates a level of overfitting that will be discussed later. Secondly, the ANN architecture has a clear effect on model accuracy, although it is difficult to determine this relationship between training and testing data. Thirdly, increasing PSD time had a positive effect on model accuracy for both training and testing. Finally, most of the inputs seem to have similar importance in training and testing, although it is noted that many of the inputs above input number 55 have little effect.
5. Pavement Feature Recognition during Stable Driving Conditions

Figure 5.20: Effect of model architecture for training data (3 second, 67 input, 5 output model)

Figure 5.21: Effect of sample time for training data (67 input, 5 output model with 75x40 architecture)

Figure 5.22: Training input parameter importance (3 second, 67 input, 5 output model with 75x40 architecture)

Figure 5.23: Effect of model architecture for testing data (3 second, 67 input, 5 output model)

Figure 5.24: Effect of sample time for testing data (67 input, 5 output model with 50x0 architecture)

Figure 5.25: Testing input parameter importance (3 second, 67 input, 5 output model with 50x0 architecture)
The accuracy of these models greater than the models designed during the initial investigation, although the different output methods employed make direct comparison difficult. Nonetheless, significant performance advances could be found by testing different models based on these findings. As such, a number of models were developed along the same principles as this one in the hope of obtaining greater accuracy.

5.5.2.4 Input and Architecture Tests
Four more investigations, identical in method to the 67 input model, were carried out in the interests of increasing accuracy. In particular, the dynamics of the proceeding models and their importance analyses were used to progressively test different ideas within the scope of ANN inputs and outputs used for the 138 and 67 input models. The computation required to complete the following FFBP tests runs into over four months of dedicated computer time, and as such represents the limit that could be conducted within the scope of this investigation.

The first model that was attempted increases the number of inputs to 150 by reducing the intervals between PSD mean and standard deviation values. It also places these values into an approximate logarithmic scale in an attempt to alter the way the ANN sees the data, as in Table 5.4. These inputs were utilised in a 150 input 5 output ANN model in a similar way to the 67 input model, with the results shown in Figure 5.26 to Figure 5.31.

| All Values | 1 = SP | 30 = dSP/dt | 69 = d²SP/dt² | 90 = d³SP/dt³ | 109 = d⁴SP/dt⁴ | 138 = d⁵SP/dt⁵ |
| All Std. Dev. | 2 = SP | 31 = dSP/dt | 60 = d²SP/dt² | 90 = d³SP/dt³ | 110 = d⁴SP/dt⁴ | 131 = d⁵SP/dt⁵ |
| 1Hz Values | 3 = SP | 32 = dSP/dt | 61 = d²SP/dt² | 90 = d³SP/dt³ | 111 = d⁴SP/dt⁴ | 132 = d⁵SP/dt⁵ |
| 2Hz Values | 4 = SP | 33 = dSP/dt | 62 = d²SP/dt² | 91 = d³SP/dt³ | 112 = d⁴SP/dt⁴ | 133 = d⁵SP/dt⁵ |
| 3Hz Values | 5 = SP | 34 = dSP/dt | 63 = d²SP/dt² | 92 = d³SP/dt³ | 113 = d⁴SP/dt⁴ | 134 = d⁵SP/dt⁵ |
| 4Hz Values | 6 = SP | 35 = dSP/dt | 64 = d²SP/dt² | 93 = d³SP/dt³ | 114 = d⁴SP/dt⁴ | 135 = d⁵SP/dt⁵ |
| 5Hz Values | 7 = SP | 36 = dSP/dt | 65 = d²SP/dt² | 94 = d³SP/dt³ | 115 = d⁴SP/dt⁴ | 136 = d⁵SP/dt⁵ |
| 6Hz Values | 8 = SP | 37 = dSP/dt | 66 = d²SP/dt² | 95 = d³SP/dt³ | 116 = d⁴SP/dt⁴ | 137 = d⁵SP/dt⁵ |
| 7Hz Values | 9 = SP | 38 = dSP/dt | 67 = d²SP/dt² | 96 = d³SP/dt³ | 117 = d⁴SP/dt⁴ | 138 = d⁵SP/dt⁵ |
| 8Hz Values | 10 = SP | 39 = dSP/dt | 68 = d²SP/dt² | 97 = d³SP/dt³ | 118 = d⁴SP/dt⁴ | 139 = d⁵SP/dt⁵ |
| 9.15Hz Mean | 11 = SP | 40 = dSP/dt | 69 = d²SP/dt² | 98 = d³SP/dt³ | 119 = d⁴SP/dt⁴ | 140 = d⁵SP/dt⁵ |
| 9.15Hz Std. Dev. | 12 = SP | 41 = dSP/dt | 70 = d²SP/dt² | 99 = d³SP/dt³ | 120 = d⁴SP/dt⁴ | 141 = d⁵SP/dt⁵ |
| 15-23Hz Mean | 13 = SP | 42 = dSP/dt | 71 = d²SP/dt² | 100 = d³SP/dt³ | 121 = d⁴SP/dt⁴ | 142 = d⁵SP/dt⁵ |
| 15-23Hz Std. Dev. | 14 = SP | 43 = dSP/dt | 72 = d²SP/dt² | 101 = d³SP/dt³ | 122 = d⁴SP/dt⁴ | 143 = d⁵SP/dt⁵ |
| 23-30Hz Mean | 15 = SP | 44 = dSP/dt | 73 = d²SP/dt² | 102 = d³SP/dt³ | 123 = d⁴SP/dt⁴ | 144 = d⁵SP/dt⁵ |
| 23-30Hz Std. Dev. | 16 = SP | 45 = dSP/dt | 74 = d²SP/dt² | 103 = d³SP/dt³ | 124 = d⁴SP/dt⁴ | 145 = d⁵SP/dt⁵ |
| 36-57Hz Mean | 17 = SP | 46 = dSP/dt | 75 = d²SP/dt² | 104 = d³SP/dt³ | 125 = d⁴SP/dt⁴ | 146 = d⁵SP/dt⁵ |
| 36-57Hz Std. Dev. | 18 = SP | 47 = dSP/dt | 76 = d²SP/dt² | 105 = d³SP/dt³ | 126 = d⁴SP/dt⁴ | 147 = d⁵SP/dt⁵ |
| 57-90Hz Mean | 19 = SP | 48 = dSP/dt | 77 = d²SP/dt² | 106 = d³SP/dt³ | 127 = d⁴SP/dt⁴ | 148 = d⁵SP/dt⁵ |
| 57-90Hz Std. Dev. | 20 = SP | 49 = dSP/dt | 78 = d²SP/dt² | 107 = d³SP/dt³ | 128 = d⁴SP/dt⁴ | 149 = d⁵SP/dt⁵ |
| 90-142Hz Mean | 21 = SP | 50 = dSP/dt | 79 = d²SP/dt² | 108 = d³SP/dt³ | 129 = d⁴SP/dt⁴ | 150 = d⁵SP/dt⁵ |
| 90-142Hz Std. Dev. | 22 = SP | 51 = dSP/dt | 80 = d²SP/dt² | 109 = d³SP/dt³ | 130 = d⁴SP/dt⁴ | 151 = d⁵SP/dt⁵ |
| 142-222Hz Mean | 23 = SP | 52 = dSP/dt | 81 = d²SP/dt² | 110 = d³SP/dt³ | 131 = d⁴SP/dt⁴ | 152 = d⁵SP/dt⁵ |
| 142-222Hz Std. Dev. | 24 = SP | 53 = dSP/dt | 82 = d²SP/dt² | 111 = d³SP/dt³ | 132 = d⁴SP/dt⁴ | 153 = d⁵SP/dt⁵ |
| 224-355Hz Mean | 25 = SP | 54 = dSP/dt | 83 = d²SP/dt² | 112 = d³SP/dt³ | 133 = d⁴SP/dt⁴ | 154 = d⁵SP/dt⁵ |
| 224-355Hz Std. Dev. | 26 = SP | 55 = dSP/dt | 84 = d²SP/dt² | 113 = d³SP/dt³ | 134 = d⁴SP/dt⁴ | 155 = d⁵SP/dt⁵ |
| 355-520Hz Mean | 27 = SP | 56 = dSP/dt | 85 = d²SP/dt² | 114 = d³SP/dt³ | 135 = d⁴SP/dt⁴ | 156 = d⁵SP/dt⁵ |
| 355-520Hz Std. Dev. | 28 = SP | 57 = dSP/dt | 86 = d²SP/dt² | 115 = d³SP/dt³ | 136 = d⁴SP/dt⁴ | 157 = d⁵SP/dt⁵ |
| 520-800Hz Mean | 29 = SP | 58 = dSP/dt | 87 = d²SP/dt² | 116 = d³SP/dt³ | 137 = d⁴SP/dt⁴ | 158 = d⁵SP/dt⁵ |

Table 5.4: PSD characteristic input numbers for 150in model
Figure 5.26: Effect of model architecture for training data (3 second, 150 input, 5 output model)

Figure 5.27: Effect of sample time for training data (150 input, 5 output model with 40x20)

Figure 5.28: Training input parameter importance (3 second, 150 input, 5 output model with 40x20 architecture)

Figure 5.29: Effect of model architecture for testing data (3 second, 150 input, 5 output model)

Figure 5.30: Effect of sample time for testing data (150 input, 5 output model with 40x0)

Figure 5.31: Testing input parameter importance (3 second, 150 input, 5 output model with 40x0 architecture)
The 150 input models provide greater training accuracy than the 67 input models, but the testing error is greater. This highlights increased overfitting, which is undesirable, and in both cases increases of internal neurons make this effect more pronounced. As such, it appears that decreasing ANN size increases model accuracy. In addition, the importance analysis of both training and testing data shows that PSD curves above 100Hz have little effect on the model, and could be removed in further work. The 3 second PSD timescale that has been used to this point produces at least twice the error of using a 5 second interval. As such, accuracy could be improved by using this as a new value.

Moreover, the logarithmic scale used in this case did not have a great effect on accuracy. It was also observed that the use of two derivatives of each of the measured values may be overkill, and work was completed to see whether some of the sets of values could be removed. These observations lead to the selection of new model inputs, as shown in Table 5.5, which clearly abandon suspension speed, suspension acceleration and wheel rate of acceleration used previously. The model size was increased accordingly to obtain more resolution of the PSD curve approximation in an attempt to see if this has an effect.

<table>
<thead>
<tr>
<th>Value</th>
<th>All Mean</th>
<th>1 = SP</th>
<th>44 = WS</th>
<th>67 = dAWS/dt</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>2 = SP</td>
<td>4 = SP</td>
<td>8 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>All</td>
<td>3 = SP</td>
<td>4 = SP</td>
<td>8 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>0.5Hz</td>
<td>4 = SP</td>
<td>4 = SP</td>
<td>5 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>0.5Hz</td>
<td>5 = SP</td>
<td>4 = SP</td>
<td>8 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>1.0Hz</td>
<td>6 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>1.0Hz</td>
<td>7 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>4.5Hz</td>
<td>8 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>4.5Hz</td>
<td>9 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>10Hz</td>
<td>10 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>10Hz</td>
<td>11 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>20Hz</td>
<td>12 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>20Hz</td>
<td>13 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>25Hz</td>
<td>14 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>25Hz</td>
<td>15 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>30Hz</td>
<td>16 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>30Hz</td>
<td>17 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>35Hz</td>
<td>18 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>35Hz</td>
<td>19 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>40Hz</td>
<td>20 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>40Hz</td>
<td>21 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>45Hz</td>
<td>22 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
<tr>
<td>45Hz</td>
<td>23 = SP</td>
<td>4 = SP</td>
<td>9 = SP</td>
<td>1 = dAWS/dt</td>
</tr>
</tbody>
</table>

Table 5.5: PSD characteristic input numbers for 127in model

Previous investigation shows that ANN architectures containing a second hidden layer (second layer neurons > 0) have similar modelling ability when compared to single hidden layer ANNs. As such, further work abandoned investigations of architectures that contained second hidden layers to reduce the level of computation required. The investigation of PSD timescales conducted previously was considered sufficient for this work, with 5 second segments used for the remainder. As such, the testing results for the 129 input ANN are given in Figure 5.32 and Figure 5.33. The training results are not
given because they provide little useful information above what has already been presented.

These results provide a clear increase in accuracy above the 150 input model, but are not quite as good as the 67 input model, although the change in PSD timescale makes direct comparison difficult. Nonetheless, it can be assumed that generally smaller numbers of inputs result in increases in ANN accuracy. As such, further reduction of inputs was considered an important investigation avenue. In this manner, a number of PSD inputs used in the 127 input model investigation were removed, based on the input importance analysis results. This resulted a model that consisted of suspension acceleration and suspension acceleration inputs only, and is shown in Table 5.6. The results of this model are given in Figure 5.34 and Figure 5.35.

<table>
<thead>
<tr>
<th>All</th>
<th>Values</th>
<th>1 = \text{d}^{2}\text{SPD}^{2}</th>
<th>44 = \text{dMS}^{2}\text{Hz}</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Mean</td>
<td>2 = \text{d}^{2}\text{SPD}^{2}</td>
<td>46 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>All</td>
<td>Std. Dev.</td>
<td>3 = \text{d}^{2}\text{SPD}^{2}</td>
<td>46 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>0.5 Hz</td>
<td>Mean</td>
<td>4 = \text{d}^{2}\text{SPD}^{2}</td>
<td>47 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>0.5 Hz</td>
<td>Std. Dev.</td>
<td>5 = \text{d}^{2}\text{SPD}^{2}</td>
<td>48 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>5 Hz</td>
<td>Mean</td>
<td>6 = \text{d}^{2}\text{SPD}^{2}</td>
<td>49 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>5 Hz</td>
<td>Std. Dev.</td>
<td>7 = \text{d}^{2}\text{SPD}^{2}</td>
<td>50 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>10 Hz</td>
<td>Mean</td>
<td>8 = \text{d}^{2}\text{SPD}^{2}</td>
<td>51 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>10 Hz</td>
<td>Std. Dev.</td>
<td>9 = \text{d}^{2}\text{SPD}^{2}</td>
<td>52 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>15 Hz</td>
<td>Mean</td>
<td>10 = \text{d}^{2}\text{SPD}^{2}</td>
<td>53 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>15 Hz</td>
<td>Std. Dev.</td>
<td>11 = \text{d}^{2}\text{SPD}^{2}</td>
<td>54 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>20 Hz</td>
<td>Mean</td>
<td>12 = \text{d}^{2}\text{SPD}^{2}</td>
<td>55 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>20 Hz</td>
<td>Std. Dev.</td>
<td>13 = \text{d}^{2}\text{SPD}^{2}</td>
<td>56 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>25 Hz</td>
<td>Mean</td>
<td>14 = \text{d}^{2}\text{SPD}^{2}</td>
<td>57 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>25 Hz</td>
<td>Std. Dev.</td>
<td>15 = \text{d}^{2}\text{SPD}^{2}</td>
<td>58 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>30 Hz</td>
<td>Mean</td>
<td>16 = \text{d}^{2}\text{SPD}^{2}</td>
<td>59 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>30 Hz</td>
<td>Std. Dev.</td>
<td>17 = \text{d}^{2}\text{SPD}^{2}</td>
<td>60 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>35 Hz</td>
<td>Mean</td>
<td>18 = \text{d}^{2}\text{SPD}^{2}</td>
<td>61 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>35 Hz</td>
<td>Std. Dev.</td>
<td>19 = \text{d}^{2}\text{SPD}^{2}</td>
<td>62 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>40 Hz</td>
<td>Mean</td>
<td>20 = \text{d}^{2}\text{SPD}^{2}</td>
<td>63 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>40 Hz</td>
<td>Std. Dev.</td>
<td>21 = \text{d}^{2}\text{SPD}^{2}</td>
<td>64 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>45 Hz</td>
<td>Mean</td>
<td>22 = \text{d}^{2}\text{SPD}^{2}</td>
<td>65 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>45 Hz</td>
<td>Std. Dev.</td>
<td>23 = \text{d}^{2}\text{SPD}^{2}</td>
<td>66 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>50 Hz</td>
<td>Mean</td>
<td>24 = \text{d}^{2}\text{SPD}^{2}</td>
<td>67 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>50 Hz</td>
<td>Std. Dev.</td>
<td>25 = \text{d}^{2}\text{SPD}^{2}</td>
<td>68 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>55 Hz</td>
<td>Mean</td>
<td>26 = \text{d}^{2}\text{SPD}^{2}</td>
<td>69 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>55 Hz</td>
<td>Std. Dev.</td>
<td>27 = \text{d}^{2}\text{SPD}^{2}</td>
<td>70 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>60 Hz</td>
<td>Mean</td>
<td>28 = \text{d}^{2}\text{SPD}^{2}</td>
<td>71 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>60 Hz</td>
<td>Std. Dev.</td>
<td>29 = \text{d}^{2}\text{SPD}^{2}</td>
<td>72 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>65 Hz</td>
<td>Mean</td>
<td>30 = \text{d}^{2}\text{SPD}^{2}</td>
<td>73 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>65 Hz</td>
<td>Std. Dev.</td>
<td>31 = \text{d}^{2}\text{SPD}^{2}</td>
<td>74 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>70 Hz</td>
<td>Mean</td>
<td>32 = \text{d}^{2}\text{SPD}^{2}</td>
<td>75 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>70 Hz</td>
<td>Std. Dev.</td>
<td>33 = \text{d}^{2}\text{SPD}^{2}</td>
<td>76 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>75 Hz</td>
<td>Mean</td>
<td>34 = \text{d}^{2}\text{SPD}^{2}</td>
<td>77 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>75 Hz</td>
<td>Std. Dev.</td>
<td>35 = \text{d}^{2}\text{SPD}^{2}</td>
<td>78 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>80 Hz</td>
<td>Mean</td>
<td>36 = \text{d}^{2}\text{SPD}^{2}</td>
<td>79 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>80 Hz</td>
<td>Std. Dev.</td>
<td>37 = \text{d}^{2}\text{SPD}^{2}</td>
<td>80 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>85 Hz</td>
<td>Mean</td>
<td>38 = \text{d}^{2}\text{SPD}^{2}</td>
<td>81 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>85 Hz</td>
<td>Std. Dev.</td>
<td>39 = \text{d}^{2}\text{SPD}^{2}</td>
<td>82 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>90 Hz</td>
<td>Mean</td>
<td>40 = \text{d}^{2}\text{SPD}^{2}</td>
<td>83 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>90 Hz</td>
<td>Std. Dev.</td>
<td>41 = \text{d}^{2}\text{SPD}^{2}</td>
<td>84 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>95 Hz</td>
<td>Mean</td>
<td>42 = \text{d}^{2}\text{SPD}^{2}</td>
<td>85 = \text{dMS}^{2}\text{Hz}</td>
</tr>
<tr>
<td>95 Hz</td>
<td>Std. Dev.</td>
<td>43 = \text{d}^{2}\text{SPD}^{2}</td>
<td>86 = \text{dMS}^{2}\text{Hz}</td>
</tr>
</tbody>
</table>

Table 5.6: PSD characteristic input numbers for 86in model
The accuracy of this model is similar to the 150 input model, and can be considered low. Furthermore, the importance analysis shows that the wheel acceleration inputs have little effect on the result, which is in contrast to all previous investigation. This is an interesting result, because previous investigation has shown that relying on suspension data alone for the surface prediction should result in decreased accuracy. It is possible, however, that these parameters are partially interrelated, and that the size of the ANN has impaired its ability. As such, another model was produced that decreased the number of inputs even further by noting the input importances of the 129 input model. This model is shown in Table 5.7, and the results in Figure 5.36 and Figure 5.37.

<table>
<thead>
<tr>
<th>All</th>
<th>Std. Dev.</th>
<th>1$ = \delta^2$</th>
<th>23$ = \delta W_\text{Dry}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5Hz</td>
<td>Mean</td>
<td>2$ = \delta^2$</td>
<td>24$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>10.05Hz</td>
<td>Mean</td>
<td>3$ = \delta^2$</td>
<td>25$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>15.0Hz</td>
<td>Mean</td>
<td>5$ = \delta^2$</td>
<td>26$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>15.0Hz</td>
<td>Std. Dev.</td>
<td>6$ = \delta^2$</td>
<td>27$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>20.25Hz</td>
<td>Mean</td>
<td>7$ = \delta^2$</td>
<td>28$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>20.25Hz</td>
<td>Std. Dev.</td>
<td>8$ = \delta^2$</td>
<td>29$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>25.5Hz</td>
<td>Mean</td>
<td>9$ = \delta^2$</td>
<td>30$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>25.5Hz</td>
<td>Std. Dev.</td>
<td>10$ = \delta^2$</td>
<td>31$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>30.5Hz</td>
<td>Mean</td>
<td>11$ = \delta^2$</td>
<td>32$ = W_\text{Dry}$</td>
</tr>
<tr>
<td>30.5Hz</td>
<td>Std. Dev.</td>
<td>12$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>35.4Hz</td>
<td>Mean</td>
<td>13$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>35.4Hz</td>
<td>Std. Dev.</td>
<td>14$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>40.4Hz</td>
<td>Mean</td>
<td>15$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>40.4Hz</td>
<td>Std. Dev.</td>
<td>16$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>45.5Hz</td>
<td>Mean</td>
<td>17$ = \delta^2$</td>
<td></td>
</tr>
<tr>
<td>45.5Hz</td>
<td>Std. Dev.</td>
<td>18$ = \delta^2$</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7: PSD characteristic input numbers for 44in model
The smaller model produces less error than the 86 input model, but more than the 129 input one. This suggests that the small size of the ANN has increased its generalisation abilities to avoid overfitting, but the inputs it uses are not adequate to fully represent the important vibration processes. As such, more work is needed to determine a suitable balance between appropriate model inputs and reduced ANN size. This is beyond the scope of this work however, and represents a very long investigation.

Instead, it was observed that from the six different models investigated here, the most accurate one could be selected for further research. However, each of the models are not directly comparable because of differences in training and testing files, and in PSD time scales. This was not ideal, and to make a fair comparison it was necessary to retrain each of them under similar conditions. The logged data was used to construct new training and testing files that were all similar sizes, and all at the 5 second PSD timescale. The ANN models were trained with identical numbers of training iterations (2 million) and ANN constants. The best results for each of the models are given in the Table 5.8 summary.

![Figure 5.36: Effect of model architecture for testing data (5 second, 44 input, 5 output model)](image1)

![Figure 5.37: Testing input parameter importance (5 second, 44 input, 5 output model with 30x0 architecture)](image2)

<table>
<thead>
<tr>
<th>No. Inputs</th>
<th>No. 1st Layer</th>
<th>No. 2nd Layer</th>
<th>Smooth Asphalt Error (%)</th>
<th>Rough Asphalt Error (%)</th>
<th>Cement Error (%)</th>
<th>Unsealed Error (%)</th>
<th>Wet Error (%)</th>
<th>Total Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>138</td>
<td>30</td>
<td>0</td>
<td>38.4</td>
<td>38.7</td>
<td>34.9</td>
<td>38.6</td>
<td>21.2</td>
<td>35.0</td>
</tr>
<tr>
<td>67</td>
<td>50</td>
<td>0</td>
<td>37.2</td>
<td>36.6</td>
<td>41.4</td>
<td>30.3</td>
<td>21.6</td>
<td>34.1</td>
</tr>
<tr>
<td>150</td>
<td>40</td>
<td>0</td>
<td>33.8</td>
<td>45.5</td>
<td>39.4</td>
<td>34.8</td>
<td>26.7</td>
<td>36.6</td>
</tr>
<tr>
<td>129</td>
<td>40</td>
<td>0</td>
<td>36.2</td>
<td>40.1</td>
<td>40.1</td>
<td>32.5</td>
<td>5.3</td>
<td>33.5</td>
</tr>
<tr>
<td>86</td>
<td>50</td>
<td>0</td>
<td>35.6</td>
<td>42.2</td>
<td>45.1</td>
<td>31.8</td>
<td>14.5</td>
<td>35.5</td>
</tr>
<tr>
<td>44</td>
<td>30</td>
<td>0</td>
<td>35.1</td>
<td>41.6</td>
<td>43.7</td>
<td>30.7</td>
<td>14.1</td>
<td>34.7</td>
</tr>
</tbody>
</table>

Table 5.8: Best surface prediction results for 5 output, 5 second models
The 129 input model was found to have the best accuracy overall, mostly due to its ability to predict whether the road was wet or dry with only 5.3% error. The lowest error rates for each of the inputs was also recorded as 33.8% for smooth asphalt; 36.6% for rough asphalt; 34.9% for cement and; 30.3% for unsealed.

5.5.2.5 Reduced Number of Outputs

It was noted through the course of this investigation much of the error arose from the need to differentiate between asphalt and cement. The “smooth asphalt”, “rough asphalt” and “cement” tests are each difficult for the model to determine, and have limited application to stability controllers. Furthermore, the presence of these outputs increased the size of the ANN models, which was shown to have a negative effect. Instead, it was proposed to replace these three outputs with a single “rough or smooth rough” test. In combination with the “sealed or unsealed” and “wet or dry” tests this provides the capacity to predict all of the cases with the five output model, except for being able to differentiate between cement and asphalt surfaces. As such, cement, rough asphalt and gravel are characterised as rough, and smooth asphalt as smooth.

An investigation was carried out for the six models developed previously, by replacing the five outputs with these three outputs. This investigation procedure was identical to the five output case, with different neural architectures used to determine the best arrangement and input importances. The results of this investigation are given in Table 5.9.

<table>
<thead>
<tr>
<th>No. Inputs</th>
<th>No. 1st Layer</th>
<th>No. 2nd Layer</th>
<th>Rough Test Error (%)</th>
<th>Sealed Test Error (%)</th>
<th>Wet Test Error (%)</th>
<th>Total Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>138</td>
<td>60</td>
<td>0</td>
<td>38.9</td>
<td>35.9</td>
<td>24.4</td>
<td>33.6</td>
</tr>
<tr>
<td>67</td>
<td>20</td>
<td>0</td>
<td>37.5</td>
<td>27.3</td>
<td>21.6</td>
<td>29.5</td>
</tr>
<tr>
<td>150</td>
<td>30</td>
<td>0</td>
<td>34.4</td>
<td>32.9</td>
<td>23.2</td>
<td>30.6</td>
</tr>
<tr>
<td>129</td>
<td>30</td>
<td>0</td>
<td>36.7</td>
<td>29.6</td>
<td>6.9</td>
<td>27.5</td>
</tr>
<tr>
<td>86</td>
<td>40</td>
<td>0</td>
<td>36.4</td>
<td>32.8</td>
<td>12.9</td>
<td>29.2</td>
</tr>
<tr>
<td>44</td>
<td>30</td>
<td>0</td>
<td>36.3</td>
<td>28.2</td>
<td>3.5</td>
<td>26.6</td>
</tr>
</tbody>
</table>

*Table 5.9: Best surface prediction results for 3 output, 5 second models*

In this case, the 44 input model was shown to have the best performance, followed by the 126 input model. The best predictions, therefore, contain errors of 36.3% for the smooth/rough test, 28.2% for the sealed/unsealed test and 3.5% for the wet/dry test. This can be directly compared to the prediction error of the same features within the best 5 output model of 36.2%, 32.5% and 5.3% respectively, which produces a total RMS error
of 28.3%. As such, the 3 output model performs 1.6% better than the 5 output model, and the 44 input 3 output model can be considered to provide the best FFBP prediction of road condition.

5.5.2.6 OLL Model
Each of the ANN models presented above use the Feed Forward Back Propagation approach, which is robust but slow to train. Optimised Layer by Layer ANNs are a newer technology, and have offered considerable savings in training time and reduced error in previous applications. It was stated earlier that OLL investigation would form part of this research.

In practice, it was found that the advantages of OLL could only be obtained using relatively simple models. The large models and large training sets used in this surface identification investigation were found to significantly slow the OLL training to a point where they could not be completed. For example, the 44 input x 30 first layer x 0 second layer x 3 output model that was highlighted as the best performing FFBP model took approximately 24 hours to train using a standard PC. The identical model using an OLL algorithm had very little success after 4 weeks of training.

The reason for this limitation seems clear. The OLL algorithm relies on calculating large matrices to determine weight updates, while the FFBP performs many small calculations in series. In a relatively small ANN model, the training time for the OLL is much shorter because the matrices provide fast mathematical convergence with reduced computation. However, as ANN models grow in size the OLL computations increase at a faster rate than the FFBP ones, and consume vast qualities of memory. For instance, the OLL optimal hidden layer matrix $\tilde{A}$ has a defined size of (no. inputs + 1)*(no. hidden layer) square. In the 44x30x3 single hidden layer model this equates to 1.7 million elements. Furthermore, each of these elements must be calculated by summing the output responses of the ANN to the entire training data set. With a typical training set of over 40,000 patterns, this equates to approximately a trillion array seeks, 840 billion summations and multiplications, and 210 billion array updates to determine each matrix $\tilde{A}$ iteration. Similar scales of calculation are also required for other OLL functions, with the array seeking and updating functions and the possible use of ‘virtual memory’ considered as the mostly likely bottlenecks. As such, no OLL models are provided here.
Output Analysis

The 44 input x 30 first layer x 0 second layer x 3 output FFBP ANN model was found to produce the best accuracy for surface identification within the scope of this research. To further evaluate the performance of the model, it is useful to observe its ability to predict road surface features as a time history. Figure 5.38, Figure 5.39 and Figure 5.40 provide this information, with the ANN model used to predict road surface features employing logged data that is wholly separate from the training set. It should be noted that the transition from one surface characteristic to another is the result of combining data, rather than as a transition observed while driving.

The RMS errors mirror the results in Table 5.9. However, the average error is much less than these values, at 0.3%, 9.9% and 14.1% for the wet, sealed and rough tests respectively. Scatter in the results is also evident, particularly in the sealed and rough tests. The scatter is a product of the analogue nature of the ANN outputs, whereas the surfaces tests that need to be completed are binary. In this respect, it is useful to convert the analogue signals into binary ones to observe the absolute prediction error. At its simplest, this conversion could be made so that values below 0.5 are false, and values above 0.5 are true. However, it is also observed that values near 0.5 represent a condition of little confidence in the predicted output, while values near 0 or 1 indicate high confidence. This can be used to incorporate confidence limits into the analogue to binary conversion, and hence increase prediction accuracy. In this way, predictions that do not meet the desired confidence levels can be excluded from the results.

Figure 5.41, Figure 5.42 and Figure 5.43 show the results of filtering using this method. In particular, upper and lower confidence limits are chosen for each output to minimise RMS error. Predicted values above the upper limit are considered true, while values below the lower limit are considered false. All values that exist between the upper and lower limits are excluded from the predictions, with the previous filtered prediction assumed to be constant in these cases. Each of the figures illustrate the confidence limits which are used, the percentage of predictions that were outside these confidence bounds, and the resulting RMS and average errors. The average errors were found to be 0.0%, 6.4% and 13.2% for the wet, sealed and rough tests respectively.
5. Pavement Feature Recognition during Stable Driving Conditions -215-

Figure 5.38: Unfiltered prediction of “Wet” output (44 input, 30x0, 3 output model)

Figure 5.39: Unfiltered prediction of “Sealed” output (44 input, 30x0, 3 output model)

Figure 5.40: Unfiltered prediction of “Rough” output (44 input, 30x0, 3 output model)

Figure 5.41: Upper and lower limit filtered prediction of “Wet” output (44 input, 30x0, 3 output model)

Figure 5.42: Upper and lower limit filtered prediction of “Sealed” output (44 input, 30x0, 3 output model)

Figure 5.43: Upper and lower limit filtered prediction of “Rough” output (44 input, 30x0, 3 output model)
Due to the binary nature of these outputs, the average error can be considered to be the frequency of error. In this light, the model is capable of correct prediction of wet or dry 100% of the time, sealed or unsealed 93.6% of the time, and rough or smooth 86.8% of the time.

5.6 Remarks

The investigation brings together two ideas presented in literature to produce a new road surface feature identification technology. Suspension position and wheel speed vibrations are used within ANN models to predict important pavement characteristics under stable driving conditions. This capability is new on three fronts. Firstly, literature considers road surface identification in stable conditions to be impossible using traditional techniques. This investigation shows that it can be done. Secondly, this research brings together the two separate literature topics of suspension position vibration and wheel speed vibration for road-limited road feature identification. These are integrated to provide more predictive functionality. Thirdly, the investigation is carried out on a real vehicle, and in real conditions. This is relatively unusual in automotive ANN research, and provides a useful foundation for the results. The development and use of ANN control programs within LabVIEW is considered a new application for the software, and represents a significant proportion of the investigation effort.

The final outcomes of 100% predictive accuracy of whether a road is wet or dry, 94% of whether a road is sealed or unsealed and 87% of whether a road is rough or smooth provide promising scope for future work. Particular emphasis should be placed on reducing RMS error in the future, which appears to be the result of “overfitting” of the ANN models due to their size. It is assumed that a significant part of the error is the result of the complex tyre elastic properties within the linear region, which is the literature-cited reason why surface prediction in stable conditions is not considered possible. It is expected that further investigation will enhance these research outcomes on many fronts.